

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

D'ALFONSO, SIMON

Title:

Towards a framework for semantic information

Date:

2012

Citation:

D'Alfonso, S. (2012). Towards a framework for semantic information. PhD thesis, Dept. of Philosophy, School of Historical and Philosophical Studies, The University of Melbourne.

Persistent Link:

<https://hdl.handle.net/11343/37818>

Terms and Conditions:

Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.

Towards a Framework for Semantic Information

Simon D'Alfonso

Submitted in total fulfilment of the requirements of the degree of Doctor of
Philosophy
Department of Philosophy
The University of Melbourne

September 2012

Produced on archival quality paper

Contents

Abstract	ix
Declaration	xi
Preface	xiii
Acknowledgements	xv
1 An Introductory Overview of Information	1
1.1 What is Information?	2
1.1.1 Data	4
1.1.2 The Mathematical Theory of Communication	8
1.1.3 Moving Beyond Data	16
1.2 Philosophy and Information	18
1.2.1 The Philosophy of Information	19
1.3 Semantic Information and Environmental Information	20
1.3.1 Gricean Meaning	20
1.3.2 Semantic Information	21

1.3.3	Environmental Information	27
1.4	The Alethic Nature of Semantic Information	30
1.5	Conclusion	44
2	Quantifying Semantic Information	45
2.1	Bar-Hillel and Carnap's Theory of Semantic Information	47
2.1.1	Some Comments on the Theory of Classical Semantic Information	50
2.1.2	The Bar-Hillel-Carnap Paradox and Paraconsistent Logic	51
2.2	CSI and Truth	54
2.2.1	Epistemic Utility	54
2.2.2	CSI, Truth and Scoring Rules	55
2.3	Floridi's Theory of Strongly Semantic Information	56
2.3.1	Some Comments on Floridi's Theory	61
2.4	Information Quantification via Truthlikeness	62
2.4.1	The Basic Feature Approach to Truthlikeness	63
2.4.2	The Tichy/Oddie Approach to Truthlikeness	65
2.4.3	Truthlikeness Adequacy Conditions and Information Conditions	78
2.4.4	Niiniluoto on Truthlikeness	79
2.4.5	An Interpretation of the Tichy/Oddie Measure	80
2.4.6	Adjusting State Description Utility Values	87
2.4.7	Another Interpretation of the Tichy/Oddie Measure	88
2.5	Another Method to Quantify Semantic Information	89

2.5.1	Adequacy Conditions	91
2.5.2	Adjusting Utilities	96
2.5.3	Misinformation	97
2.6	Estimated Information	99
2.7	Combining Measures	102
2.8	Conclusion	105
3	Agent-Relative Informativeness	107
3.1	AGM Belief Change	108
3.2	Combining Information Measurement and Belief Revision	113
3.2.1	True Database Content and True Input	113
3.2.2	False Database Content and True Input	115
3.2.3	True Inputs Guaranteed to Increase Information Yield	116
3.3	Agent-Relative Informativeness	125
3.3.1	Adding Information to Information	126
3.3.2	Adding Information to Information/Misinformation	128
3.3.3	Adding Misinformation/Information to Misinformation/Information	131
3.4	Dealing with Uncertainty and Conflicting Input	133
3.4.1	Paraconsistent Approaches	137
3.5	Applications	152
3.5.1	Extension to Other Spaces	152
3.5.2	The Value of Information	153

3.5.3	Lottery-Style Scenarios	159
3.5.4	The Conjunction Fallacy	163
3.5.5	Quantifying Epistemic and Doxastic Content	164
3.5.6	Agent-Oriented Relevance	165
3.6	Conclusion	170
4	Environmental Information and Information Flow	173
4.1	Dretske on Information	174
4.1.1	Dretske's Account and Properties of Information Flow	186
4.2	Probabilistic Information	188
4.3	A Counterfactual Theory of Information	194
4.3.1	The Logic of Counterfactuals and Information Flow Properties	201
4.3.2	Transitivity, Monotonicity and Contraposition	203
4.4	A Modal Logical Account of Information Flow	208
4.4.1	Variable Information Flow Contexts	211
4.4.2	Agent-Relative Information Flow	218
4.4.3	Information Closure	219
4.4.4	Variable Relevant Alternatives and the Other Information Flow Prop- erties	243
4.5	Another Requirement on Semantic Information?	251
4.6	Conclusion	254
5	Information and Knowledge	257

5.1	Some Background	257
5.2	Dretske on Knowledge	261
5.2.1	Testing and Supplementing Dretske's Informational Epistemology . .	266
5.2.2	How reliable does an information source have to be?	284
5.2.3	Dealing with Knowledge of Necessary Truths	293
5.2.4	Knowledge, Information and Testimony	296
5.3	Epistemic Logics for Informational Epistemologies	302
5.3.1	Alternative Sets of Relevant Alternatives and Multi-Modal Logic . . .	303
5.3.2	'Knows that' as a semi-penetrating operator	306
5.3.3	One Approach to Epistemic Logic and Relevant Alternatives	306
5.3.4	Going Non-normal	309
5.3.5	A Logic for Dretskean Epistemology	315
5.4	The Value of Information and Knowledge	318
5.4.1	Knowledge and True Belief Generation	320
5.4.2	The Value of Knowledge	326
5.5	Conclusion	339
A	Quantifying Semantic Information	345
A.1	Adequacy Condition Proofs for the Value Aggregate Method	347
A.2	Translation Invariance	350
A.3	Formula-Based Approaches	353
	Appendices	345

B Agent-Relative Informativeness	357
B.1 Combining Information Measurement and Belief Revision	357
B.1.1 True Content and False Input	357
B.1.2 False Content and False Input	357
B.2 True Inputs Guaranteed to Increase Information Yield	358
B.3 Paraconsistent Approaches	361
 C Environmental Information and Information Flow	 367
C.1 Probabilistic Information	367
C.2 The Arrow of Information Flow	368

Abstract

This thesis addresses some important questions regarding an account of semantic information. Starting with the contention that semantic information is to be understood as truthful meaningful data, several key elements for an account of semantic information are developed. After an introductory overview of information, the thesis is developed over four chapters. ‘Quantifying Semantic Information’ looks at the quantification of semantic information as represented in terms of propositional logic. The main objective is to investigate how traditional inverse probabilistic approaches to quantifying semantic information can be replaced with approaches based on the notion of truthlikeness. In ‘Agent-Relative Informativeness’ the results of the previous chapter are combined with belief revision in order to construct a formal framework in which to, amongst other things, measure agent-relative informativeness; how informative some piece of information is relative to a given agent. ‘Environmental Information and Information Flow’ analyses several existing accounts of environmental information and information flow before using this investigation to develop a better account of and explicate these notions. Finally, ‘Information and Knowledge’ contributes towards the case for an informational epistemology, based on Fred Dretske’s information-theoretic account of knowledge.

Declaration

This is to certify that

- i the thesis comprises only my original work towards the PhD except where indicated in the Preface,
- ii due acknowledgement has been made in the text to all other material used,
- iii the thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Simon D'Alfonso

Preface

This PhD thesis is the result of research conducted over the last three or so years. Whilst searching for a thesis topic, I became interested in philosophical work on information after coming across some of Luciano Floridi's work in the philosophy of information. This discovery led me to several other pieces of literature in the field that served as a starting point for my research. A notable mention goes to the introductory text *Information and Information Flow* [22], which introduced me to and provided an accessible overview of the areas that came to form my research agenda.

Whilst 'information' is a term that everyone is familiar with, the notion of information is one that I had not thought much about. Struck by the richness of the simple question 'what is information?', my investigation was initiated by the curiosity that it raised. The novelty of and my interest in this general question has remained a driving factor in my research.

The notion of information presents a vast conceptual labyrinth and there is a plethora of research avenues one could take in researching it. My background and preliminary readings resulted in my focus on semantic conceptions of information and the investigation carried out in this thesis. Thus the aim of this thesis is basically to establish a definition of semantic information, address some questions regarding semantic information that have been raised and develop several key elements for an account of semantic information.

Not too long after Claude Shannon introduced in the middle of the twentieth century what is known as the mathematical theory of communication or information theory, philosophers started taking a serious interest in information. By the end of the first decade of the twenty-first century a substantial body of work on the philosophy of information has accumulated. This thesis makes a modest, yet I hope worthwhile, contribution to this field of research by expanding upon and adding to this body of work.

After an introductory overview of information, this thesis is developed over four chapters. 'Quantifying Semantic Information' investigates ways to quantitatively measure semantic information as represented in terms of propositional logic. In 'Agent-Relative Informativeness'

the results of the previous chapter are combined with belief revision to construct a formal account of measuring how informative some piece of information is to a given agent. ‘Environmental Information and Information Flow’ analyses several existing accounts of environmental information and information flow before using this investigation to develop a better account of and explicate these notions. Finally, with contributions from some of the previous chapter’s results, ‘Information and Knowledge’ contributes towards the case for an informational epistemology.

Beyond their relevance to the philosophy of information, some of the results in this thesis will be of particular relevance to and potentially find applicability in other areas. Two examples worth mentioning are the chapter on information and knowledge, which provides responses to some general questions in epistemology and the chapter on agent-relative informativeness, which deals with topics that overlap with formal epistemology and belief/database revision. With regard to the latter, the link between recent literature on theory change and my formal account of agent-relative informativeness, in that they both look at combining truthlikeness with belief revision, is a good example of convergent evolution in research; similar problems leading to the development of common outcomes and solutions.

Acknowledgements

I would firstly like to thank my supervisor Greg Restall for his support throughout my candidature. It was in the University of Melbourne philosophy department where I ‘cut my philosophical teeth’, starting as an undergraduate in the year 2000. Thus to Greg and the other members of the department who have contributed to my philosophical upbringing I would also like to offer my thanks. The environment provided by the department over the years has significantly influenced the approaches I take and tools that I apply to my research in philosophy.

Secondly, I would like to thank those members of the international philosophy community whose work on information I have used as a starting point and benefited from. In particular Luciano Floridi for his work in establishing the philosophy of information field and Fred Dretske for his innovative and influential work on information and knowledge. Furthermore, I would like to thank those members of the philosophy of information community in general whom I have had the opportunity to engage with during my thesis.

Thirdly, I would like to thank those who made possible or contributed to the presentations of my research in development; those responsible for organising seminars/workshops/conferences at which I was able to present and those who provided me with feedback. I also am grateful to the journals in which some of this work has been published.

I would lastly like to thank those family members and friends who have contributed to my development and supported me throughout this thesis, particularly Simone Schmidt.

Chapter 1

An Introductory Overview of Information

The term ‘information’ has become ubiquitous. In fact the notion of information is arguably amongst the most important of our ‘Information Age’. But just what exactly is information? This is a question without a straightforward response, particularly as information is a polysemantic concept, applied to a range of phenomena across a range of disciplines. The notion of information is associated with many central concerns of philosophy and has been used in various ways. Dealings with information from within philosophy include:

- Work on conceptions and analyses of information, as exemplified by recent work in the philosophy of information [74].
- The application of information to philosophical issues, two examples being:
 1. The use of information to develop accounts of knowledge, as exemplified in Fred Dretske’s information-theoretic epistemology [51].
 2. Informational semantics for logic, particularly relevant logic [130, 156, 11].
- Information ethics, “the branch of ethics that focuses on the relationship between the creation, organization, dissemination, and use of information, and the ethical standards and moral codes governing human conduct in society” [155].

Further to this, conceptions of information within other disciplines such as biology and physics can be and have been of interest within philosophy [83, 24].

Notably with the advent of the Information Age, information has increasingly come to be

seen as an important and useful notion within philosophy. This has reached a point where the field of the philosophy of information has been established [76].

In this thesis I focus on certain conceptions of information of particular interest to philosophers, so called semantic (non-natural) information and environmental (natural) information, which can be seen as roughly correlating to the Gricean notions of non-natural and natural meaning respectively (more on this to follow).

1.1 What is Information?

Information is applied in a variety of ways across a variety of disciplines. The computer scientist speaks of digital information, the biologist speaks of genetic information, the network engineer speaks of information channels, the cognitive scientist speaks of sensory information and the physicist speaks of physical information. Ordinarily we say that a watch provides us with information about the time, or that a newspaper provides us with information about the weather. If we want to prepare a dish we have not made before, we seek out a recipe to provide us with information on how to make it. These are but some of the senses in which we use the term.

The French mathematician Rene Thom neatly captured this polymorphic nature, by calling ‘information’ a ‘semantic chameleon’, something that changes itself easily to correspond to the environment. But “the plethora of different analyses can be confusing. Complaints about misunderstandings and misuses of the very idea of information are frequently expressed”, with some criticising others for laxity in use of the term ‘information’ [73]. Indeed plethoric usage of the term can be somewhat overwhelming and caution should be exercised to avoid ‘information’ being used as a buzzword, placeholder or term synonymous with ‘stuff’.

Given its numerous definitions and applications, the question naturally arises: is a grand unified theory of information possible [64]? In this thesis I do not in the least intend to offer a definitive account or deliver a grand unified theory of information. So far as concerns me, I think it a good idea to heed ‘Shannon’s Premonition’:

The word ‘information’ has been given different meanings by various writers in the general field of information theory. It is likely that at least a number of these will prove sufficiently useful in certain applications to deserve further study and permanent recognition. It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications

of this general field. [166, p. 180]

This being said, I shall soon have to offer my own analysis and stipulate my own usage of the term ‘information’ in order to lay the foundations for this thesis.

On the issue of a grand unified theory of information, Luciano Floridi writes:

The reductionist approach holds that we can extract what is essential to understanding the concept of information and its dynamics from the wide variety of models, theories and explanations proposed. The non-reductionist argues that we are probably facing a network of logically interdependent but mutually irreducible concepts. ... Both approaches, as well as any other solution in between, are confronted by the difficulty of clarifying how the various meanings of information are related, and whether some concepts of information are more central or fundamental than others and should be privileged. Waving a Wittgensteinian suggestion of family resemblance means acknowledging the problem, not solving it. [64, p. 563]

For my part I am inclined to adopt a nonreductionist stance. Rather than trying to develop a reductionist account, for me the pertinent task would be to investigate why the term ‘information’ is used in such a variety of ways. Rather than something like “what is information?” or “what is the set of elements common to the different types of information?”, the question should be something like “why is the word ‘information’ so widely and variously used?”. There are three broad options regarding a response to this question, listed as follows:

1. There is one or more characteristic that every usage of the term ‘information’ relates to
2. There is no one characteristic, but rather a family resemblance
3. Neither of the above two hold. Still, the set of characteristics associated with a conception of information can be explained via some means (e.g. etymology, linguistic convention, origins and applicability within areas such as science, technology and philosophy, etc)

Leaving this ambitious question aside, we will now turn towards laying the foundations for this thesis. Weaver [175, p. 11] offered the following tripartite analysis of information:

1. technical aspects concerning the quantification of information and dealt with by the mathematical theory of communication
2. semantic aspects relating to meaning and truth
3. pragmatic (what he terms ‘influential’) aspects concerning the impact and effectiveness of information on human activity

This thesis revolves around 2 and to a lesser extent 3, which are both built upon 1. Figure 1.1 depicts a useful map of information given by Floridi [73], which we will refer to in the following discussion.

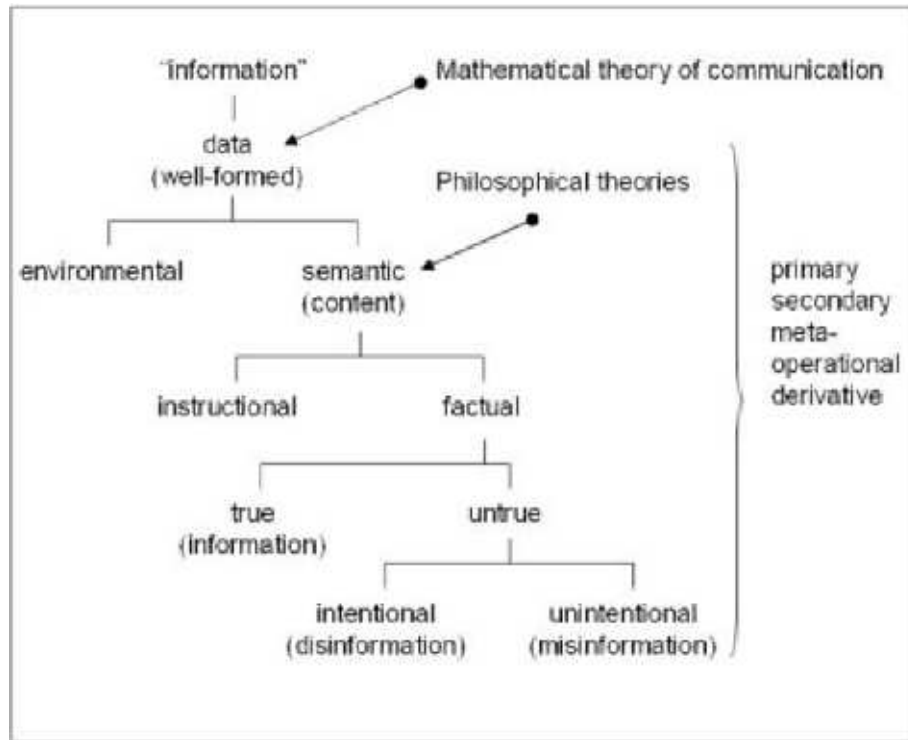


Figure 1.1: An informational map ([73])

1.1.1 Data

Whatever information is exactly, it seems right to think of it as consisting of data. Thus in this work information will be defined in terms of data. Whilst ‘data’ and ‘information’ are sometimes treated as synonymous terms, as will become clearer a separation of the two is required. Data is prior to information and “information cannot be dataless but, in the simplest case, it can consist of a single datum” [73]. So information implies data but

data does not imply information. This separation will not only increase the specificity and descriptive power of these terms, but will facilitate the construction of a concept hierarchy¹.

A general definition of a datum, what Floridi calls the ‘The Diaphoric Definition of Data’ [DDD] (*diaphora* being the Greek word for ‘difference’), is:

A datum is a putative fact regarding some difference or lack of uniformity within some context. [73]

When Gregory Bateson said that information ‘is a difference which makes a difference’, the first difference refers to data.

The following examples and discussion will serve to elucidate this definition:

- Take a system consisting solely of a single sheet of unmarked white paper. Without a way to mark the paper, there is no way to create a distinction between a blank sheet and a marked sheet. Thus there is uniformity and no possibility of difference; the only way the sheet can be is blank. However, if a pen were added to the system and used to mark the sheet, then there would be a lack of uniformity. Now that a distinction between a blank white sheet and marked sheets can be made, there is now the possibility of data. When marked, the white background plus the marks would constitute data.
- The previous example raises an important point. A blank page can still be a datum, as long as there is something like the possibility that something could be written on it. There is no data only when there is no possibility but a blank sheet of paper. In much the same way, the absence of a signal from a system can still constitute a (informative) datum. For example, the silence of a smoke alarm constitutes a datum and carries the information that there is no smoke within the vicinity. In general, an unmarked medium, silence, or lack of signal can still constitute a datum, as long as it is one possibility amongst two or more possibilities.
- As another example, consider a unary alphabet that consists solely of the symbol ‘0’. Using this alphabet it is not possible to generate data, for there could be no difference or lack of uniformity in output. However, if the alphabet were expanded to include the symbol ‘1’ as well as the symbol ‘0’, then it would be possible for the source to emit data, by using both instances of the ‘0’ symbol and instances of the ‘1’ symbol.

¹For example, the Data-Information-Knowledge-Wisdom hierarchy associated with knowledge management and systems analysis

Imagine trying to type a message to someone over the internet and the only button working on the keyboard was the ‘0’ button. It would be impossible to communicate any information to this person because such a system does not provide the means to generate any data.

- This idea extends beyond human semantic and communication systems. For example, genetic information can be translated into an alphabet consisting of four elements: ‘A’, ‘C’, ‘G’ and ‘T’. So there must be data in biological systems before there can be any information.

It is now time to discuss some more detailed matters concerning data. Whilst a position on such matters is not necessary for the main purposes of this thesis, their outlining forms a valuable part of its theoretical background.

As suggested by Floridi, the Diaphoric Definition of Data can be interpreted in three ways.

- data as lacks of uniformity in the external world, as diaphora *de re* (Floridi terms them *dedomena*, the word for data in Greek). They are the differences out there in the real world, prior to epistemic interpretation, which are empirically inferred via experience (think Kant’s *noumena*). Data *de re* are “whatever lack of uniformity in the world is the source of (what looks to information systems like us as) data” [73].
- data as lacks of uniformity between at least two physical states, as diaphora *de signo*. Here the difference occurs at the level of epistemic interpretation and perception, such as when one reads a message written in English. The data are the characters from the English alphabet that form the message.
- data as lacks of uniformity between symbols, as diaphora *de dicto*. These are the pure differences between symbols in some system, such as the numerical digits 1 and 2.

“Depending on one’s position with respect to the thesis of ontological neutrality and the nature of environmental information [see below] *dedomena* in (1) may be either identical with, or what makes possible signals in (2), and signals in (2) are what make possible the coding of symbols in (3)” [73].

Floridi also identifies four types of neutrality associated with DDD.

1. **Taxonomic Neutrality (TaN):** - A datum is a relational entity.

- According to TaN, nothing is a datum *per se*. Rather, it is the relation between two or more things that constitutes a datum. In the paper and marker example above, neither a black dot mark nor the white background of the paper (the two relata) is the datum. Rather both, along with the fundamental relation of inequality between the dot and the background constitute the datum.

2. **Typological Neutrality (TyN):** - Information can consist of different types of data as relata.

- Following are five standard classifications. They are not mutually exclusive and more than one classification might apply to the same data, depending on factors such as the type of analysis conducted and at what level of abstraction:

- 1 Primary data - The principal data of an information system, whose creation and transmission is the system's purpose. A smoke alarm's audio signal is primary data.
- 2 Secondary data - The converse of primary data, these are data resulting from absence. The failure of a smoke alarm to sound amidst a smoke-filled environment provides secondary information that it is not functioning.
- 3 Metadata - Data about data, describing certain properties. For example, web pages often include metadata in the form of meta tags. These tags describe certain properties of the web pages, such as the document's creator, document keywords and document description.
- 4 Operational data - Data regarding the operation and performance of a whole data/information system. For example, a computer might have a set of monitoring tools, one of which monitors the status of the memory hardware. A signal that one of the two memory sticks has died would provide operational data/information indicating why the computer has slowed down.
- 5 Derivative data - Data that can be derived from a collection of data. For example, say a football team has a collection of documents recording the performance statistics of each player for each game played for the previous season. The presence of player A on the list for the last game is primary data that they played. The absence of player B on the list for the last game is secondary data that they did not play. If one were to extract the pattern that player C underperforms when their side goes into the last quarter with the lower score, this would be a piece of derivative data/information.

3. **Ontological Neutrality (ON):** - There can be no information without data representation.

4. **Genetic Neutrality (GeN):** - Data (as relata) can have a semantics independently of any informee.

- According to GeN, there can be semantic data (information) without an informee. Meaning at least partly exists outside the minds of semantic agents. To use Floridi's example, the Rosetta Stone already contained semantic data/information prior to its accessibility upon the discovery of an interface between Greek and Egyptian. Note though, this is not to say that data can have a semantics without being created by a semantic agent, that semantic data is independent of intention.

This discussion of data places us at the root of Figure 1.1. We shall now turn to an outline of the Mathematical Theory of Communication (MTC), which deals with information as data communication.²

1.1.2 The Mathematical Theory of Communication

The *Mathematical Theory of Communication* (or *Information Theory* as it is also known as) was developed primarily by Claude Shannon in the 1940s [168]. It measures the information (structured data) generated by an event (outcome) in terms of the event's statistical probability and is concerned with the transmission of such structured data over (noisy) communication channels.

The Shannon/Weaver communication model with which MTC is concerned is given in Figure 1.2.³

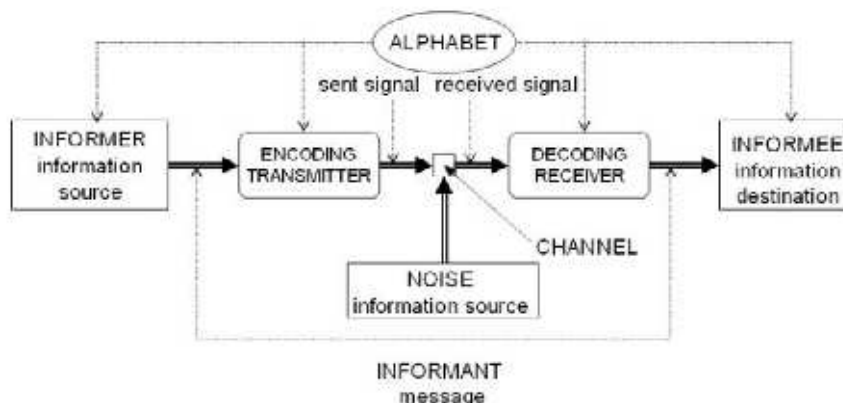


Figure 1.2: Shannon Weaver communication model ([73])

²MTC is the most important and widely known mathematical approach to information. Its ideas and application have also made their way into philosophy. Another approach of particular significance within philosophy is Algorithmic Information Theory (AIT) [30].

³The applicability of this fundamental general communication model goes beyond MTC and it underlies many accounts of information and its transmission, including those found throughout this thesis.

A good example of this model in action is Internet telephony. John says ‘Hello Sally’ in starting a conversation with Sally over Skype. John is the informer or information source and the words he utters constitute the message. His computer receives this message via its microphone and digitally encodes it in preparation for transmission. The encoding is done in a binary alphabet, consisting conceptually of ‘0’ and ‘1’s. The signal for this encoded message is sent over the Internet, which is the communication channel. Along the way some noise is added to the message, which interferes with the data corresponding to ‘Sally’. The received signal is decoded by Sally’s computer, converted into audio and played through the speakers. Sally, the informee at the information destination, hears ‘Hello Sal**’, where * stands for unintelligible crackles due to the noise in the decoded signal.⁴

In order to successfully carry out such communication, there are several factors that need to be worked out. What is the (minimum) amount of information required for the message and how can it be encoded? How can unwanted equivocation and noise in the communication channel be dealt with? What is the channel’s capacity and how does this determine the ultimate rate of data transmission? Since MTC addresses these questions it plays a central role in achieving the execution of this model. Given its foundational importance, we will now briefly go over the basic mathematical ideas behind MTC.⁵

Let $\mathbb{S}$ stand for some event/outcome/source which generates/emits symbols in some alphabet $\mathbb{A}$ which consists of n symbols. As three examples of this template, consider the following:

1. $\mathbb{S}$ is the tossing of a coin. $\mathbb{A}$ consists of two symbols, ‘heads’ and ‘tails’.
2. $\mathbb{S}$ is the rolling of a die. $\mathbb{A}$ consists of six symbols, the numbers 1-6.
3. $\mathbb{S}$ is the drawing of a name in an eight-person raffle. $\mathbb{A}$ consists of each of the eight participants names.

The information measure associated with an event is proportional to the amount of certainty it reduces. For an event where all symbols have an equal probability of occurring, the probability of any one event occurring is $\frac{1}{n}$. The greater n is to begin with, the greater the number of initial possibilities, therefore the greater the reduction in uncertainty or data deficit. This is made mathematically precise with the following formulation. Given an

⁴Despite the simplicity of this high-level account, it adequately illustrates the fundamental components involved. Beyond this account there are richer ones to be given. Firstly, each process can be explained in greater degrees of detail. Secondly, this communication model can apply to other processes in the whole picture; for example, the communication involved in the transmission of sound from Sally’s ear to a signal in her brain.

⁵For a very accessible introduction to MTC, see [144]. For a considerably in-depth textbook see [40]. Shannon’s original paper is [168].

alphabet of n equiprobable symbols, the information measure or entropy of the source is calculated with the following:

$$\log_2(n)\text{bits} \quad (1.1)$$

Going back to the above three examples:

1. The outcome of a coin toss generates $\log_2(2) = 1$ bit of information
2. The outcome of a die roll generates $\log_2(6) = 2.585$ bits of information
3. The outcome of an eight-person raffle generates $\log_2(8) = 3$ bits of information

In cases where there is only one possible outcome uncertainty is zero and thus so is the information measurement. In terms of playing card types, the random selection of a card from a standard 52-card deck generates $\log_2(52) = 5.7$ bits of information. But the selection of a card from a deck consisting of 52 cards, all king of spades, generates $\log_2(1) = 0$ bits of information.

Skipping over the technical details and derivations, the general formula for the entropy (H) of a source, the average quantity of information it produces (in bits per symbol), is given by

$$H = - \sum_{i=1}^n \text{Pr}(i) \log_2 \text{Pr}(i) \text{ bits per symbol} \quad (1.2)$$

for each of the n possible outcomes/symbols i . When all outcomes are equiprobable, this equation reduces to that of formula 1.1. When the source's outcomes are not all equiprobable things become more interesting.

Let us start with a fair coin, so that $\text{Pr}(\text{'heads'}) = \text{Pr}(\text{'tails'}) = 0.5$. Plugging these figures into Equation 1.2, we get:

$$H = -(\frac{1}{2} \times \log_2(\frac{1}{2}) + \frac{1}{2} \times \log_2(\frac{1}{2})) = -(\frac{1}{2} \times -1 + \frac{1}{2} \times -1) = 1 \text{ bit}$$

which is the same as $\log_2(2) = 1$ bit.

But now consider a biased coin, such that $\text{Pr}(\text{'heads'}) = 0.3$ and $\text{Pr}(\text{'tails'}) = 0.7$. Plugging these figures into equation 1.2, we get:

$H = -(0.3 \times \log_2(0.3) + 0.7 \times \log_2(0.7)) = -((0.3 \times -1.737) + (0.7 \times -0.515)) = 0.8816$ bits

So the biased coin generates less information than the fair coin. This is because the overall uncertainty in the biased coin case is less than in the fair coin case; with the former case there is a higher chance of ‘tails’ and lower chance of ‘heads’ so in a sense any outcome is less surprising. This is all mathematically determined by the structure of the formula for H . The occurrence of some particular symbol generates some amount of information; the lower the probability of it occurring the higher the information generated. This is represented with the $\log_2 \Pr(i)$ part. Although a lower probability means more information on an individual basis, with the average calculation this is regulated and diminished with the multiplication by its own probability. This balance is why H takes its highest value when all of a source’s potential symbols are equiprobable.

Equation 1.2 represents a fundamental limit. It represents the lower limit on the expected number of symbols (‘0’s and ‘1’s)⁶ required to devise a coding scheme for the outcomes of an event, irrespective of the coding method employed. It represents the most efficient way that the signals for an event can be encoded. It is in this sense that H is the unique measure of information quantity.

This point can be appreciated with the simplest of examples. Take the tossing of two fair coins (h = heads, t = tails). John is to toss the coins and communicate the outcome to Sally by sending her a binary digital message. Since the coins are fair, $\Pr((h, h)) = \Pr((h, t)) = \Pr((t, h)) = \Pr((t, t)) = 0.25$. The number of bits required to code for the tossing of these coins is two ($H = 2$); it is simply not possible on average to encode this information in less than two bits. Given this, John and Sally agree on the following encoding scheme:

- $(h, h) = 00$
- $(h, t) = 01$
- $(t, h) = 10$
- $(t, t) = 11$

As an example, the string which encodes the four outcomes (h, h) , (t, t) , (h, t) and (t, h) is ‘00110100’.

⁶This is the standard set of symbols for a binary alphabet ($n = 2$). n is the same as the base of the logarithm in H , which does not have to be two. A base that has also been used is ten (in this case the unit of information is called a ‘Hartley’, in honour of Ralph Hartley, who originated the measure). Natural logarithms have also been used (in this case the unit of information is called a ‘nat’).

Now, modify this coin scenario so that the outcomes have the following probabilities:

- $\Pr(h, h) = 0.5$
- $\Pr(h, t) = 0.25$
- $\Pr(t, h) = 0.125$
- $\Pr(t, t) = 0.125$

Given this probability distribution, $H = 1.75$ bits. As we have just seen, this means that the lower limit on the average number of symbols required to code each tossing of the two coins is 1.75. How would such an encoding go? The basic idea is to assign fewer bits to the encoding of more probable outcomes and more bits to the encoding of less probable outcomes. Since (h, h) is the most probable outcome, fewer bits should be used to encode it. This way, the number of expected bits required is minimised.

The most efficient coding to capture this connection between higher probability of an outcome and more economical representation is:

- $(h, h) = 0$
- $(h, t) = 10$
- $(t, h) = 110$
- $(t, t) = 111$

If we treat the number of bits for each outcome as the information associated with that outcome, then we can plug these figures into the following formula and also get a calculation of 1.75:

- $\sum_{i=1}^n \Pr(i) \times \text{number of bits to represent } i \text{ (for each sequence } i)$
- $(0.5 \times 1) + (0.25 \times 2) + (0.125 \times 3) + (0.125 \times 3) = 1.75 \text{ bits}$

In comparison to the previous example, the string which represents the four outcomes (h, h) , (t, t) , (h, t) and (t, h) using this encoding scheme is the shorter '0111100'. This

optimal encoding scheme is the same as that which results from the Shannon-Fano coding method.⁷

The discussion of MTC thus far has involved fairly simple examples without any of the complications and complexities typically involved in realistic communication. To begin with, it has only considered the information source of perfect communication channels, where data is received if and only if it is sent. In real conditions, communications channels are subject to *equivocation* and *noise*. The former is data that is sent but never received and the latter is data that is received but not sent. The communication system as a whole involves both the possible outcomes that can originate from the information source $S = \{s_1, s_2, \dots, s_m\}$ and the possible signals that can be received at the information destination $R = \{r_1, r_2, \dots, r_n\}$. It is the statistical relations between S and R (the conditional probability that an element in one set occurs given the occurrence of an element from the other set) that determine the communication channel. Some of the technical details and example calculations regarding these factors will be covered in Chapters 4 and 5.

Redundancy refers to the difference between the number of bits used to transmit a message and the number of bits of fundamental information in the message as per the entropy information calculations. Whilst redundancy minimisation through data compression is desirable, redundancy can also be a good thing, as it is used to deal with noise and equivocation. As the simplest of examples, if John says ‘hello hello’ to Sally, the second hello is redundant. But if the first hello becomes unintelligible due to noise/equivocation, then an intelligible second hello will serve to counter the noise/equivocation and communicate the information of original message. In technical digital communication, sophisticated error detection and correction algorithms economically use desired redundancy.

Another factor to briefly mention is that the probability distribution of the source can be conditional. In our examples, the probability distributions were fixed and the probability of one outcome was independent of any preceding outcome. The term for such a system is *ergodic*. Many realistic systems are non-ergodic. For example, you are about to be sent an English message character by character. At the start there is a probability distribution across the range of symbols (i.e. English alphabet characters). If an ‘h’ occurs as the first character in the message then the probability distribution changes. For example, the probability that the next character is a vowel would increase and the probabilities for ‘h’ and ‘k’ would decrease, effectively to zero, since there are no valid constructions in the English language with ‘hh’ or ‘hk’. Whilst such complications and complexities are covered by MTC,

⁷Whilst it produces an optimal encoding in this case, in general this method is suboptimal, in that it does not achieve the lowest possible expected code word length. Another method which does generally achieve the lowest possible expected code word length is Huffman coding (http://en.wikipedia.org/wiki/Shannon-Fano_coding).

the details are unnecessary for our purposes and need not detain us.

Continuing on, once a message is encoded it can be transmitted through a communication channel.⁸ Shannon came up with the following two theorems concerning information transmission rates over communication channels. Let C stand for the transmission rate of a channel, measured in bits per second (bps). Firstly, there is Shannon's theorem for noiseless channels:

Shannon's Theorem for Noiseless Channels: Let a source have entropy H (bits per symbol) and a channel have a capacity C (bits per second). Then it is possible to encode the output of the source in such a way as to transmit at the average rate of $C/H - \epsilon$ symbols per second over the channel where ϵ is arbitrarily small. It is not possible to transmit at an average rate greater than C/H . [167, p. 59]

To deal with the presence of noise in practical applications, there is the corresponding theorem for a discrete channel with noise:

Shannon's Theorem for Discrete Channels: Let a discrete channel have the capacity C and a discrete source the entropy per second H . If $H \leq C$ there exists a coding system such that the output of the source can be transmitted over the channel with an arbitrarily small frequency of errors (or an arbitrarily small equivocation). If $H > C$ it is possible to encode the source so that the equivocation is less than $H - C + \epsilon$ where ϵ is arbitrarily small. There is no method of encoding which gives an equivocation less than $H - C$. [167, p. 71]

With these fundamental theorems stated, we now come to conclude this outline of MTC. As can be gathered, MTC covers several properties associated with an intuitive conception of information:

- information is quantifiable
- information quantity is inversely related to probability
- information can be encoded
- information is non-negative
- information is additive

⁸Research into the practical concerns of communication was a key factor for Shannon, whose interest resulted from his work at AT&T Bell Labs. As a telephone company, they wanted to know what minimum capacities their networks needed in order to efficiently handle the amounts of traffic (data) they were expecting to deal with.

Ultimately however MTC is a syntactic treatment of information that is not really concerned with semantic aspects. Although it deals with structured data that is potentially meaningful, any such meaning has no bearing on MTC's domain. Theories of semantic information on the other hand deal with data that is meaningful and the use of such data by semantic agents. The following examples serve to illustrate these points:

1. For MTC information is generated when one symbol is selected from a set of potential symbols. As we have seen, entropy, the measure of information associated with a source, is inversely related to probability.

If three-letter strings were being generated by randomly selecting English alphabetical characters, then since 'xjk' is just as probable as 'dog', it yields just as much MTC-information, despite the former being gibberish and the latter being a meaningful English word.

Or suppose that three-character words were being randomly selected from a novel written in English featuring an animal. Although the word 'fur' is less probable than 'dog', the latter is semantically more informative than the former.

2. Take a network over which statements in the English language are encoded into ASCII⁹ messages and transmitted. The encoding of each character requires 7 bits. Now, consider the following three strings:

- the an two green four cat !?down downx
- Colourless green ideas sleep furiously
- The native grass grew nicely in spring

Although each message uses the same number of bits ($7 \times 38 = 266$), within the English language the first is meaningless and not well-formed and the second is well-formed but is not meaningful. Only the third is well-formed and meaningful and hence can be considered to be semantic information.

3. Consider a basic propositional logic framework. Say that for each symbol in a statement 1 unit of data is required in its encoded message. Consider the following three strings:

- $A \neg B$
- $A \vee B$
- $A \wedge B$

Each of these statements contains the same quantity of syntactic data. The first however, is not well-formed. Whilst the second and third are well-formed, according to

⁹American Standard Code for Information Interchange: <http://en.wikipedia.org/wiki/ASCII>

the meanings of the connectives $\vee$ and $\wedge$, there is a sense in which $A \wedge B$ is more informative than $A \vee B$ ¹⁰.

So MTC is ultimately about the quantification and communication of syntactic information or data. It

approaches information as physical phenomenon, syntactically. It is interested not in the usefulness, relevance, interpretation, or aboutness of data but in the level of detail and frequency in the uninterpreted data (signals or messages). It provides a successful mathematical theory because its central question is whether and how much data, not what information is conveyed. [64, p. 561]

Whilst it is not meant to deal with the semantic aspects of information, since MTC deals with the data that constitutes semantic information it is still relevant and provides mathematical/technical constraints for theories of semantic information and a philosophy of information [75]. Ideas from MTC have actually found application in methods of quantitatively measuring semantic information in communication [13]. Furthermore, as will be covered in Chapters 4 and 5, MTC can serve as a starting point and guide for a semantical theory of information.

Chapman [31] argues for a stronger link between Shannon information and semantic information. Whether, as suggested by some, and to what extent, the semantic level of information can be reduced to the syntactic level is an interesting question. However it is not something to be preoccupied with here. To begin with, there is still much work to be done on understanding semantic information. Secondly, I think such reductionist ideas at this stage amount to tentative suggestions, and could still very well ultimately “belong to an old-fashioned, perfectly respectable but also bankrupted tradition of attempting to squeeze semantics out of syntax” [75, p. 259].

1.1.3 Moving Beyond Data

Moving towards the right from the root of Figure 1.1, we move towards forms of information as semantic content. It is accounts within this region which address aspects left by mathematical treatments of information as data and which will be investigated throughout this thesis. I will go over this region very briefly here as it will be covered in greater detail later on. Information as semantic content is basically information as meaningful data. This is

¹⁰Since it is less probable and $A \wedge B \vdash A \vee B$.

information in the ordinary or common sense and is amenable to a propositional analysis. When Bateson wrote that information ‘is a difference which makes a difference’, the second difference suggests that the data is (potentially) meaningful.

There are two types of semantic content, factual and instructional. The proposition ‘Canberra is the capital city of Australia’ is factual semantic content. The content of a cake recipe constitutes an instance of instructional information. Factual semantic content can be true or false (untrue). False semantic content is *misinformation* if it is unintentionally spread, *disinformation* if it is intentionally spread.

Some further comments on this right branch:

- It implies a difference in the informational status of true semantic content and false semantic content. We will look at this in detail a little further on.
- The choice of ‘untrue’ instead of ‘false’ might be somewhat problematic and it would be better to substitute the latter for the former and perhaps add another branch. Given a classical truth system, this is straightforward since ‘untrue’ is equivalent to ‘false’. But given a different system, for example, one in which it is possible to have meaningful propositions that are neither true nor false, the map would be unsound, since such statements are neutral and should not be classed as disinformation or misinformation.
- I would modify the disinformation and misinformation leaves, so that misinformation is false semantic content in general and disinformation is a subclass consisting of intentional false semantic content.

Environmental information on the left branch refers to the sense of information we use when we say that something occurring carries information that something else is occurring; for example, when we say that the presence of smoke carries the information that there is a fire. This notion of information will be introduced in more detail below.

Given this outline, a good information classification system to keep in mind is as follows [64, p. 560]:

1. information as reality (e.g. patterns of physical signals, which are neither true nor false)
2. information about reality (alethically qualifiable semantic information)
3. information for reality (instructions, construction manuals, genetic information)

1.2 Philosophy and Information

As has been established by now, there is interest in the notion of information from within philosophy. This interest includes the technical, semantic and pragmatic aspects of information. Notable work on information from within the domain of philosophy has occurred sporadically. Some of the key developments are listed below. By no means comprehensive, the list will at least provide some historical insight and serve as a rough lineage for the work in this thesis.¹¹

- Observing the limitations of MTC, around the middle of the 20th century Yehoshua Bar-Hillel and Rudolf Carnap gave a quantitative account of semantic information. This work will serve as a starting point for Chapter 2. As will also be covered in Chapter 2, Hintikka expanded upon some of this work and also made other contributions in applying information to philosophy and logic.
- In his influential book *Knowledge and the Flow of Information* [51] Fred Dretske provides a semantic theory of information and gives an account of information flow, what it is for one thing to carry information about another thing. With this account of information he gives a definition of knowledge. He also explains perception in relation to information and attempts to develop a theory of meaning by viewing meaning as a certain kind of information-carrying role.
- Accounts of information have been developed using the framework of situation semantics. In *Logic and Information* [46], Keith Devlin introduces the concept of *infor* and merges it with situation theory. Later on, Jon Barwise and Jerry Seligman developed situation semantics into a formal model of information flow [16]. Perry and Israel [142, 143] are two other notable figures in this school.
- One example of an independent, relatively early investigation into information is Christopher Fox's. *Information and Misinformation: An Investigation of the Notions of Information, Misinformation, Informing, and Misinforming* [79]. Fox employs an ordinary language analysis of information to get some insight into its nature. He develops the notions of information and misinformation to serve as part of the foundation for an information science.

¹¹For a good survey of 'the informational turn' in philosophy see [2]

1.2.1 The Philosophy of Information

The philosophy of information (PI) is the area of research that studies conceptual issues arising at the intersection of computer science, information technology, and philosophy. It concerns [76]:

1. the critical investigation of the conceptual nature and basic principles of information, including its dynamics, utilisation and sciences
2. the elaboration and application of information-theoretic and computational methodologies to philosophical problems.

This field has emerged relatively recently and was largely initiated by Luciano Floridi, whose work over the past decade or so has helped to establish the field. This is certainly not to say that Floridi is responsible for introducing philosophy to the notion of information. But as noted by Michael Dunn he does deserve credit “for his insight in establishing the very concept of the ‘philosophy of information’. His books and various papers really legitimated this as a major area in philosophy and not just a minor topic” [62].

There is some contention surrounding the naming of this field. From my observations I take it that some would prefer something like ‘*a* philosophy of information’ rather than ‘*the* philosophy of information’.¹² I think that such concerns are unwarranted. Floridi is not equating ‘the philosophy of information’ with the ‘be all and end all of information’. Nor is he suggesting that his own positions on information matters are law/lore. Rather, the way I see it he has worked and continues working to establish a field of philosophy whose subject matter is information, in the same way that the philosophy of language is the field of philosophy whose subject matter is language. Thus as there is a variety of issues, perspectives and subtopics within any philosophy of *X*, so it is the case with the philosophy of information. Also, in the same way that *the* philosophy of language exists alongside linguistics, the philosophy of information can exist alongside other fields such as information science. Furthermore, their relationship can be a mutually beneficial one.

Wrapping up this small foray, the philosophy of information can be seen as the field of philosophy that philosophers philosophising about and using information work under. It is in this field that my work is situated.

¹²See <http://theoccasionalinformationist.com/2011/08/16/the-philosophy-or-a-philosophy-of-information/> for example

1.3 Semantic Information and Environmental Information

It is now time for a detailed introduction to the two conceptions of information central to this thesis, semantic information and environmental information. For the sake of historical perspective, I would like to begin by outlining Paul Grice's two definitions of meaning, which can be seen as rough correlates of these two conceptions of information.

1.3.1 Gricean Meaning

In his paper 'Meaning' [90], Grice begins by making a distinction between what he terms *natural* meaning (meaning_N) and *non-natural* meaning (meaning_{NN}). Here are two examples he provides to illustrate these two senses of meaning:

(N) Those spots mean measles

(NN) Those three rings on the bell (of the bus) mean that the bus is full.

The correlations between our notions of information and Grice's senses of meaning are obvious, with natural meaning corresponding to environmental information and non-natural meaning corresponding to semantic information.

A whistling kettle (naturally) means that the water has boiled in the sense that a whistling kettle carries the information that the water is boiling. Likewise, smoke (naturally) means fire in the sense that smoke carries the information that there is fire. The presence of this meaning/information involves regularity and is independent of semantics.

A tick on a student's essay (nonnaturally) means that the student has done well. The sign of a tick signifies good work, irrespective of whether or not the work is actually good. In this way, a tick provides semantic information that the marker commended the piece. Likewise, an exclamation mark at the end of a sentence (nonnaturally) means that the sentence is exclaimed. By symbolic convention '!' signifies and provides the semantic information that the writer exclaims the sentence. So non-natural meaning requires some semantic framework: "the presence of meaning_{NN} is dependent on a framework provided by a linguistic, or at least a communication-engaged community" [91, p. 350].

Grice maintains that sentences like (N) are factive, while sentences like (NN) are not. Grice notes that it would be contradictory to say:

(N*) Those spots mean measles, but he hasn't got measles

So in the case of natural meaning, sentences of the form ' x means that p ' entail the truth of p . Non-natural meaning on the other hand is non-factive; sentences of the form ' x means that p ' do not entail the truth of p . If someone rings a bus' bell three times, it non-naturally means that the bus is full according to the standard communication framework, even if the bus is not actually full.

This outline of meaning serves to initiate some points on information. Firstly, environmental information will also be taken to be factive, in the sense that if A carries the information that B , and it is the case that A , then it is also the case that B . Secondly, although starting off with a conception of semantic information as alethically neutral semantic (propositional) content, I will subsequently endorse a conception of semantic information that implies truth.

1.3.2 Semantic Information

We start off with a *General Definition of (Semantic) Information* (GDI) as *data + meaning*. The following tripartite definition is taken from [73]:

The General Definition of Information (GDI):

σ is an instance of information, understood as semantic content, if and only if:

1. σ consists of one or more *data*
2. the data in σ are *well-formed*
3. the well-formed data in σ are *meaningful*

Condition 1 simply states that data are the stuff of which information is made, as covered earlier in Section 1.1.1.

With condition 2, 'well-formed' means that the data are composed according to the rules (syntax) governing the chosen system, code or language being analysed. Syntax here is to be understood generally, not just linguistically, as what determines the form, construction, composition or structuring of something [73]. The string 'the an two green four cat !?down downx' is not well-formed in accordance with the rules of the English language, so therefore cannot be an instance of semantic content in the English language. Or, to take another

example, the string ‘ $A \neg B$ ’ is not well-formed in accordance with the rules of the language of propositional logic, so therefore cannot be an instance of semantic content in propositional logic.

With condition 3, ‘meaningful’ means that the well-formed data must comply with the meanings (semantics) of the chosen system, code or language in question. For example, the well-formed string ‘Colourless green ideas sleep furiously’ cannot be semantic content in the English language because it is meaningless; it does not correspond to anything. An example of a string which fulfils conditions 1, 2 and 3 is ‘The native grass grew nicely in spring’.

There are two main types of information, understood as semantic content: *factual* and *instructional*. Put simply, factual information represents facts; when the next train is coming, what the capital of a country is, how many people are in the room, etc. Although our main interest here lies with factual information, firstly a brief discussion of instructional information is in order.

Unlike factual information, “instructional information is not about a situation, a fact, or a state of affairs w and does not model, or describe or represent w . Rather, it is meant to (help to) bring about w ” [73]. The information contained in a cake recipe is an example of instructional information; the instructions in the recipe help bring about the production of a cake, by informing the baker about what needs to be done. Another example is sheet music. The notes written on a page do not provide facts. They are instructions that inform the musician how to play a piece of music. Although not factual, these instances of information are semantic in nature and they still have to be well-formed and meaningful to an informee.

It does however seem that instructional information is reducible to factual information. Fox distinguishes between these two senses of the term ‘information’: information-how (i.e. instructional information) and information-that (i.e. factual information) [79, p. 14]. Information-how is information that consists of instructions about how to carry out a task or achieve a goal and it is carried by imperative sentences. In contrast, information-that is information to the effect that some state of affairs obtains and is carried by indicative sentences.

No doubt there are parallels between the instructional/factual information dichotomy and the knowledge-that/knowledge-how dichotomy in epistemology. With the latter, it is sometimes argued that knowledge-how is ultimately reducible to knowledge-that. According to Fox, “whether this is so or not, it certainly is the case that the parallel reduction of information-how to information-that can be carried out” [79, p. 16].

According to his reduction method, information on how to do some task T is a sequence

of instruction $t_1, t_2, \dots, t_n$, all in the imperative mood. These instructions can be converted into an indicative statement containing information-that, by constructing a sentence of the following form: ‘Task t can be accomplished by carrying out the following instructions: $t_1, t_2, \dots, t_n$ ’. In converting information-how to information-that here, there is no loss, since any task that could be accomplished using the information-how of the sequence of instructions $t_1, t_2, \dots, t_n$, can still be carried out using the indicative equivalent. Thus information-how is reducible to information-that.

As Fox points out, the converse reduction is generally not possible. For example, the information that lemons are yellow is not amenable to reduction in terms of a sequence of instructions. This reducibility of information-how and the irreducibility of information-that indicates that the latter sense of information is the more fundamental one. It is this primacy, along with a few other reasons he briefly mentions, which result in his focus on semantic information-that. Although the nature of information-how and its connection with information-that might be more involved than Fox’s coverage, in short, I agree with this position.

Instantiations of factual semantic content occur in a variety of ways. Here are some examples:

- A map of Europe contains the true semantic content that Germany is north of Italy, in the language of cartography. The data that this semantic content is made of is identified with the sheet of paper on which the map is printed plus the various markings on the page. This data is well-formed; among other things, the North-South-East-West coordinates are correctly positioned and no countries are marked as overlapping each other. Finally, this data is meaningful; bordered and coloured parts of the paper correspond to countries, thin blue lines mean rivers, etc.
- A person’s nod contains the true semantic content that they are in agreement, in certain human body languages. The data that this semantic content is made of is identified with the variation in head position. This data is well-formed; head movement is a legitimate expression in the language. This data is also meaningful; this particular expression means ‘yes’ or ‘positive’.
- The content of an Encyclopaedia Britannica entry on Australia will contain the true semantic content that Canberra is the capital of Australia, in the language of English. The data that this semantic content is made of are the varied strings of English alphabetical symbols. This data is well-formed as it accords with the syntax of the English language and is also meaningful to an English language reader.
- The content of a book which says that there are nine planets in the solar system is

false semantic content. The data that this semantic content consists of are the varied strings of English alphabetical symbols. This data is well-formed as it accords with the syntax of the English and is also meaningful to an English language reader.

As can be seen, semantic information/content is often, but need not be, linguistic.

Clearly instructional information is not alethically qualifiable. Factual information on the other hand can be either true or false; truth and falsity supervene on information as factual semantic content.

We now come to the establishment of an important point regarding our analysis of semantic information. As has already been mentioned, according to GDI, for something to count as information, it needs to be an instance of well-formed, meaningful data. For it to be factual, it needs to be about some state of affairs, about some fact, whether it is true or false. Factual information comes in a variety of forms. A map of the world contains the factual information that Germany is north of Italy. A person's nod contains the factual information that they are in agreement. The content of an encyclopaedia entry on Australia will contain the information that Canberra is the capital of Australia. These various forms of semantic information can ultimately be expressed propositionally.¹³ Thus in this thesis factual semantic information is identified with propositions. If *i* is factual information, then it can be expressed in the form 'the information that *i*'. So although sentences can be said to carry information, this information is ultimately to be identified with the propositions that the sentences correspond to:

The *information* carried by a sentence *S* is a proposition appropriately associated with *S*. [79, p. 84]

Given this, the following sentences

- Two moons circle Mars
- The number of moons of Mars is the first prime number
- Marte ha due lune

¹³The reducibility of different kinds of semantic information to propositional form is, or at least I think it is, straightforward. For a discussion of this point, see [71, p. 153]. Fox cogently argues that information should be understood in terms of propositions, what he calls the propositional analysis of information [79, p. 75].

all instantiate the same information; although different sentences, they are not different pieces of information. Likewise, a picture of Mars with two moons around it would also be an instance of this information.¹⁴

Whilst the GDI and propositional analysis of information are straightforward enough and I do not wish to become engrossed in a discussion of any associated philosophical conundrums, I shall briefly raise a few points before closing this section.

To begin with, is condition 2 of GDI (well-formed) unnecessary or redundant? Is it possible to have data that is meaningful but not well-formed? If not, then the condition of meaningfulness in GDI renders the condition of well-formation redundant. Also, a stipulation of well-formation would rule out counting the not well-formed string ‘the car red won the race’ (String₁) as information in the English language, even though it is potentially meaningful. In such a case, given the propositional analysis String₁ could be considered a piece of data that can be mapped to the proposition represented by ‘the red car won the race’. Given the propositional analysis of information, perhaps data need only correspond to a proposition to count as an instance of information.

As mentioned earlier the rejection of dataless information leads to the following modest thesis of ontological neutrality:

There can be no information without data representation.

This thesis can be, and often is, interpreted materialistically, with the equation of representation and physical implementation leading to the following:

There can be no data representation without physical implementation.

These two imply that:

There can be no information without physical implementation.

If propositions are immaterial entities, then how can this statement be reconciled with the propositional analysis of information? The information *I* could be identified with a tuple

¹⁴Since information is identified with propositions and propositions here are identified with sets of possible worlds, the result is a coarse-grained account of semantic information. Thus ‘ $x = 2$ ’ and ‘ $x =$ the first prime number’ represent the same piece of information. For the purposes of this work such an analysis will suffice because the focus is on synthetic/empirical semantic information. However the issue of analytic/logical information and the sense in which analytic/logical truths (such as ‘ $2 =$ the first prime number’) can be informative will be touched upon.

(X, Y) , where X is a proposition and Y is a physical representation corresponding to that proposition. We then would have:

- $I_1 = (X_1, Y_1)$
- $I_2 = (X_2, Y_2)$
- $I_1 = I_2$ if and only if $X_1 = X_2$

Alternatively, one could reject the thesis that information (and data) requires physical implementation; “some philosophers have been able to accept the thesis that there can be no information without data representation while rejecting the thesis that information requires physical implementation” [73]. If this were the case, data could perhaps be identified with immaterial patterns of disuniformity. For example, according to this account the strings ‘110010’ and ‘BBAABA’ would both represent the same data.

The Logic of Data and Semantic Content

Since semantic content is propositional in nature it can be dealt with using a propositional logic. For example, if p is true semantic content and q is true semantic content, then $p \wedge q$ is true semantic content. Or if p is false semantic content, then $p \wedge q$ is false semantic content. But what happens when semantic content is connected with data that is not semantic content? For example, what is the status of the conjunction:

Colourless green ideas sleep furiously and The native grass grew nicely in spring

Bochvar’s 3-valued logic can be used as a logic to reason about data in general (semantic content and meaningless data). With Bochvar’s system, in addition to the classical values t (true) and f (false), a third value $*$ that represents ‘meaninglessness’ is introduced. Its purpose is to “avoid logical paradoxes such as Russell’s and Grelling’s by declaring the crucial sentences involving them to be meaningless” [174, p. 75]. For our purposes, $*$ is used to evaluate meaningless data. We let the usual propositional variables range over data and call them data variables. If a data variable is assigned the value t or the value f , then it qualifies as semantic content, since only data that is also semantic content can be alethically qualified. If a data variable is assigned the value $*$, then it is meaningless and fails to be semantic content.

The truth functions for the connectives behave as follows:

- return the same as classical logic when only classical truth values are input
- return $*$ whenever $*$ is input (a meaningless part ‘infects’ the whole)

Let A stand for ‘Colourless green ideas sleep furiously’ and B stand for ‘The native grass grew nicely in spring’. Since $v(A) = *$, $v(A \wedge B) = *$, irrespective of the value for B .

So in this system semantic content is output only if no meaningless data is input:

- $\neg A$ is semantic content iff A is semantic content
- $A \wedge B$ is semantic content iff A is semantic content and B is semantic content
- $A \vee B$ is semantic content iff A is semantic content and B is semantic content
- $A \supset B$ is semantic content iff A is semantic content and B is semantic content

Further to this core system, 3 external one-place operators can be added. The first one, I , is such that $I p$ is to be read as ‘ p is a piece of information’. p is information if and only if it is true semantic content. This essentially gives an operator that is the same as Bochvar’s assertion operator. The second operator, M , is such that $M p$ is to be read as ‘ p is a piece of misinformation’. p is misinformation if and only if it is false semantic content. The third operator, S , is such that $S p$ is to be read as ‘ p is a piece of semantic content’. If p is semantic content, then it is true or false. If p is meaningless, then it is not semantic content. All this gives the following:

f_I		f_M		f_S	
t	t	t	f	t	t
$*$	f	$*$	f	$*$	f
f	f	f	t	f	t

One can easily verify that the S operator formally satisfies the above list of properties regarding semantic content and connectives. For example, $S(p \wedge q) \equiv S p \wedge S q$.

1.3.3 Environmental Information

It is now time for a proper introduction to the notion of environmental information, which will be analysed later on in this thesis. The gist of environmental information is an easily familiar

one. When the doorbell in your home sounds, this auditory signal carries the information that someone is at the door. The presence of fingerprints at a crime scene carries the information that so and so participated in the crime. When the output of a square root function is 7, this carries the information that the input was 49. As can be gathered, environmental does not suggest *natural* here, but rather that the information results from connections within some environment or system. A general definition of environmental information is as follows:

Environmental information: Two systems a and b are coupled in such a way that a 's being (of type, or in state) F is correlated to b being (of type, or in state) G , thus carrying for the information agent the information that b is G . [73]

Or to put it another way, environmental information is a result of regularities that exist within a distributed system [16, p. 7].

As we will see, environmental information can be defined relative to an agent; what information a signal carries *for an agent* depends on what they already know, or what information they already have. In the examples above, there is a semantic element involved; the agents receiving the information are semantic agents who process the signal and give it a semantic significance. Yet it is important to emphasise that environmental information need not involve any semantics at all.

It may consist of (networks or patterns of) correlated data understood as mere differences or constraining affordances. Plants (e.g., a sunflower), animals (e.g., an amoeba) and mechanisms (e.g., a photocell) are certainly capable of making practical use of environmental information even in the absence of any (semantic processing of) meaningful data. [73]

A great example of this phenomenon is the insectivorous Venus flytrap plant. A Venus flytrap

lures its victim with sweet-smelling nectar, secreted on its steel-trap-shaped leaves. Unsuspecting prey land on the leaf in search of a reward but instead trip the bristly trigger hairs on the leaf and find themselves imprisoned behind the interlocking teeth of the leaf edges. There are between three and six trigger hairs on the surface of each leaf. If the same hair is touched twice or if two hairs are touched within a 20-second interval, the cells on the outer surface of the leaf expand rapidly, and the trap snaps shut instantly. [119]

The redundant triggering in this mechanism serves as a safeguard against a waste of energy in trapping inanimate objects that are not insects and have no nutritional value for the plant. If the trapped object is an insect chemical detection will occur and the plant will digest its prey.

In informational terms, the right succession of hair contact carries for the plant the information that there is a certain type of object moving along its leaves. The plant does not semantically process this information, but still makes essential use of the correlation between hair contact and object presence. So semantic information and environmental information are two separate conceptions of information. The former is meaningful data and thus requires semantics whilst the latter simply involves the regularity between two or more things in a system.

Whilst they are distinct notions, environmental information and information as semantic content are generally concurrent in our information systems and environments. For example, a properly functioning smoke alarm involves both types of information. An activated smoke alarm carries the environmental information that there is smoke (and possibly fire). That the smoke alarm is sensitive to smoke and beeps in response to its presence is why it carries this environmental information. Also, the smoke alarm's high-pitched beep signifies the presence of smoke and is an instance of semantic information.

Despite this general concurrence we must not forget to distinguish between the two separable notions of information. Replace the beep with an inaudible or undetectable signal and the alarm will still carry environmental information without providing semantic information. Damage the smoke alarm so that it malfunctions and frequently activates in cases where there is no smoke and the environmental information is lost whilst the semantic information remains.

Whilst the two types of information are independent, our dealings with environmental information throughout this work involve semantic information (given some *semantic* signal A and some fact B , A carries the information that B). Eventually the two will be linked up, so that semantic information will be defined such that it requires environmental information. I think that the following quote from Mingers is one way to express this idea: "A sign [signal] is caused by an event and carries that information. When it is taken, by an observer, as a sign of the event then it is said to have 'signification'. The sign signifies the causal event. This is essentially semantic information." [132, p. 6]

1.4 The Alethic Nature of Semantic Information

In the previous section a general definition of semantic information as well-formed, meaningful data (semantic content) was given, which led to the establishment of a propositional analysis of information. One extra aspect to consider is the alethic nature of information; does factual semantic content need to be true in order to qualify as semantic information or does any semantic information, true or false, count as information? According to Fox:

... ‘ x informs y that p ’ does not entail that p [and since] ... we may expect to be justified in extending many of our conclusions about ‘inform’ to conclusions about ‘information’ [it follows that] ... informing does not require truth, and information need not be true. [79, pp. 160–1, 189, 193]

Fox thus advocates some form of the Alethic Neutrality (AN) principle:

meaningful and well-formed data qualify as information, no matter whether they represent or convey a truth or a falsehood or have no alethic value at all. [66, p. 359]

In the discussion that follows consideration of factual semantic content that has no alethic value is generally disregarded. The prime issue here is whether or not information requires truth.

According to AN, since semantic content already qualifies as semantic information, the conditions of GDI are sufficient. Despite this, there has recently been some debate on the alethic nature of semantic information and a questioning of whether these conditions are sufficient. This debate was initiated by Floridi’s advocacy of a veridicality requirement for semantic information.¹⁵ According to the veridicality thesis (VT), in order for semantic content to be counted as information it must also be true: semantic information is well-formed, meaningful and *veridical/truthful* data. In other words, only true propositions count as genuine semantic information. Bear in mind that this veridicality requirement applies only to factual semantic content and not instructional semantic content, which is not alethically qualifiable.¹⁶

¹⁵See [66], [67] and [68]. Floridi himself acknowledges that this veridicality requirement is hardly a novel idea (some precedents are listed below). Nonetheless (re)ignition of the debate shows that discussion of the issue remains to be had.

¹⁶On a side note, given a dialethic picture a distinction arises between the veridicality thesis (i.e. that information must be true) and the non-falsity thesis (i.e. that information cannot be false) [10]. Take the following liar cycle: (A) The following sentence is true (B) The previous sentence is false. If (A) is both truth and false, can it be classed as information?

Other notable advocates of a veridicality condition for information are Dretske [51], Barwise and Seligman [16], Graham [87] and Grice [91], who offers the following direct characterisation of this position: “false information [misinformation] is not an inferior kind of information; it just is not information” [91, p. 371]. Thus the prefix ‘mis’ in ‘misinformation’ is treated as a negation.

In this dissertation the veridicality thesis is endorsed and semantic information is taken to be truthful semantic content. Admittedly, there is probably no objective fact about the world that will serve to decide this dispute. But whilst it might seem that the debate is just a trivial terminological one, there is arguably more to it. As I will show there is a host of good and legitimate reasons for adopting the veridicality thesis. Some will be covered in this section and some will unfold throughout this thesis.

Let us give this discussion some perspective by introducing a scale between 0 and 10. On this scale, we may identify three possible, mutually exclusive positions in the debate:

- A Genuine semantic information requires truth
- B Any legitimate conception of semantic information will not have truth as a requirement
- C There is more than one legitimate conception of semantic information. Some require truth and others don’t.

Let position C be located in the middle at 5, position A located at 10 and position B located at 0. I am confident that the arguments presented in this section and throughout this thesis¹⁷ suffice to establish that wherever we should settle on this scale, it should be no lower than 5. Furthermore, given the range of arguments offered, it is not fanciful to think that we are perhaps nudging a little towards 10.

My choice for adopting VT is largely motivated by the work to which I will put the notion of information I am employing. To begin with, an account according to which information encapsulates truth is more in line with the ordinary conception of factual information I am trying to capture; the sense in which information is a success word. Also, the specificity with which I am using the term will aid my investigation and the adoption of this position will provide a disambiguation of the terms ‘information’ and ‘misinformation’. Figure 1.3 shows the simple terminological hierarchy that will be employed in this thesis.

¹⁷Since each chapter relies in some way on the veridicality thesis, their successful development will contribute towards a case for it.

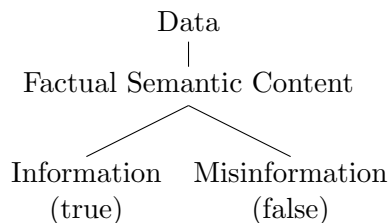


Figure 1.3: Terminological hierarchy

So factual/propositional semantic content consists of data, information is true semantic content and misinformation is false semantic content.¹⁸

Continuing on, as I will show there are at least several other technical or practical reasons for adopting VT. Here are a few of them:

- Standard traditional accounts of quantifying semantic information treat semantic information as alethically neutral and measure information as inversely related to probability. As a consequence, contradictions are problematically assigned maximal informativeness. VT paves the way for a method to quantitatively measure information in terms of truthlikeness and thus avoid such issues.
- VT facilitates attempts to define knowledge in terms of information.

A pragmatic rather than a principled motivation for understanding declarative, objective and semantic information as truthful, well-formed meaningful data derives from the constitutive role of acquiring information as a means to attain knowledge. Provided knowledge gets its usual factive reading instead of the ultra-loose sense it gets in information-science, it surely makes sense to apply the same veridical standard to information. [9, p. 5]¹⁹

- As suggested by the Gricean notion of natural meaning, environmental information is factive. If A carries the information that B , then if A is the case then B is the case. Given that semantic information will be linked with environmental information, this further implies the veridicality of such semantic information. Not only does being

¹⁸It is the structure and relationships of this hierarchy which are essential. In a way one could get away with making ‘information’ synonymous with ‘semantic content’ on the condition that they replace the left leaf ‘information’ with another (neologistic?) term which would then stand for what would then be true semantic content = true information

¹⁹Patrick Allo has raised in personal communication a related point that the VT advocate could make. Floridi stresses the ‘relational’ nature of information by comparison with food. Something does not count as food in general, but only for a certain type of organism: “it makes little sense to talk about the presence and nature of food without any reference to the specific type of feeder” [71, p. 163]. In this sense one could claim that food that isn’t edible by some organism just isn’t food for that organism. Similarly, misinformation just isn’t information for a cognitive agent interested in knowledge.

informed that B imply holding the true semantic content that B , but it also implies that B was acquired in a way that guarantees its truth, such as by receiving A .

Apart from my own considerations and motivations, there is a varied collection of arguments for and against the veridicality thesis, some of which we shall now look at. Let us begin with Dretske, whose thoughts on the matter I find pertinent and agreeable. Under his information-theoretic epistemology information is a necessary element of knowledge, so his aim is to formulate a theory of information where information, like knowledge, entails truth. He intends to respect some ordinary intuitions about what information is, where meaning (semantic content), which need not be true, is distinguished from information, which must be true. In his own words:

As the name suggests, information booths are supposed to dispense information. The ones in airports and train stations are supposed to provide answers to questions about when planes and trains arrive and depart. But not just any answers. True answers. They are not there to entertain patrons with meaningful sentences on the general topic of trains, planes, and time. Meaning is fine. You can't have truth without it. False statements, though, are as meaningful as true statements. They are not, however, what information booths have the function of providing. Their purpose is to dispense truths, and that is because information, unlike meaning, has to be true. If nothing you are told about the trains is true, you haven't been given information about the trains. At best, you have been given misinformation, and misinformation is not a kind of information anymore than decoy ducks are a kind of duck. If nothing you are told is true, you may leave an information booth with a lot of false beliefs, but you won't leave with knowledge. You won't leave with knowledge because you haven't been given what you need to know: information. [59, p. 2]

As pointed out by Dretske, other disciplines such as the computing and information sciences freely employ the term 'information' to refer to data or statements in general. In these cases truth seems to be irrelevant and anything that can be processed or stored in a database is counted as information. One reason

to maintain that nonnatural [semantic] false information is information too mirrors our reason to posit nonnatural information in the first place: it allows us to capture important uses of the term "information". It is only by tracking such disparate uses that we can make sense of the central role information plays in

the descriptive and explanatory activities of cognitive scientists and computer scientists, which partially overlap with the descriptive and explanatory activities of ordinary folk. [162, p. 323]

Whilst this may be the case, ‘important uses of the term’ are not necessarily justified or correct uses of the term. The term ‘vegetable’ has important culinary and cultural uses, although technically it is incorrectly applied in some cases. For example, eggplants and tomatoes are botanically speaking fruits, not vegetables. As I have suggested and will discuss further below, it seems preferable to develop a concept and terminology hierarchy in which data, semantic content and information are distinct. Under such a framework, cognitive, computer and information scientists would generally speaking traffic in data and semantic content, despite their use of the term ‘information’.

Whilst for computational purposes the data ‘Germany is in Europe’ and ‘Mexico is in Europe’ might be indistinguishable (they will be input, stored, manipulated and retrieved in the same way), it does not follow from this that the true datum counts as information if and only if the other does. Such an approach to information

blithely skates over absolutely fundamental distinctions between truth and falsity, between meaning and information. Perhaps, for some purposes, these distinctions can be ignored. Perhaps, for some purposes, they should be ignored. You cannot, however, build a science of knowledge, a cognitive science, and ignore them. For knowledge is knowledge of the truth. That is why, no matter how fervently you might believe it, you cannot know that Paris is the capital of Italy, that pigs can fly or that there is a Santa Claus. You can, to be sure, put these “facts”, these false sentences, into a computer’s data base (or a person’s head for that matter), but that doesn’t make them true. It doesn’t make them information. It just makes them sentences that, given the machine’s limitations (or the person’s ignorance), the machine (or person) treats as information. [59, p. 2]

Michael Dunn responds to the line of argument exemplified in the above quotes from Dretske with a clever counterexample:

I have heard a similar defense in a story of the “Information Booth” in a railway station and how it would be misnamed if it gave out false information. But note that I said “false information” in a very natural way. I think it is part of the pragmatics of the word “information” that when one asks for information, one expects to get true information, but it is not part of the semantics, the literal

meaning of the term. If there is a booth in the train station advertising “food”, one expects to get edible, safe food, not rotten or poisoned food. But rotten food is still food. [62]

As noted by Dunn, “false information” can be said in a very natural way. So what to make of the term ‘false information’? Well, Dunn’s point can be addressed with an argument originally made by Floridi [66], which offers a way to explain how it is that “false information” can be said in a very natural way whilst adopting VT and maintaining that “false information” is pseudo-information. The crucial distinction to make is that between *attributive* and *predicative* uses of “false”. Apparently this distinction was already known to medieval logicians and revived by Geach.

Take two adjectives like “male” and “good”. A male constable is a person who is both male and employed as a policeman. A good constable, however, is not a good person who is also employed as a member of the police force, but rather a person who performs all the duties of a constable well. “Male” is being used as a predicative adjective, whereas “good” modifies “constable” and is being used as an attributive adjective. On this distinction one can build the following test: if an adjective in a compound is attributive, the latter cannot be split up. This property of indivisibility means that we cannot safely predicate of an attributively-modified x what we predicate of an x . [66, p. 363]

Two further points are required to complete this argument:

1. some adjectives can be used attributively or predicatively depending on the context.
2. the attributive use can be either positive or negative. Positive attributively-used adjectives further qualify their reference x as y . Negative, attributively-used adjectives negate one or more of the qualities necessary for x to be x . They can be treated as logically equivalent to “not”.

Applying all of this to some examples, consider the following two compound terms:

1. false proposition
2. false policeman

A false proposition is still a proposition; it is something that is both false and a proposition. On the other hand, when we say that someone is a false policeman, we are in effect saying that they are not a policeman²⁰.

In Geachian terms, with 1 ‘false’ is being used predicatively and with 2 ‘false’ is a negative, attributively-used adjective. If p is a false proposition, we can split this compound term up “without any semantic loss or confusion”; ‘ p is false’ and ‘ p is a proposition’. If John is a false policeman, it makes no sense to split up the compound term. This distinction allows us to treat the term ‘false information’ like ‘false policeman’; when we say that something is false information, we are in effect saying that it is not information.

Admittedly, this argument does not really settle much. For example, take the proposition ‘the earth has two moons’. According to Floridi’s analysis, this is a false proposition in the predicative sense and a piece of false information in the attributive sense. But this

requires the brute intuition that that the earth has two moons is not information. The content of this intuition is nothing but an instance of the general thesis to be established. Thus, the argument is question-begging. No independent reason to reject instances of false information as information is given. Whether false information passes [this test] depends on whether we accept that a false p can constitute [semantic] information. We do! [162, p. 321]

This is a fair point. Furthermore, note that there is a difference between ‘false policeman’ and ‘false information’. With ‘false policeman’, we have established that ‘false’ is used attributively and is in effect a negation. If ‘false’ were to be used predicatively, then this would necessarily be a category mistake; policeman are not truthbearers. However if false were to be used predicatively in ‘false information’, this would not necessarily be a category mistake, since information as semantic content is a truthbearer.

So the claim that ‘false’ is attributive as opposed to predicative when it is applied to ‘information’ is unsettled and does not provide a conclusive argument for VT in its own right; it leaves us at a stalemate and other reasons are needed to support a definition of information with truth built in as a condition. This is related to the fact that the prefix ‘mis’ in ‘misinformation’ can be applied as a negation or to mean something like ‘ill’ or ‘wrong’. All this being said, I do though think that this argument revolving around the predicative/attributive distinction is an important one, because it, contra Dunn, allows proponents of VT to account for the term “false information” in a legitimate way. Given such a

²⁰‘false policeman’ is an example given by Floridi. To some English language readers this might seem like a strange phrase. A common equivalent way of saying this would be ‘fake policeman’. At any rate, there are other good example terms that could be used to illustrate this point, such as ‘false economy’.

definition adhering to VT, ‘information’ will be more like ‘tautology’. With ‘false tautology’ in the predicative sense, this is a contradiction in terms, since tautologies are by definition true. With ‘false tautology’ in the attributive sense, this means ‘not actually a tautology’, in the way that an intuitionist might say that $A \vee \neg A$ is a false tautology.

Apart from the overarching reasons for VT which form part of my thesis and will be covered in subsequent chapters, in the remainder of this section I would like to go over some extra, peripheral points.

Recall the Alethic Neutrality (AN) principle:

meaningful and well-formed data qualify as information, no matter whether they represent or convey a truth or a falsehood or have no alethic value at all.

From an unrestricted Alethic Neutrality principle it follows that:

- TA) tautologies qualify as semantic information
- FI) false information or misinformation (including contradictions) is a genuine type of semantic information, not pseudo-information

The acceptance or rejection of TA could go either way. As will be covered in Chapters 2 and 3, tautologies are not informative in that they do not provide any new information about the world; but neither do they misinform. Furthermore, in some sense tautological deductive inferences can also be said to yield information. So whilst one option is to reject TA because tautologies are never informative, it is perhaps also reasonable to represent tautologies as “instances of information devoid of any informativeness” [68, p. 36]. Whilst this can sound strange, consider it this way. If semantic information is defined as true semantic content, then since tautologies are instances of true semantic content, they are instances of information. But they are not informative given that an informational agent has this information ‘by default’, in the same way that a person’s name is not informative to them.

We now turn our focus to FI. Before looking at some reasons for rejecting FI and supporting a veridicality requirement, let us reject some reasons for supporting FI. The following discussion is based on a list found in [66].

1. *False information can include genuine information*

This simply means that such a piece of misinformation is a compound which includes one or more true constituents and it is only these constituents that qualify as information. For example, if A were true and B were true, then the statement $A \wedge \neg B$, although false, would still include some genuine information, namely A . As we will come to see in Chapter 2, a false statement can still have a positive measure of truthlikeness and yield some information.

2. *False information can entail genuine information*

Even if some true semantic content can be inferred from some false semantic content, what counts as information is the inferred true semantic content, not the false semantic content from which it was inferred. For example, if someone is *misinformed* that Barcelona is the capital of Spain, they can still infer the *information* that Barcelona is in Europe.

Also, contradictions logically entail every true statement, but we don't want to count contradictions as information.

3. *False information can support decision-making processes*

Just because false semantic content can be pragmatically useful this does not mean that it suffices to be counted as genuine information, much in the same way that mere true belief or even false belief can be pragmatically useful although they are not knowledge. Yes misinformation can lead to true beliefs, but such true beliefs are not knowledge. Knowledge requires something more, namely genuine information.

while false information might very well be presented as if it were a solid basis for knowledge, it can never be the stepping stone to knowledge that epistemological theorizing requires. [9, p. 5]

4. *If false information does not count as information, what is it? Assuming that p is false “if S only thinks he or she has information that p , then what does S really have? Another cognitive category beyond information or knowledge would be necessary to answer this question. But another cognitive category is not required because we already have language that covers the situation: S only thinks he or she has knowledge that p , and actually has only information that p .” [38, p. 468]*

Firstly, another cognitive category exists already, namely misinformation.

Secondly, contrary to the above quote and as we will see, the difference between knowing that p and being informed that p is not a difference regarding the truth or falsity of p .²¹

²¹[77] and [51] are two places where accounts of the distinction between information and knowledge are given.

Now that these reasons for supporting FI have been addressed, it is time to turn to some more reasons for supporting a veridicality requirement and rejecting FI.

To begin with, those who reject VT might prefer a modified form of FI, one which excludes contradictions so that only contingently false semantic content is included. This could lead to some issues though. An important question is how to measure the informativeness of contradictions. Standardly they are assigned an information yield measure of zero. But take the following principle of information aggregation, according to which the information yield (info) of two combined pieces of information is never lower than the information yield of either single piece: if I_1 and I_2 are instances of information, then $\text{info}(I_1 + I_2) \geq \text{info}(I_1)$ and $\text{info}(I_1 + I_2) \geq \text{info}(I_2)$. Now, if I_2 is the negation of I_1 , then their addition forms a contradiction. But if contradictions are assigned a measure of zero then this principle of information aggregation is violated. Thus it seems that the $\text{info}()$ measure would have to be a partial function such that $\text{info}(C)$ is undefined when C is a contradiction.

On a similar note, it is fair to say that information follows a principle of conjunction: for any two propositions A and B , if A is information and B is information, then the compound proposition $A \wedge B$ is information. However, this principle together with the modified FI leads to the problematic result that when A is information and $\neg A$ is information, $A \wedge \neg A$ is both information and not information.

Floridi argues that if any type of well-formed, meaningful data counts as information, “information becomes semantically indestructible and the informative content of a repository can decrease only by physical and syntactical manipulation of data” [66]. Take a tourist information pamphlet. If it were to be shredded, then the information contained in the pamphlet would be destroyed by physical manipulation. If the text of the pamphlet were to be jumbled up and randomly rearranged, then the information contained in the pamphlet would be destroyed by syntactical manipulation.

But it would also seem that information can be semantically depleted. A good example of this which comes to mind involves the changing of a datum in a database, and how this change can affect the alethic value of another datum. If, on Tuesday January 1 2013 it is actually raining and a database contains the datum ‘today is Tuesday January 1 2013’ and ‘today it is raining’, both pieces of data count as information. If on Wednesday January 2 2013 it is not raining and the database contains the datum ‘today is Wednesday January 2 2013’ whilst still holding the unrevised datum ‘today it is raining’, then the latter is no longer information, and *ceteris paribus* there is an information decrease. The point here is that collections of information should be sensitive to factual changes, with the possibility of semantic decrease as well as increase, rather than just any piece of semantic content indiscriminately being

considered information. The qualification of truth thus gives semantic information an extra quantitative dimension.

An attempt to account for this semantic loss whilst rejecting VT is given by Scarantino and Piccinini [162, p. 321]. Basically, under their approach the notion of semantic loss can be accommodated if treated as a *qualitative* rather than *quantitative* loss. They give an example where all the true propositions in a chemistry manuscript are transformed into their negations. According to their approach, such a situation would involve loss of the original information, in that “the information-carrying vehicles in the repository no longer carry the same information they used to carry” [162, p. 322]. Although the new information would be of the same amount, there would also be a qualitative semantic loss of information. Also, the resulting information would be of a lower *epistemic value*: negating a true proposition causes information loss by semantic means since false information is epistemically inferior to true information. As they sum up:

rejecting VTNN [VT for semantic information] is compatible with accounting for information loss “by semantic means” in the two senses - the qualitative and epistemic-value senses - that matter most for epistemic purposes. Moreover, our distinctions allow us to neatly distinguish between physical and syntactic information loss on the one hand and semantic information loss on the other. In the first two cases, information is destroyed but not replaced with any new nonnatural [semantic] information. There is information loss in the quantitative sense, in the qualitative sense, and in the epistemic-value sense. In the third case, information is destroyed and replaced with new (false, and thus epistemically inferior) nonnatural information. There is information loss in the qualitative sense and in the epistemic-value sense, though not in the quantitative sense. [162, p. 323]

It would seem then that we are once again at a stalemate; there are those such as Scarantino and Piccini who think that misinformation is just an inferior kind of information and those such as Floridi and Grice who think that misinformation is not an inferior kind of information, it is just not information. Depending on the position, information loss by semantic manipulation can be characterised as quantitative or qualitative. For those adhering to VT, quantitative semantic loss occurs when true semantic content is replaced by false content or when certain semantic content is weakened, for example, when a true conjunction is replaced by its corresponding disjunction.²² Qualitative semantic loss would occur when one truth is replaced by another truth which is less valuable in some sense. For those rejecting

²²This second type will be formally explicated and made precise in Chapter 2.

VT, quantitative semantic loss presumably occurs only when semantic content is logically weakened ²³. On the other hand, qualitative semantic loss would occur when one truth is replaced by another truth which is less valuable in some sense or when true content is replaced by false content.

This completes the argument concerning semantic loss. One further consideration in favour of a veridicality condition for information involves instructional information; basically, VT will give a tighter link between factual and instructional information. For something to be instructional information to bring about some state of affairs x , the instructions need to be correct, in that their execution will actually bring about x . As outlined in Section 1.3.2, this instructional information can be converted to a corresponding piece of factual information. Say we have some instructional information for task t consisting of three instructions t_1 , t_2 and t_3 . Converting this to its corresponding piece of factual information, we get ‘Task t can be accomplished by carrying out the following instructions: t_1 , t_2 and t_3 ’. Now take a piece of false semantic content: ‘Task t can be accomplished by carrying out the following instructions: t_1' , t_2' and t_3' ’. Converting this false content to a set of instructions using the reverse of this method we get a set of instructions which will not bring about the indicated state of affairs. Does this incorrect set of instructions constitute an instance of instructional information? Some, particularly proponents of VT, might say that if a set of instructions is to constitute an instance of genuine instructional information then the instructions must be correct. If this is the case, then we have gone from a case of information (false semantic content) to non-information (incorrect instructions). On the other hand, some, particularly opponents of VT, might say that incorrect instructions still constitute instructional information because they still bring about some state of affairs. Instructions for freezing water that say to place it in a kettle are incorrect but still result in the water being boiled. In this way the false factual semantic content ‘Water can be frozen by placing it in a kettle’ and its corresponding incorrect set of instructions are both instances of information. However this position is significantly undermined by instances of false factual semantic content that correspond to instructions that bring about an incoherent state of affairs or no state of affairs at all. For example, take the following instructions, which simply cannot be executed: in order to get from Moscow to New York, take a boat from Moscow to Vienna and then drive from Vienna to New York. It is fair to say that these instructions are not *instructional information* since they cannot be executed at all. Or further still, take the following set of instructions: in order to activate the machine, press the button for no more than one second and no less than three seconds. Clearly these inconsistent instructions bring about no state of affairs at all and so do not constitute an instance of genuine instructional information. But ‘junk’ instructions such as these can be converted to some piece of false factual semantic

²³Scarantino and Piccini would still need a way to make sense of quantifying semantic information. Presumably they would adopt an inverse probabilistic content approach such as that outlined in the next chapter. As we will see, this also has some issues.

content. To treat such false content as information would mean that it is possible to move from information to non-information and vice-versa. This would in a sense deprive one's account of a certain connection between factual and instructional information, which would otherwise hold if only correct instructions and true factual semantic content were treated as information. Putting this line of thought more precisely, take something like the following principle:

1. if x is a form of information and a sound operation is performed on it to get y , then y is also a form of information.

This principle is plausible. Something like it is already at play in the case of logical deduction: if A is true (information) and B is soundly inferred from A , then B is true (information). But if false semantic content were classed as information then this principle would be violated. Given this, false semantic content should not be classed as information. This modest argument simply adds one more reason to accept VT and one more reason to question the rejection of VT.

In closing this discussion on the alethic nature of information, I would like to begin by re-emphasising that although 'information' is a flexible term, this flexibility is not unbounded. Caution must be exercised in its employment lest it become an indistinct 'wildcard word' and any appeal to the notion of information should be justified through argument and application.

In the case of 'false information', a good guide to explain and accommodate its usage can be found in the following quote:

[In other cases, people are] talking about information in a non-semantic sense; some other times, they may just be using a familiar synecdoche, in which the part (semantic information) stands for the whole (semantic information and misinformation), as when we speak in logic of the truth-value of a formula, really meaning its truth or falsehood. Often, they are using information as synonymous for data, or representations, or contents, or signals, or messages, or neurophysiologic patterns, depending on the context, without any loss of clarity or precision. [70, p. 406]

For our purposes an efficient and unambiguous terminological hierarchy such as that presented in Figure 1.3 is desirable. Irrespective though of whether the term 'information' is defined as semantic content or true semantic content, what is fundamental is the structure

of Figure 1.3 and it is essential that the four categories are identified and mapped to specific terms. So even if ‘information’ was mapped to the category of semantic content, the vocabulary would need to be completed with a direct term for true semantic content; there would be a hole in the vocabulary if it were to have ‘information’ for semantic content in general and ‘misinformation’ for false semantic content, but not have a direct term for true semantic content. Not only is this to have a specific word for an important category, but a terminology whereby false semantic content has a direct term but true semantic content does not would be undesirably asymmetric.

In conclusion, a strong case has been given for adopting the veridicality thesis and establishing a legitimate conception of *premium* semantic information which encapsulates truth (we are at least at 5 on the scale given at the beginning of this section). Thus this conception of information justifiably stands amongst the myriad of conceptions of information. We defer to Floridi for the final word:

Bananas are not fruit, and tomatoes are not vegetables, but we know where to find them in the supermarket, and not even a philosopher should complain about their taxonomically wrong locations. ... Now, if one wishes to talk rather loosely of information from the beginning to the end of this journey and all the way through it, that is fine. We know our way in the supermarket, we can certainly handle loose talk about information. There is no need to be so fussy about words: the tomatoes will be found next to the salad, and the bananas next to the apples. ... from such a “supermarket approach”, the veridicality thesis is untenable, since truth or falsehood plays absolutely no role in “information” for a long while during the journey from electromagnetic radiation to “Sorry, dear, the battery is flat”. Of course, this leaves open the option of being conceptually careful when dealing with semantic information itself, the end product of the whole process. Botanically, tomatoes and courgettes are fruit, and bananas are female flowers of a giant herbaceous plant. Likewise, in philosophy of information semantic information is well formed, meaningful and truthful data. If you still find the veridicality thesis as counterintuitive as the fruity tomatoes, just assume that [those who endorse it] are being botanically precise and talking about *premium semantic information*. [70, p. 408]

1.5 Conclusion

We now come to the end of this first chapter, in which a general introduction to the ‘conceptual labyrinth’ that is information was provided. Further to this, the conceptions of information central to this thesis were identified and established. In the following chapters a framework for (semantic) information based on these conceptions will be developed. The next two chapters are closely linked and will show how information as true semantic content can be logically represented and demonstrate how some of its properties can be formally accounted for. In particular, the second chapter ‘Quantifying Semantic Information’ investigates ways to measure the notion of information quantity/informativeness. In line with VT, traditional inverse probabilistic approaches to quantifying semantic information are replaced with approaches based on the idea of truthlikeness. In the third chapter ‘Agent-Relative Informativeness’ results from the second chapter are combined with belief revision to construct a formal account of measuring how informative some piece of information is to a given agent. The fourth chapter ‘Environmental Information and Information Flow’ analyses several existing accounts of environmental information and information flow before using this investigation to develop a better account of and explicate these notions. Finally, with contributions from some of the results from the fourth chapter, the fifth chapter ‘Information and Knowledge’ contributes towards the case for an informational epistemology. In developing these chapters, the significance of these forms of semantic information and their applicability will be demonstrated. Importantly, they will be distinguished from related notions such as data, meaning and evidence.

Chapter 2

Quantifying Semantic Information

In this chapter we look at some existing methods of semantic information quantification and suggest some alternatives. We begin with an outline of Bar-Hillel and Carnap’s theory of semantic information before going on to look at Floridi’s theory of strongly semantic information. The latter then serves to initiate an in-depth exploration into the idea of utilising the notion of truthlikeness to quantify semantic information. The main outcomes of this investigation are:

- a few approaches to measuring truthlikeness are drawn from the literature and investigated, with a focus on their applicability to semantic information quantification.
- a new approach to measure truthlikeness/information is presented.
- the extension of these methods to the quantification of misinformation.
- a method to estimate truthlikeness/information is given.

Continuing on from the previous chapter, semantic information is understood in terms of propositions and represented using statements in a basic propositional logic (*i.e.*, the information that p). Furthermore, the approaches looked at can be seen to fall under the logical approach to semantic information theory [25, p. 315].

Semantic information is intuitively something that can be quantified. Judgements such as ‘statement A yields more information or is more informative than statement B ’ are naturally made¹. The motivations for and the aims of this work on ways to quantify semantic information can be easily appreciated. The following cases and considerations will serve to illustrate:

¹The terms ‘informativeness’ and ‘information yield’ are used synonymously in this investigation.

- Imagine a situation in which a six-sided die is rolled and it lands on 4. Given the following collection of statements describing the outcome of the roll, which is the most informative? Which is the least informative?

- The die landed on 1
- The die landed on 1 or 2 or 4
- The die landed on 4
- The die landed on 4 or 5
- The die did not land on 4

A formal framework for the quantification of semantic information could assign a numerical measure to each statement and rank accordingly.

- Take a simple domain of inquiry, involving the following two questions:
 - Is Berlin the capital city of Germany? (A)
 - Is Paris the capital city of France? (B)

Now, there is a coherent sense in which the following ranking of answer statements in terms of highest to lowest informativeness is right:

1. $A \wedge B$
2. A
3. $A \vee B$
4. $\neg A \wedge \neg B$

$A \wedge B$ precisely describes the truth of the domain, it gives the correct answer to both questions. A is also true but only provides information about one of the items, giving the correct answer to the first question. $A \vee B$ only tells us that the answer to at least one of the questions is ‘yes’, but not which one. $\neg A \wedge \neg B$ on the other hand is false and would provide an incorrect answer to both questions.

What is a suitable formal framework for the quantification of semantic information that can rigorously capture these intuitions and provide a way to measure the complete range of statements within a domain of inquiry?

- Take a more general example of a logical database consisting of a collection of facts. How can the information content of the database be measured? Why does one database contain more information than another?
- Information is often treated as a commodity and one factor that will determine the value of a piece of information is going to be its quantitative measure. A formal method to measure the information yield of statements could therefore contribute to its informational valuation.

Thus motivations and aims for the task of this chapter are clear.

2.1 Bar-Hillel and Carnap's Theory of Semantic Information

Bar-Hillel and Carnap's seminal account of semantic information [15, 14] measures the information yield of a statement within a given language in terms of the set of possible states it rules out and a logical probability space over those states.

The general idea is based on the *Inverse Relationship Principle*, according to which the amount of information associated with a proposition is inversely related to the probability of that proposition. This account will henceforth be referred to as the theory of *Classical Semantic Information* (CSI) [163]. Using some *a priori* logical probability measure $\text{Pr}()$ on the space of possible states, two measures of information [$\text{cont}()$ and $\text{inf}()$] are provided, such that:

$$\text{cont}(A) =_{df} 1 - \text{Pr}(A)$$

and

$$\text{inf}(A) =_{df} -\log_2(\text{Pr}(A))$$

In order to demonstrate these definitions, take a very simple propositional logical space consisting of three atoms ². The space concerns a weather framework, where the three properties of a situation being considered are whether or not it will be (1) hot (h), (2) rainy (r) and (3) windy (w) ³. Since there are three variable propositions involved, there are eight logically possible ways things can be, there are eight possible states. A truth table for this is shown in table 2.1:

Now, each possible state can be represented using a *state description*, a conjunction of atomic statements consisting of each atom in the logical space or its negation, but never both. For example, the state description for w_1 is $h \wedge r \wedge w$ and the state description for w_8 is $\neg h \wedge \neg r \wedge \neg w$. The probability distribution used will be a uniform logical one; that is, since there are eight possible states, each state has a probability of $\frac{1}{8}$ that it will obtain

²The spaces considered in this work are finite

³This is based on a canonical example in the truthlikeness literature, which will be looked at in Section 2.4

State	h	r	w
w ₁	T	T	T
w ₂	T	T	F
w ₃	T	F	T
w ₄	T	F	F
w ₅	F	T	T
w ₆	F	T	F
w ₇	F	F	T
w ₈	F	F	F

Table 2.1: Truth table for 3-proposition weather example.

or be the actual state ⁴. Using these parameters, some results are listed in table 2.2. In a space with n atoms, there are $n^2 + 1$ different classes of information yield. Table 2.3 lists the range of classes for $n = 3$.

Note that with the formulas for $\text{cont}()$ and $\text{inf}()$, it is clear that the following relationships hold between logical entailment and semantic information, simply due to the fact that $A \vdash B \Rightarrow \text{Pr}(A) \leq \text{Pr}(B)$:

- $A \vdash B \Rightarrow \text{cont}(A) \geq \text{cont}(B)$
- $A \vdash B \Rightarrow \text{inf}(A) \geq \text{inf}(B)$

Statement (A)	$\text{cont}(A)$	$\text{inf}(A)$
$h \wedge w \wedge r$	0.875	3
$\neg h \wedge \neg w \wedge \neg r$	0.875	3
$h \wedge r$	0.75	2
$h \vee w$	0.25	0.415
$h \vee \neg h$	0	0
$h \wedge \neg h$	1	∞

Table 2.2: Results using CSI measures

Bar-Hillel and Carnap also provided some formulas to measure relative information and this work was subsequently expanded upon by Jaakko Hintikka. Amongst such varieties of information, firstly, there is *incremental information* (add), which measures “the change in the informational status of a proposition $[A]$ brought about by another one $[B]$ ” [106, p. 2]. Definitions are as follows:

⁴Other, disuniform probability distributions could be used. For example, if $\text{Pr}(h) > \text{Pr}(\neg h)$, then *ceteris paribus* the measure for $\neg h \wedge r$ would be greater than the measure for $h \wedge r$. A uniform logical distribution is standard and appropriate for measuring pure content; when it is uniform two statements of the same logical strength have the same measure. When different atoms and states are assigned different probabilities then the inverse relationship between probability and information means that some statement A that is logically weaker than some statement B could nonetheless be assigned a higher information measure. Carnap discusses various types of more sophisticated logical probability measures. See [108] for more.

#states true in	cardinality	standard cont measure	inf measure
0	1	1	∞
1	8	0.875	3
2	28	0.75	2
3	56	0.625	1.415
4	70	0.5	1
5	56	0.375	0.678
6	28	0.25	0.415
7	8	0.125	0.193
8	1	0	0

Table 2.3: Classes of information for CSI with $n = 3$

- $\text{cont}_{add}(A|B) =_{df} \text{cont}(A \wedge B) - \text{cont}(B) = \text{cont}(B \supset A)$
- $\text{inf}_{add}(A|B) =_{df} \text{inf}(A \wedge B) - \text{inf}(B)$

Secondly there is *conditional information* (cond), “the informational status of a proposition A on the assumption that we know another one, say B ” [106, p. 2]. Definitions are as follows:

- $\text{cont}_{cond}(A|B) =_{df} 1 - \text{Pr}(A|B)$
- $\text{inf}_{cond}(A|B) =_{df} -\log_2(\text{Pr}(A|B))$

Given these definitions, it so happens that $\text{inf}_{add}(A|B) = \text{inf}_{cond}(A|B)$, so the subscripts can be dropped. No similar equivalence holds for $\text{cont}_{add}(A|B)$ and $\text{cont}_{cond}(A|B)$.

A third sense of relative information is *transmitted information*, which “measures the information B conveys concerning A ” [106, p. 3]. Definitions are as follows:

- $\text{transcont}_{add}(A|B) =_{df} \text{cont}(A) - \text{cont}_{add}(A|B) = 1 - \text{Pr}(A \vee B)$
- $\text{transcont}_{cond}(A|B) =_{df} \text{cont}(A) - \text{cont}_{cond}(A|B) = \text{Pr}(A|B) - \text{Pr}(A)$
- $\text{transinf}(A|B) =_{df} \text{inf}(A) - \text{inf}(A|B) = \log_2\left(\frac{\text{Pr}(A|B)}{\text{Pr}(A)}\right)$

As can be seen, these measures are clearly related to Bayesian measures of confirmation. For example, $\text{transcont}_{cond}(A|B)$ simply equates to the confirmation theoretic measure associated with Carnap [26] and $\text{transinf}(A|B)$ equates to the confirmation theoretic measure associated with Keynes [115].

Amongst other contributions made by Hintikka to this school of semantic information quantification are:

- Extension of the CSI account to full polyadic first-order logic, making use of his distributive normal forms.
- Introduction of his distinction between *surface* information and *depth* information. The latter is defined similarly to the approach used by Bar-Hillel and Carnap. The former is used to address the ‘scandal of deduction’: according to Bar-Hillel and Carnap’s account logical truths or proofs of logical validity give no information. In order to accommodate a sense in which they do give information, Hintikka developed a way to measure the information yield of deductive inferences.

Readers further interested in Hintikka’s contributions can consult [103, 104, 106].

2.1.1 Some Comments on the Theory of Classical Semantic Information

Does the CSI account overall provide an acceptable measure of semantic information, in line with the motivations and aims discussed in the introduction? To a certain extent its results accord with our intuitions and are expected. True state descriptions yield much information; a statement which narrows things down to just one of the possible states yields a lot of information. By contrast, if a statement tells us that it will be hot and rainy but tells us nothing about whether or not it will be windy ($h \wedge r$), it would yield less information, as the calculations agree. Further still, a statement which only says that it will be hot or windy ($h \vee w$) yields relatively little information. Finally, a tautological statement, which gives no reduction of the original eight possibilities, appropriately yields no information or is not informative. Despite these results, there are significant cons that indicate the inadequacy of the CSI account in accommodating certain criteria associated with the type of quantitative account of semantic information that is in line with present motivations and aims.

The most prominent issue concerns its assignment of maximal information yield to contradictions, what has elsewhere been termed the *Bar-Hillel-Carnap Paradox* [73]. Can no non-contradiction be more informative than a contradiction? Surely this is not the case. Furthermore, do contradictions yield any information at all? These questions will be discussed shortly.

Bar-Hillel and Carnap provide the following commentary on the matter:

It might perhaps, at first, seem strange that a self-contradictory sentence, hence one which no ideal receiver would accept, is regarded as carrying with it the most inclusive information. It should, however, be emphasised that semantic information is here not meant as implying truth. A false sentence which happens to say much is thereby highly informative in our sense. Whether the information it carries is true or false, scientifically valuable or not, and so forth, does not concern us. A self-contradictory sentence asserts too much; it is too informative to be true [15, p. 229].

No doubt the notion of information expounded in the above passage is at odds with the ordinary sense of information as discussed in the introduction. To begin with, ‘too informative to be true’ does sound somewhat odd. It is fair to say that the more a sentence says the more informative it is. But when does a sentence ‘say much’? We intuitively judge the statement $A \wedge B$ to say more than the statement $A \vee B$ not just because it is less probable or excludes more possible states, but also because it does a better, more detailed job of describing how presumably things are. For any two true statements A and B such that $\Pr(A) < \Pr(B)$, it is fair to say that A yields more information than B . On the other hand, not only is $A \wedge \neg A$ false, but it does not at all do a good job of describing how things presumably are or could be. It does not discriminate and selectively narrow down on a potential state of affairs (unless a contradiction does actually occur!).

Further to this issue, the CSI account’s indifference to truth and falsity means that it cannot distinguish between true and false statements with the same probability. If it actually is hot, rainy and windy (*i.e.*, $h \wedge r \wedge w$ is the true state description), then in the sense of information we are interested in, the statement $h \wedge r \wedge w$ is more informative than the statement $\neg h \wedge \neg r \wedge \neg w$. Even the statement $h \wedge r$, which has a higher probability, is more informative than $\neg h \wedge \neg r \wedge \neg w$, since it contains more truth. These considerations suggest that at the least semantic information and informativeness is not just about content and should incorporate truth.

2.1.2 The Bar-Hillel-Carnap Paradox and Paraconsistent Logic

Before continuing on, a brief aside is in order. Paraconsistent logics [149] do not validate the classical validity $A \wedge \neg A \vdash B$. So not everything follows from a contradiction and systems built upon such logics are able to successfully tolerate inconsistency. Given that they are able to successfully handle contradictions, might the use of a paraconsistent logic resolve the Bar-Hillel-Carnap Paradox (BCP)? As suggested by Allo, “without going into the technicalities of different paraconsistent logics, it is clear that (BCP) is no longer a necessary conclusion

$f_{\neg}$		$f_{\wedge}$	t	b	f	$f_{\vee}$	t	b	f
t	f	t	t	b	f	t	t	t	t
b	b	b	b	b	f	b	t	b	b
f	t	f	f	f	f	f	t	b	f

Figure 2.1: Truth tables for LP

of [the $\text{cont}()$ measure]” [8].

Whilst seemingly a promising idea, it is far from certain that such a move would be successful. To begin with, as evidenced by the following demonstration, paraconsistency *per se* is not a solution to BCP, so any successful strategy would require a particular paraconsistent logic.

The paraconsistent logic *Logic of Paradox* (*LP*), a many-valued logic with a possible worlds/states reading, captures the idea of paraconsistency in a simple yet fundamentally effective way. This logic has three truth values; the classical ‘true’ (t) and ‘false’ (f) plus the value ‘both’ (b), which represents ‘both true and false’. t and b are the designated values. Figure 2.1 contains truth tables for the negation, conjunction and disjunction connectives in this logic:

Given its truth-tabular semantics, *LP* is easily applied to form the basis of a paraconsistent probability function and thus is easily amenable to classical semantic information calculations. Now, take the following truth table:

p	q	$p \wedge \neg p$	$p \wedge q$	$(p \wedge \neg p) \vee (q \wedge \neg q)$
t	t	f	t	f
t	b	f	b	b
t	f	f	f	f
b	t	b	b	b
b	b	b	b	b
b	f	b	f	b
f	t	f	f	f
f	b	f	f	b
f	f	f	f	f

Table 2.4: Truth table for LP example

The basic contradiction $p \wedge \neg p$ still yields the most information according to calculations. Here $\text{cont}(p \wedge \neg p) = 1 - \frac{3}{9} = \frac{6}{9}$. In fact, for $n = 2$ the highest possible $\text{cont}()$ value for a classically satisfiable statement is $\text{cont}(p \wedge q) = 1 - \frac{4}{9} = \frac{5}{9}$; there is no classically satisfiable statement S such that $\text{cont}(S) > \frac{5}{9}$, Here’s why:

- Define a classical state as one with only t and f atomic valuations ($v(a) = t$ or $v(a) = f$ for all atoms a). If a formula is classically satisfiable, then it is true in at least one classical possible state.
- Each classical possible state C has three non-classical possible states (states with b valuations) that correspond to it. A non-classical state N corresponds to C if and only if N can be obtained by replacing one or more valuations of C with b . For example, with reference to the above example in Table 2.4 take the classical state such that $v(p) = t$ and $v(q) = t$. Its three corresponding non-classical states are:

1. $v(p) = t, v(q) = b$
2. $v(p) = b, v(q) = t$
3. $v(p) = b, v(q) = b$

- If non-classical world N corresponds to classical world C , it means that no classically satisfiable statement can distinguish between C and N ; statement S receives a designated value in C if and only if it receives a designated value in N . Here's why:

- Convert any LP formula to conjunctive normal form [e.g. $p_1 \wedge (p_2 \vee p_3 \vee p_4)$]. If a statement is classically true, then all of its conjuncts are true and in turn for every conjunct at least one of the disjuncts is true. This statement is also designated in all the non-classical corresponding alternatives. With every true conjunct, get one of the disjuncts d that is true. Its classical valuation of $v(d) = t$ can be either:

1. left alone to remain the same, if $v(d) = t$ is in the non-classical alternative.
2. replaced by $v(d) = b$, in which case d is still designated.

If it is of the form $\neg d$, then its classical valuation of $v(d) = f$ can be either:

1. left alone to remain the same, if $v(d) = f$ is in the non-classical alternative.
2. replaced by $v(d) = b$, in which case $\neg d$ is still designated.

- Therefore, consistent formulas will always be satisfied in at least one more state than some contradiction.

It can be seen that this demonstration simply generalises to all n .

Despite this result, it is possible for some contradictions to yield less information than some non-contradictions. For example, with reference to the above example in Table 2.4, $\text{cont}((p \wedge \neg p) \vee (q \wedge \neg q)) = 1 - \frac{5}{9} = \frac{4}{9} < \text{cont}(p \wedge q) = \frac{5}{9}$. To find a formula of the form $\phi \wedge \neg \phi$ such that $\text{cont}(\phi \wedge \neg \phi) < \frac{5}{9}$, ϕ must be a compound formula. Here is an example: $\text{cont}(((p \wedge \neg p) \vee (q \wedge \neg q)) \wedge \neg((p \wedge \neg p) \vee (q \wedge \neg q))) = \frac{4}{9}$.

Whilst this demonstration casts doubt on the viability of a paraconsistency solution to BCP, more work remains to be done on investigating the usage of other paraconsistent logics.

2.2 CSI and Truth

Given our conclusions thus far, it is time to start looking into combining CSI with the factor of truth.

2.2.1 Epistemic Utility

Both truth and content are epistemic utilities of inquiry, but alone neither is adequate. If truth were the only criterion, then a tautology would be just as good as any other true statement. If content were the only criterion then a false state description would be just as good as a true state description. So it is clear that a balance between the two should be struck.

Isaac Levi [122] has proposed a measure of epistemic utility $[U()]$ that is a weighted combination of the truth value of some statement s $[tv(s)]$ and the information content of s $[I(s)]$. Setting $I(s)$ as $cont(s)$, the resulting measure is:

$$U_1(s) = acont(s) + (1 - a)tv(s)$$

where $0 < a < \frac{1}{2}$ is a parameter set to reflect “how much the scientist is willing to risk error, or to ‘gamble with truth,’ in his attempt to be relieved from agnosticism” [136].

Adding to this formula, given some piece of evidence e , the expected epistemic utility of s is:

$$U_1(s, e) = acont(s) + (1 - a)Pr(s|e)$$

Similar to these formulations, according to the following measure [136], the content of s is gained if s is true whilst the content of the true $\neg s$ is lost if s is false:

$$U_2(s) = \begin{cases} cont(s) & \text{if } s \text{ is true} \\ -cont(\neg s) & \text{if } s \text{ is false} \end{cases}$$

Expected epistemic utility using this formulation becomes:

$$U_2(s, e) = \Pr(s|e) - \Pr(s)$$

Table 2.5 lists some examples using these measures.

Statement	$U_1(a = 0.5)$	U_2
$h \wedge r \wedge w$	0.9375	0.875
h	0.75	0.5
$h \vee r \vee w$	0.5625	0.125
$\top$ ⁵	0.5	0
$h \wedge r \wedge \neg w$	0.4375	-0.125
$\neg h \wedge \neg r \wedge \neg w$	0.4375	-0.125
$h \wedge \neg r$	0.375	-0.25
$\neg h \wedge \neg r$	0.375	-0.25
$\neg h \vee \neg r \vee \neg w$	0.0625	-0.875

Table 2.5: Epistemic Utility Sample Results

Whilst these measures might be good indicators of epistemic utility, it is clear that their orderings and lack of discrimination amongst false statements makes them inadequate to serve as measures of information as true semantic content.

2.2.2 CSI, Truth and Scoring Rules

Table 2.6 depicts a rough conceptual principle that can be extracted from notion of truth implying information we are trying to capture. The four possible ways of combining the two factors under consideration are represented using a standard 2x2 matrix, with each possible way being assigned a unique number between 1 and 4. 1 represents ‘best’ for information yield and 4 represents ‘worst’.

	More Truth than Falsity	More Falsity than Truth
High Reduction of Possibilities	1	4
Low Reduction of Possibilities	2	3

Table 2.6: Possibility reduction and truth/falsity combinations.

Therefore, narrowing down the range of possibilities is a good thing for information yield when it leads to truth. On the other hand, if it leads to falsity it is a bad thing. Borrowing an

⁵As we will see, the mid-range placement of tautologies is a feature common to measures that combine truth with content.

idea from decision theory, this type of reward/punishment system can be formally captured with a simple *scoring rule*. Take the following, where for a statement A :

- if A is true then give it an information measure of $\text{cont}(A)$
- if A is false then give it an information measure of $-\text{cont}(A)$

Whilst this approach would give acceptable results for true statements, when it comes to false statements it is too coarse and does not make the required distinctions between the different classes of false statements. A state description in which every atom is false is rightly assigned the lowest measure. Furthermore, other false statements which are true in the state corresponding to that false state description will be assigned acceptable measures.

But this approach also has it that any false state description will be accorded the lowest measure. So the state description $h \wedge r \wedge \neg w$ (w_2) would be assigned the same measure as $\neg h \wedge \neg r \wedge \neg w$ (w_8), which is clearly inappropriate. Furthermore, equal magnitude supersets of the states correlating to these state descriptions (e.g., w_2, w_4 and w_4, w_8) would also be assigned equal measures. Therefore, whilst the orderings amongst false statements is better, sufficient discrimination is still lacking.

Given the aim to factor in considerations of truth value, the problem with any account of information quantification based on a CSI-style inverse probabilistic approach is related to the failure of Popper's content approach to truthlikeness. As Niiniluoto nicely puts it:

among false propositions, increase or decrease of logical strength is neither a sufficient nor necessary condition for increase of truthlikeness. Therefore, any attempt to define truthlikeness merely in terms of truth value and logical deduction fails. More generally, the same holds for definitions in terms of logical probability and information content. [134, p. 296]

2.3 Floridi's Theory of Strongly Semantic Information

Luciano Floridi has developed a *Theory of Strongly Semantic Information* (TSSI), which differs fundamentally from CSI. It is termed 'strongly semantic' in contrast to Bar-Hillel and Carnap's 'weakly semantic' theory because unlike the latter, where truth values do not play a role, with the former semantic information encapsulates truth.

Floridi's basic idea is that the more accurately a statement corresponds to the way things actually are, the more informative it is. Thus information is tied in with the notion of *truthlikeness*. If a statement perfectly corresponds to the way things actually are, if it completely describes the truth of a domain of inquiry, then it is maximally informative. Then there are two extremes. On the one hand, if a statement is necessarily true in virtue of it being a tautology, its informativeness is zero. On the other, if a statement is necessarily false in virtue of it being a contradiction, its informativeness is also zero. Between these two extremes and perfect correspondence there are contingently true statements and contingently false statements, which have varying degrees of informativeness.

Using the weather framework introduced above, let the actual state be w_1 . The following ranking of propositions, from highest to lowest informativeness illustrates this idea:

1. $h \wedge r \wedge w$
2. $h \wedge r$
3. $\neg h \wedge \neg r$
4. $\neg h \wedge \neg r \wedge \neg w$
5. $h \wedge \neg h, h \vee \neg h$

We now come to take a look in more detail at Floridi's account as given in [65]. The informativeness of each statement s ⁶ is evaluated as a function of (1) the truth value of s and (2) the degree of semantic deviation (degree of discrepancy) between s and the actual situation w . This degree of discrepancy from the actual situation is measured by a function f , which takes the statement as input and outputs some value in the interval $[-1, 1]$. This allows for the expression of both positive (when s is true) and negative (when s is false) degrees of discrepancy. Basically, the more a statement deviates from 0, the less informative it is.

Floridi stipulates five conditions, that "any feasible and satisfactory metric will have to satisfy" [65, p. 206]:

Condition 1. For a true s that conforms most precisely and accurately to the actual situation w , $f(s) = 0$.

⁶Whilst we are using statements (s), Floridi uses situation theoretic infons (σ). This distinction is of no concern for the purposes of this work, so talk of infons will be replaced by talk of statements here.

Condition 2. For an s that is made true by every situation (*i.e.*, a tautology), $f(s) = 1$.

Condition 3. For an s that is made true in no situation (*i.e.*, a contradiction), $f(s) = -1$.

Condition 4. For a contingently false s , $-1 < f(s) < 0$ ⁷.

Condition 5. For a contingently true s that is also made true by situations other than the actual one (so does not conform to w with the highest degree of precision), $0 < f(s) < 1$.

For cases that fall under the fourth condition, f measures *degree of inaccuracy* and is calculated as the negative ratio between the number of false atomic statements (e) in the given statement s and the length (l) of s .

$$\text{Inaccuracy} : f(s) = -\frac{e(s)}{l(s)}$$

To get values for e and l , it seems that some simplifying assumptions need to be made about a statement's form. Statements to be analysed in this way are in conjunctive normal form, with each conjunct being an atomic statement. e is the number of false conjuncts and l is the total number of conjuncts. This point will be discussed in more detail shortly.

For cases that fall under the fifth condition, f measures *degree of vacuity* and is calculated as the ratio between the number of situations, including the actual situation, with which s is consistent (n) and the total number of possible situations (S) or states.

$$\text{Vacuity} : f(s) = \frac{n(s)}{S}$$

Of the inaccuracy metric Floridi writes, “[it] allows us to partition $s = s^l$ into l disjoint classes of inaccurate s $\{\text{Inac}_1, \dots, \text{Inac}_l\}$ and map each class to its corresponding degree of inaccuracy” [65, p. 208].

The model Floridi uses to illustrate his account, denoted E , has two predicates (G and H) and three objects (a , b and c). So in all there is a set of 64 possible states $W = w_1, \dots, w_{64}$. An application of the inaccuracy metric to the model E is presented in Table 2.7 (degree of informativeness calculations will be explained below).

⁷This should technically be $-1 \leq f(s) < 0$

# errors	Class of inaccuracy	Cardinality of Inac_i	Degree of inaccuracy	Degree of informativeness
1	Inac_1	6	$-\frac{1}{6}$	≈ 0.972
2	Inac_2	15	$-\frac{1}{3}$	≈ 0.888
3	Inac_3	20	$-\frac{1}{2}$	0.75
4	Inac_4	14	$-\frac{2}{3}$	≈ 0.555
5	Inac_5	7	$-\frac{5}{6}$	≈ 0.305
6	Inac_6	1	-1	0

Table 2.7: Classes of inaccuracy.

For vacuous statements, Floridi writes “[the vacuity metric] and the previous method of *semantic weakening* allow us to partition s into $l - 1$ disjoint classes $\text{Vac} = \{\text{Vac}_1, \dots, \text{Vac}_{l-1}\}$, and map each class to its corresponding degree of vacuity” [65, p. 209]. The semantic weakening method referred to consists of generating a set of statements by the following process. In this case, the number of all atomic propositions to be used in a statement is 6. Start with a statement consisting of 5 disjunctions, such as

$$Ga \vee Gb \vee Gc \vee Ha \vee Hb \vee Hc$$

This is the weakest type of statement and corresponds to Vac_1 . Next, replace one of the disjunctions with a conjunction, to get something like

$$(Ga \vee Gb \vee Gc \vee Ha \vee Hb) \wedge Hc$$

This is the second weakest type of statement and corresponds to Vac_2 . Continue this generation process until only one disjunction remains. Table 2.8 summarises an application of this process to the model E :

Suppose that the actual situation corresponds to the state description $Ga \wedge Ha \wedge Gb \wedge Hb \wedge Gc \wedge Hc$. Table 2.9 summarises an example member of each class:

With a way to calculate degrees of vacuity and inaccuracy at hand Floridi then provides a straightforward way to calculate degrees of informativeness (g), by using the following formula, where once again f stands for the degree of vacuity/inaccuracy function:

$$g(s) = 1 - f(s)^2$$

# compatible situations	com- patible situations	Class vacuity	of Cardinality of Vac_i	Degree vacuity	of Degree of infor- mativeness
63		Vac_1	63	$\frac{63}{64}$	0.031
31		Vac_2	31	$\frac{31}{64}$	0.765
15		Vac_3	15	$\frac{15}{64}$	0.945
7		Vac_4	7	$\frac{7}{64}$	0.988
3		Vac_5	3	$\frac{3}{64}$	0.998

Table 2.8: Classes of vacuity.

Thus we have the following three conditions:

1. $(f(s) = 0) \Rightarrow (g(s) = 1)$
2. $(f(s) = +1 \vee f(s) = -1) \Rightarrow (g(s) = 0)$
3. $(0 < f(s) < +1) \vee (-1 < f(s) < 0) \Rightarrow (0 < g(s) < 1)$

Class	Statement
$Inac_1$	$Ga \wedge Ha \wedge Gb \wedge Hb \wedge Gc \wedge \neg Hc$
$Inac_2$	$Ga \wedge Ha \wedge Gb \wedge Hb \wedge \neg Gc \wedge \neg Hc$
$Inac_3$	$Ga \wedge Ha \wedge Gb \wedge \neg Hb \wedge \neg Gc \wedge \neg Hc$
$Inac_4$	$Ga \wedge Ha \wedge \neg Gb \wedge \neg Hb \wedge \neg Gc \wedge \neg Hc$
$Inac_5$	$Ga \wedge \neg Ha \wedge \neg Gb \wedge \neg Hb \wedge \neg Gc \wedge \neg Hc$
$Inac_6$	$\neg Ga \wedge \neg Ha \wedge \neg Gb \wedge \neg Hb \wedge \neg Gc \wedge \neg Hc$
Vac_1	$Ga \vee Gb \vee Gc \vee Ha \vee Hb \vee Hc$
Vac_2	$(Ga \vee Gb \vee Gc \vee Ha \vee Hb) \wedge Hc$
Vac_3	$(Ga \vee Gb \vee Gc \vee Ha) \wedge Hb \wedge Hc$
Vac_4	$(Ga \vee Gb \vee Gc) \wedge Ha \wedge Hb \wedge Hc$
Vac_5	$(Ga \vee Gb) \wedge Gc \wedge Ha \wedge Hb \wedge Hc$

Table 2.9: Classes of inaccuracy and vacuity examples.

Furthermore, this is extended and an extra way to measure amounts of semantic information is provided. As this extension is simply derivative and not essential, I will not go into it here. Suffice it to say, naturally, the higher the informativeness of s , the larger the quantity of semantic information it contains; the lower the informativeness of s , the smaller the quantity of semantic information it contains. To calculate this quantity of semantic information contained in s relative to $g(s)$, Floridi makes use of integrals and the area delimited by the equation given for degree of informativeness.

2.3.1 Some Comments on Floridi's Theory

It will be evident to the reader that the classes of inaccuracy and vacuity presented by Floridi are not comprehensive in that they do not accommodate the full range of statements which could be constructed in the logical space of E . Once again, suppose that the actual situation corresponds to $Ga \wedge Ha \wedge Gb \wedge Hb \wedge Gc \wedge Hc$.

Take the false statement $Ga \wedge Ha \wedge Gb \wedge \neg Hb \wedge \neg Gc$, consisting of 5 conjoined atoms, 2 of which are false. A simple extension of Floridi's method for dealing with inaccuracy would naturally result in several other classes and the degree of inaccuracy of this statement would be $-\frac{2}{5}$.

Or take the following false statement:

$$s = (Ga \vee Ha) \wedge Gb \wedge \neg Hb \wedge (\neg Gc \vee \neg Hc)$$

How should the formula given for inaccuracy be applied here? There is no clear-cut way to determine the values for e and l going by Floridi's description of the formula.

As for the possible classes of vacuity in E , it is clear that beyond those listed in above table, for any x such that $2 \leq x \leq 63$, it is possible to construct a statement such that the degree of vacuity is $\frac{x}{64}$.

In total the classes given by Floridi deal with 14 different types of propositions: 1 (true state description) + 1 (tautologies) + 1 (contradictions) + 6 (classes of inaccuracy) + 5 (classes of vacuity). Since there are 64 possible states in the space E , there are 2^{64} different propositions (propositions being satisfied by a set of states, so 2^{64} different sets of states). With an aim of developing a system that can deal with the complete range of propositions in a given logical space, we must look further.

Another aspect of Floridi's system which draws consideration is a certain asymmetry between the metrics for false statements and true statements. False statements are dealt with syntactically and true statements are dealt with by analysing their models. Might there be one approach that deals appropriately with both? A consequence of this separation between the false metric and true metric is that the same numerical value might be given to both a false statement and a true statement with otherwise seemingly different information measures. For example, take the following two statements:

1. $(p_1 \wedge p_2) \vee (p_2 \wedge p_3) \vee (p_3 \wedge p_4)$ (true)

2. $p_1 \wedge p_2 \wedge \neg p_3 \wedge \neg p_4$ (false)

Relative to an actual state corresponding to the state description $p_1 \wedge p_2 \wedge p_3 \wedge p_4$, both 1 and 2 are given informativeness measures of 0.75 using Floridi's measure. Yet as will be formally shown in the next section it is fair to say that the true statement here is more informative than the false one.

The central idea behind Floridi's theory is right. Both false and true statements can deviate from the precise truth. For false statements, the more falsity and less truth they contain, the greater the deviation hence the less information. For true statements, the less precise they are the greater their deviation, hence the less information. However as we have just seen it falls short of providing rigorous metrics, which can deliver appropriate and relatively consistent measures for the complete class of propositions in a logical space.

2.4 Information Quantification via Truthlikeness

As we have seen so far, the CSI approach to quantifying semantic information does not imply a veridicality condition for information. Floridi's TSSI on the other hand holds that semantic information should be meant as implying truth. This paves the way for an alternative approach to quantifying semantic information; rather than measuring information in terms of probability, the information given by a statement is to be measured in terms of its truthlikeness. Discussion in Section 2.3.1 suggested however that Floridi's contribution in [65] is perhaps of more conceptual rather than technical value, as the metrics provided can make way for some more detailed and refined improvements.

At any rate, there already is a mass of formal work on truthlikeness to draw from ⁸. Indeed the truthlikeness research enterprise has existed for at least a few decades now and there is a good variety of approaches to formalising this concept, a few of which we will now look at. Whilst seemingly simple in essence, the notion of truthlikeness has resisted a straightforward formal characterisation. Over the last few decades of research, a variety of rival approaches have developed, each with their own pros and cons. In turn it follows that information quantification via truthlikeness measures is also not going to be a straightforward matter with a definitive account.

⁸Truthlikeness is also referred to as verisimilitude, although some use the two terms distinctively, to refer to a basic distinction in the types of approaches (verisimilitude for *content* approaches and truthlikeness for *likeness* approaches). Its origins can be traced back to Popper, who motivated by his philosophy of science, was the first philosopher to take the formal problem of truthlikeness seriously. See [139] for a brief introduction to truthlikeness. [138] and [133] are major and now somewhat classic pieces of work within the truthlikeness enterprise. A great piece of relatively recent work with a good summary of all that has come before it can be found in [183]

2.4.1 The Basic Feature Approach to Truthlikeness

In [93] and [28] a simple approach to truthlikeness dealing with a restricted class of statements is introduced. According to this *basic feature* approach, the truthlikeness of a theory or statement A depends only on what A says about the ‘basic features’ of the logical space (“the world”). Using a propositional language L_n with n atoms, these basic features are described using *literals* (either an atomic statement p_i or its negation, $\neg p_i$). A literal is denoted $\pm p_i$.

Here are some key points:

- A *constituent* (i.e. *state description*) C_i describes a specific possible state and has the form $\pm p_1 \wedge \dots \wedge \pm p_n$. There is one true constituent C_* , corresponding to the one true or actual state.
- A *conjunctive statement* (c-statement) in L_n is a conjunction of the form: $\pm p_1 \wedge \dots \wedge \pm p_k$, where $0 \leq k \leq n$. So constituents are simply c-statements with $k = n$ and a tautology can be seen as a c-statement with $k = 0$. There are 3^n c-statements within language L_n (See Theorem 2.4.1 below).
- A literal $\pm p_i$ occurring as a conjunct of a c-statement A is a *basic claim* (b-claim) of A . The set of all the b-claims of A is referred to as the *basic content* (b-content) of A and is denoted A^+ .
- $t(A, C_i)$ is the subset of basic content consisting of the true b-claims of A relative to the state corresponding to constituent C_i and $f(A, C_i)$ is the subset consisting of the false b-claims of A relative to the state corresponding to constituent C_i .

So basically this approach just deals with propositional statements in conjunctive normal form, with each conjunct being either an atom or negated atom.

The quantitative notions of degree of true b-content $\text{cont}_t(A, C_i)$ and the degree of false b-content $\text{cont}_f(A, C_i)$ are defined as

$$\text{cont}_t(A, C_i) = \frac{|t(A, C_i)|}{n} \text{ and } \text{cont}_f(A, C_i) = \frac{|f(A, C_i)|}{n}$$

and the similarity of statement A to constituent C_i can be defined as a weighted average between these two:

$$s_\tau(A, C_i) = \tau \text{cont}_t(A, C_i) - (1 - \tau) \text{cont}_f(A, C_i)$$

where $0 < \tau \leq \frac{1}{2}$. Finally, the truthlikeness $[\text{Tr}_\tau]$ of A is defined as:

$$\text{Tr}_\tau(A) = s_\tau(A, C_*)$$

This function offers a very simple way to measure strongly semantic information in terms of truthlikeness. Two small points to consider if this function were employed for such a thing:

- An appropriate value range for τ would be: $0 < \tau \leq 1$
- If information yield is defined in terms of something like degrees of truth rather than truthlikeness, then one could remove the falsity component and just use $\text{cont}_t(A, C_*)$.

Following is an example of the basic feature measure:

Example 2.1

- $C_* = p_1 \wedge p_2 \wedge p_3 \wedge p_4 \wedge p_5$
- $A = p_1 \wedge p_2 \wedge p_3 \wedge \neg p_4$
- $\text{cont}_t(A, C_*) = \frac{3}{5}$
- $\text{cont}_f(A, C_*) = \frac{1}{5}$
- $\tau = \frac{1}{2}$
- $\text{Tr}_\tau(A) = \frac{1}{2} \times \frac{3}{5} - \frac{1}{2} \times \frac{1}{5} = 0.2$

□

Theorem 2.4.1. The set of c-statements expressible within L_n contains 3^n members.

Proof. Each c-statement can be seen as a normalised c-statement of the form $X_1 \wedge X_2 \wedge \dots \wedge X_n$, where each X_i is one of

- p_i

- $\neg p_i$
- $p_i \vee \neg p_i$

Since there are n conjuncts in each of these normalised c-statements and three possible values for each conjunct to take, there are 3^n different c-statements.

□

Within its range of applicability, the basic features approach is an elegant and appropriate way to quantify truthlikeness or strongly semantic information. We will touch upon it again in the next chapter and due to its nature and simplicity it will be used for demonstrative purposes further down the track. However, with an aim of finding a system that can deal with the complete range of propositions in a given logical space, we must once again look further.

2.4.2 The Tichy/Oddie Approach to Truthlikeness

The first main approach we will look at was first proposed by Pavel Tichy and expanded upon by Graham Oddie [138, p. 44]. This approach is an example of approaches that measure the truthlikeness of a statement A by firstly calculating its distance from a statement T using some distance metric Δ , where T is a state description of the actual state (so T is in a sense *the truth*). Actually, the distance metric is ultimately a function that operates on states. $\Delta(w_i, w_j)$ measures the distance between states w_i and w_j and we use the notation $\Delta(A, T)$ in effect as shorthand for an operation that reduces to Δ operation on states corresponding to A and T .

The result of this distance calculation is then used to calculate the statement's truthlikeness (Tr); the greater this distance, the less truthlike the statement and vice versa. This inverse relation is simply achieved with the following formula:

$$Tr(A, T) = 1 - \Delta(A, T)$$

To see this approach at work, consider again the canonical weather framework, with w_1 being the actual state (throughout the remainder of this work w_1 is assumed to be the actual state in all examples):

State	h	r	w
w ₁	T	T	T
w ₂	T	T	F
w ₃	T	F	T
w ₄	T	F	F
w ₅	F	T	T
w ₆	F	T	F
w ₇	F	F	T
w ₈	F	F	F

Table 2.10: Truth Table for 3-Proposition Weather Example Logical Space.

Before continuing with some examples, the introduction of some terminology is in order. We firstly recall that a *state description* is a conjunction of atomic statements consisting of each atom in the logical space or its negation, but never both and that each state description corresponds to a state. In our example $T = h \wedge r \wedge w$.

A statement is in *distributive normal form* if it consists of a disjunction of the state descriptions of states in which it is true. For example, $h \wedge r$ is true in states w₁ and w₂, so its distributive normal form is $(h \wedge r \wedge w) \vee (h \wedge r \wedge \neg w)$. For notational convenience throughout the remainder of this work, when listing a statement in its distributive normal form, state descriptions may be substituted by the states they correspond to. For example, $h \wedge r$ may be represented with the term w₁ $\vee$ w₂.

Now, take the possible states w₁, w₈ and w₅, which correspond to the state descriptions $h \wedge r \wedge w$, $\neg h \wedge \neg r \wedge \neg w$ and $\neg h \wedge r \wedge w$ respectively. The difference between w₁ and w₈ is {h, r, w}; they differ in every atomic assertion. w₅ and w₁ on the other hand differ by only {h}. So $\neg h \wedge r \wedge w$ is more truthlike than $\neg h \wedge \neg r \wedge \neg w$, because the distance between w₅ and the actual state w₁ is less than the distance between w₈ and w₁.

The general formula to calculate the distance between A and T is:

$$\Delta(A, T) = \frac{d}{|W_A|} \quad (2.1)$$

where

- Let w_T stand for the state that corresponds to T .
- W_A stands for the set of states in which A is true.
- A weight with value $\frac{1}{n}$ is assigned to every atomic element, where n is the number

of propositions in the logical space. So in our example, $\text{weight}(h) = \text{weight}(r) = \text{weight}(w) = \frac{1}{3}$. This weight is used to calculate the value of the distance between two states, with each atomic assertion difference adding $\frac{1}{n}$. So $\Delta(w_5, w_1) = \frac{1}{3}$ and $\Delta(w_8, w_1) = 1$.

- d is the sum of atomic assertion differences between A and T . That is, the sum of $\Delta(w_a, w_T)$ for each $w_a \in W_A$.

So the statement $\neg h \wedge r \wedge w$ (the state description for w_5) has a truthlikeness/information measurement of $\frac{2}{3}$ and the statement $\neg h \wedge \neg r \wedge \neg w$ (the state description for w_8) has a truthlikeness/information measurement of 0.

This approach extends to the complete range of statements involving h , r and w . According to Formula 2.1, the distance of a statement from the truth is defined as the average distance between each of the states in which the state is true and the actual state. Take the statement $h \wedge \neg r$, which makes assertions about only 2 of the 3 atomic states. It is true in both w_3 and w_4 , or $\{h, \neg r, w\}$ and $\{h, \neg r, \neg w\}$ respectively. w_3 has a distance of 0.33 and w_4 has a distance of 0.67 from w_1 so the average distance is $\frac{0.33+0.67}{2} = 0.5$. Note that henceforth the truthlikeness function term $[\text{Tr}()]$ will be replaced by an information measure function term $[\text{info}()]$, so that

$$\text{info}(A, T) = 1 - \Delta(A, T)$$

Also, given that T is set as w_1 throughout this work, $\text{info}(A, T)$ will generally be abbreviated to $\text{info}(A)$.

Table 2.11 lists some results using this method. How do these results fare as measures of informativeness? Statement #1 (complete truth) clearly is the most informative and #21 (complete falsity) clearly the least informative. #10, #11, #15 and #16 indicate that for a disjunctive statement, the more false constituents it has the less informative it is. In general, the more false constituents contained in a formula the more likely a decrease in informativeness, as indicated by the difference between #5 and #9.

Statements #7, #17 and #21 make sense; the greater the number of false conjuncts a statement has, the less informative it is, with the upper bound being a statement consisting of nothing but false conjuncts, which is accordingly assigned a measure of 0. Although #17 and #19 have the same measure, #17 has one true conjunct out of three atomic conjuncts and #19 has zero true atomic statements out of one atomic statement. From this it can be said that the assertion of a false statement detracts more from informativeness than the absence of

#	Statement (A)	T/F	info(A)
1	$h \wedge r \wedge w$	T	1
2	$h \wedge r$	T	0.83
3	$h \wedge (r \vee w)$	T	0.78
4	$h \wedge (\neg r \vee w)$	T	0.67
5	$(h \wedge r) \vee w$	T	0.67
6	h	T	0.67
7	$h \wedge r \wedge \neg w$	F	0.67
8	$h \vee r$	T	0.61
9	$(h \wedge \neg r) \vee w$	T	0.6
10	$h \vee r \vee w$	T	0.57
11	$h \vee r \vee \neg w$	T	0.52
12	$h \vee \neg r$	T	0.5
13	$h \vee \neg h$	T	0.5
14	$h \wedge \neg r$	F	0.5
15	$h \vee \neg r \vee \neg w$	T	0.48
16	$\neg h \vee \neg r \vee \neg w$	F	0.43
17	$h \wedge \neg r \wedge \neg w$	F	0.33
18	$(h \vee \neg w) \wedge \neg r$	F	0.33
19	$\neg h$	F	0.33
20	$\neg h \wedge \neg r$	F	0.17
21	$\neg h \wedge \neg r \wedge \neg w$	F	0
22	$h \wedge \neg h$	F	N/A

Table 2.11: Informativeness results using Tichy/Oddie metric.

an assertion or denial of that statement. Half of #14 is true, so it is appropriately assigned a measure of 0.5. #20 further shows that falsity proportionally detracts from informativeness. The difference between #18 and #16 is perhaps the most interesting. Although #18 has one true conjunct out of two and #16 contains no true atoms, #16 comes out as being more informative, further suggesting the price paid for asserting falsity in a conjunction. Also, note that some false statements are more informative than some true statements. An interpretation of the Tichy/Oddie method, which will provide a way to further understand some of these results, is given in Section 2.4.5.

One possible issue with this method is that tautologies are not assigned a maximum distance and hence are assigned a non-zero, positive measure. Without going into detail, it seems that this is a more significant issue for a quantitative account of information than it is for an account of truthlikeness. The implication that tautologies have a middle degree of informativeness strongly conflicts with the intuition and widely accepted position that tautologies are not informative. There are several ways to address this point. One way is something along the lines of an explanation of this result that will be discussed in Section 2.4.5. Another way, which will be discussed in the next chapter, is to base an informativeness measure of zero on the fact that the addition of tautological input to a database never changes

its content. Or the simplest way to address this point would be to just exclude tautologies from this metric and assign them a predefined distance of 1 hence informativeness measure of 0. This expedient move would put the tautology alongside its extreme counterpart the contradiction, which is also excluded. In the case of contradictions (#22), these calculations do not apply, because since they are true in no states this would mean division by zero.

For a logical space with n atoms, there are 2^n different propositions (sets of states). Table 2.12 lists an ordering of the different informativeness classes using the Tichy/Oddie measure for $n = 3$.

	info()	cardinality	# true statements	# false statements
1	1	1	1	0
2	0.83	3	3	0
3	0.78	3	3	0
4	0.75	1	1	0
5	0.67	31	24	7
6	0.61	3	3	0
7	0.6	10	10	0
8	0.583	15	12	3
9	0.57	1	1	0
10	0.56	21	12	9
11	0.53	15	12	3
12	0.476	3	3	0
13	0.5	41	21	20
14	0.524	3	3	0
15	0.47	15	9	6
16	0.44	21	6	15
17	0.429	1	0	1
18	0.417	15	3	12
19	0.4	10	1	9
20	0.39	3	0	3
21	0.33	31	0	31
22	0.25	1	0	1
23	0.22	3	0	3
24	0.17	3	0	3
25	0	1	0	1

Table 2.12: Sample list of informativeness classes using the Tichy/Oddie measure

Quantifying Misinformation

One task that Floridi leaves for a subsequent stage of research in his paper to which we shall now turn is “the extension of the quantitative analysis to the semantic concepts of quantity of misinformation” [65, p. 217]. However before undertaking this task a clarification is in

order. It has been established in this work that semantic information is defined as true semantic content and misinformation is defined as false semantic content. For example, in Table 2.11, statement #1 would be considered a piece of semantic information and statement #7 would be considered a piece of semantic misinformation. Given this, here are two issues to consider:

1. Statement #7, a piece of misinformation, has a non-zero, positive information measure;
2. Some pieces of misinformation (false statements) are more informative than some pieces of information (true statements).

In order to address these issues, I refer to some points made by Floridi in his objections to reasons to think that false information is a type of semantic information [66, p. 361]:

1.
 - **Reason 1.** Misinformation can include genuine information.
 - **Objection 1.** This just shows that misinformation is a compound in which only the true component qualifies as information.
2.
 - **Reason 2.** Misinformation can entail genuine information.
 - **Objection 2.** Even if one correctly infers only some semantically relevant and true information from misinformation, what now counts as information is the inferred true consequence, not the original misinformation. Besides, *ex falso quod libet sequitur*, so any contradiction would count as information.

Thus in this approach statements can have both a true, information component and a false, misinformation component.

Continuing on, how might the approach in this section to quantifying information be used in order to quantify misinformation? To begin with, it seems that one straightforward stipulation to make is that true statements should have a measure of 0, for they are not misinformative. So unlike information measures, where both true and false statements can have relevant measures, with misinformation attention is confined to false statements⁹. The more falsity in a false statement, the more misinformation it yields. As with the information measure just given, the first metric to devise is one which measures the degree of deviation from complete misinformation ($\Delta_{misinfo}$). A deviation of 0 will translate to maximum misinformation and a deviation of 1 will translate to no misinformation. So the metric for this deviation will at least satisfy these conditions:

⁹There is a sense in which true statements might contain misinformation, as will be discussed later

1. all true statements have a predefined deviation of 1;
2. all contingently false statements have a deviation greater than or equal to 0 and less than 1.

Two ways to go about measuring this deviation come to mind. The first $[\Delta_{misinfo1}]$ is to use the complement of the deviation for information measures: $\Delta_{misinfo}(A, T) = 1 - \Delta_{info}(A, T)$.

For the second $[\Delta_{misinfo2}]$, we shall bring in some terminology given by Oddie [138, p. 50], who discusses a reversal operation $\text{Rew}()$ on states such that:

$\text{Rew}(U) =$ the state V such that for any atomic state B , B is true in U if and only if B is false in V .

This reversal operation on states is extended to a reversal operation on propositions. The reversal of a proposition A is the image of A under $\text{Rew}()$:

$\text{Rev}(A) =$ the proposition B such that A contains state U if and only if B contains $\text{Rew}(U)$, for any state U .

Where w_T is the actual state, the second way to measure misinformation deviation would be to measure the distance of a statement from $\text{Rew}(w_T)$. In our example, $w_T = w_1$ so $\text{Rew}(w_T) = w_8$. These two approaches turn out in fact to be equivalent.

Theorem 2.4.2. $\Delta_{misinfo1} = \Delta_{misinfo2}$

Proof. See Appendix A Theorem A.0.1.

□

From here one can go on to calculate a statement's misinformation measure simply by subtracting its deviation/distance from 1; the greater the distance, the less misinformative the statement:

$$\text{misinfo}(A) = 1 - \Delta_{misinfo}(A, T)$$

From this it follows that

$$\text{info}(A) + \text{misinfo}(A) = 1$$

and the following hold:

- $\text{info}(\text{Rev}(T)) = 0$
- $\text{info}(A) + \text{info}(\text{Rev}(A)) = 1$
- $\text{misinfo}(A) = \text{info}(\text{Rev}(A))$
- $\text{info}(A) > \text{info}(B) \Leftrightarrow \text{misinfo}(A) < \text{misinfo}(B)$

As this section has shown, once truth becomes a factor in semantic information quantification, this naturally leads to a method for quantifying misinformation.

Adjusting Atom Weights

It is worth briefly mentioning the possibility of adjusting atomic weights in order to reflect differences in the informational value of atomic statements. As we have seen, where n stands for the number of propositional atoms in a logical space, each atom is assigned a standard weight of $\frac{1}{n}$ for the purposes of Δ calculation. In the 3-proposition weather example being discussed, this results in each atom being assigned a weight of $\frac{1}{3}$. As a consequence of this even distribution of weight, the three statements $h \wedge r \wedge \neg w$, $h \wedge \neg r \wedge w$ and $\neg h \wedge r \wedge w$ are all assigned the same information measure.

Beyond this there is the possibility of adjusting the weights so that the resulting assignment is non-uniform. Such a modification could perhaps be used to model cases where different atomic statements (hence respective compound statements containing them) have different informational value. There is much room to interpret just what is meant here by ‘informational value’. Statement A could have a greater informational value than B if an agent prefers the acquisition of A over the acquisition of B , or if the agent can do more with A than B . Variable weights could also perhaps be used to reflect extra-quantitative or qualitative factors. These are some preliminary thoughts.

The minimum requirement is that the sum of the values assigned to the atoms comes to 1. This is to ensure that values for Δ are appropriately distributed between 0 and 1, with 1 being the maximum, reserved for $\text{Rev}(w_T)$.

With this in mind, take the following weight assignments:

- $\text{weight}(h) = \frac{1}{6}$
- $\text{weight}(r) = \frac{2}{6}$
- $\text{weight}(w) = \frac{3}{6}$

With such an assignment, the information measurement distribution changes significantly. Some results are given in Table 2.13.

#	Statement (A)	$\text{info}(A)$
1	$h \wedge r \wedge \neg w$	0.5
2	$h \wedge \neg r \wedge w$	0.67
3	$\neg h \wedge r \wedge w$	0.83
4	$h \wedge \neg r \wedge \neg w$	0.167
5	$\neg h \wedge \neg r \wedge w$	0.5
6	$\neg h \wedge r \wedge \neg w$	0.33

Table 2.13: Results with adjusted weights.

It can be seen that although statements 1, 2 and 3 all share the same form (2 true atoms and 1 false atom), none share the same information measure. In such a case, 3 yields more information than 1 due to the fact that its two true atoms are of greater informational value than the two true atoms contained in 1.

One potential issue with such a modification is the treatment of misinformation quantification. Although 1 has a lower information yield measure than 3, it does not seem as right to say that conversely it has a greater misinformation yield, since it contains the same amount of falsity as 3. Is there a corresponding variation in misinformational value?

Contradictions

As we have just seen, it is not mathematically possible to calculate an $\text{info}()$ value for contradictions using this method (since contradictions contain no states, this means a division by 0). How then should contradictions be dealt with? One option is to simply exclude contradictions from the metrics and assign them a predefined deviation of 1 hence information measure of 0. As we have seen, this is the approach that Floridi takes.

This however is arguably too rash a move and there are good reasons to adopt an approach in which the metrics are adjusted or expanded in order to accommodate contradictions. To

begin with, the class of contradictions is not homogeneous with regards to informativeness and different contradictions can be treated as having different non-zero measures of informativeness.

Take a logical space consisting of 100 atomic propositions, $p_1, p_2 \dots p_{100}$, all true relative to the actual state. Now the statement $p_1 \wedge p_2 \wedge p_3 \wedge \dots \wedge p_{99} \wedge p_{100}$ is maximally informative. If we conjoin it with $\neg p_{100}$ to get the contradiction $p_1 \wedge p_2 \wedge p_3 \wedge \dots \wedge p_{99} \wedge p_{100} \wedge \neg p_{100}$, the original information remains and the resulting statement should not instantly be assigned an information measure of 0. To put it another way, if one contradictory atom is inserted into a database with much information, whilst this means that there is now some misinformation within the database, surely there is still a lot of information within the database. If the contents of the database were to be represented as a statement (the statement being a contradiction), it would be good to have a way to measure information that can deal with the fact that the database still contains much information and that is also sensitive to the fact that different contradictions have different informativeness measures; for example, $p_1 \wedge p_2 \wedge p_3 \wedge \dots \wedge p_{99} \wedge p_{100} \wedge \neg p_{100}$ clearly is more informative than $p_1 \wedge \neg p_1$.

Now that the case has been made, in order to accommodate contradictions and have a way to assign them positive, non-zero measures of information, I propose the following approach. All non-contradictions are dealt with as before, using the standard metrics with classical propositional logic. For contradictions, a many-valued framework such as that associated with the paraconsistent logic *LP* is employed¹⁰. In this system a third truth value, B (true and false) is introduced. Importantly, since contradictions can hold in states involving B, the denominator in calculations involving contradictions need no longer be and is in fact never 0.

In the classical framework, each atomic element is assigned a weight given by $\frac{1}{n}$ and the value of the distance between a T and an F is equal to this weight. For this extended framework, the classical distances are still the same and the distance between B and a T or an F is $\frac{1}{2n}$. So the distance between a B and a T or an F is half of the distance between a T and an F. Since B represents both true and false, this is a natural choice; B ‘half agrees’ with T/F and ‘half disagrees’. Its correctness is evident in the calculation results. Formally, the distance calculation between *A* and the true state description *T* now becomes:

$$\Delta(A, T) = \frac{d}{|W_A|} \quad (2.2)$$

¹⁰Refer back to Section 2.1.2 for an overview of this logic. The adoption of such a framework is for instrumental purposes only and is not an endorsement of paraconsistency or dialetheism

where

- Let w_T stand for the state that corresponds to T . This state will consist of only T and F truth values.
- W_A stands for the set of states in which A is true.
- where n is the number of propositions in the logical space:
 - A weight with value $\frac{1}{n}$ is assigned to every classical atomic state (T or F).
 - A weight with value $\frac{1}{2n}$ is assigned to every non-classical atomic state (B).
 - The distance function $\Delta(w_i, w_j)$ calculates the total difference between two states.
- d is the sum of atomic assertion differences between A and T ; that is, the sum of $\Delta(w_a, w_T)$, for each $w_a \in W_A$.

To see this approach at work we again consider the weather framework. The list of 27 possible states looks like this:

State	h	r	w	State	h	r	w	State	h	r	w
w ₁	T	T	T	w ₁₀	B	T	T	w ₁₉	F	T	T
w ₂	T	T	B	w ₁₁	B	T	B	w ₂₀	F	T	B
w ₃	T	T	F	w ₁₂	B	T	F	w ₂₁	F	T	F
w ₄	T	B	T	w ₁₃	B	B	T	w ₂₂	F	B	T
w ₅	T	B	B	w ₁₄	B	B	B	w ₂₃	F	B	B
w ₆	T	B	F	w ₁₅	B	B	F	w ₂₄	F	B	F
w ₇	T	F	T	w ₁₆	B	F	T	w ₂₅	F	F	T
w ₈	T	F	B	w ₁₇	B	F	B	w ₂₆	F	F	B
w ₉	T	F	F	w ₁₈	B	F	F	w ₂₇	F	F	F

Table 2.14: LP Truth Table for 3-Proposition Logical Space.

Take the statement $h \wedge \neg h$. It holds in $w_{10} - w_{18}$. In this set of states there are 6 instances of F and 15 instances of B. Therefore, the distance is:

$$\Delta(h \wedge \neg h, h \wedge r \wedge w) = \frac{(6 \times \frac{1}{3}) + (15 \times \frac{1}{6})}{9} = 0.5$$

Table 2.15 lists some results, with all statements being classical contradictions. These results are reasonable. #1 is the contradiction that gives the most information and can simply be seen as a conjunction of the maximum information yielding true state description $h \wedge r \wedge w$ and the false contradicting atom $\neg h$. Conversely, #15 is the contradiction that yields the least information and can simply be seen as a conjunction of the minimum information

yielding completely false state description $\neg h \wedge \neg r \wedge \neg w$ and the true contradicting atom h . In between these two extremes, the sensitivity of this approach to various forms of contradictions is clear. The significant number of contradictions with a measure of 0.5 emphasises how the Tichy/Oddie measure can be seen as gauging the balance between the truth and falsity contained in a statement (a detailed interpretation of the Tichy/Oddie approach, related to this balance gauging, is given in Section 2.4.5). For the basic contradiction #6, half of its atomic assertions are true and half are false; similarly for #9 and #10. Comparing #7 to #9 and #8 to #10, it can be seen that whether a collection of basic contradictions are disjunctively or conjunctively connected makes no difference and the measure remains 0.5; whether $h \wedge \neg h$ becomes $(h \wedge \neg h) \wedge (r \wedge \neg r)$ or $(h \wedge \neg h) \vee (r \wedge \neg r)$, the balance between truth and falsity remains a constant. However, as indicated by #11 and #13, there can be a difference between the conjunction of a basic contradiction with a non-basic contradiction and the corresponding disjunction.

#	Statement (A)	info(A)
1	$h \wedge \neg h \wedge r \wedge w$	0.6667
2	$h \wedge r \wedge \neg(h \wedge r)$	0.6111
3	$h \wedge \neg h \wedge r \wedge \neg r \wedge w$	0.5833
4	$((h \wedge \neg h) \vee (r \wedge \neg r)) \wedge w$	0.5833
5	$(h \wedge \neg h) \vee (r \wedge \neg r \wedge w)$	0.5256
6	$h \wedge \neg h$	0.5
7	$(h \wedge \neg h) \vee (r \wedge \neg r)$	0.5
8	$(h \wedge \neg h) \vee (r \wedge \neg r) \vee (w \wedge \neg w)$	0.5
9	$(h \wedge \neg h) \wedge (r \wedge \neg r)$	0.5
10	$(h \wedge \neg h) \wedge (r \wedge \neg r) \wedge (w \wedge \neg w)$	0.5
11	$(h \wedge \neg h \wedge \neg r) \vee (w \wedge \neg w)$	0.4744
12	$h \wedge \neg h \wedge \neg r$	0.4167
13	$(h \wedge \neg h \wedge \neg r) \wedge (w \wedge \neg w)$	0.4167
14	$\neg h \wedge \neg r \wedge \neg(\neg h \wedge \neg r)$	0.3889
15	$h \wedge \neg h \wedge \neg r \wedge \neg w$	0.3333

Table 2.15: Results for contradictions.

Whilst this approach gets things right amongst the set of contradictions, there is a measure discrepancy between contradictions and non-contradictions that could be rectified. To illustrate this issue, compare the following two statements:

1. $r \wedge w$
2. $h \wedge \neg h \wedge r \wedge w$

With regards to the balance between truth and falsity, both of these statements are equal. Yet the measure for (1) using the standard procedure for non-contradictions is 0.83 whilst

the measure for (2) using the contradiction procedure is 0.67.

Apart from any potential accommodation of this discrepancy that justifies a lower measure for (2) because contradictions are ‘punished’ more, if the original idea behind the Tichy/Oddie method of average truth/falsity balance is to be adhered to, then an adjustment is required.

One way to address this discrepancy is by treating non-contradictions the same as contradictions and dealing with them using the LP-based measurement approach. However,

the range of values would then change, with a maximum measure of $1 - \frac{(\frac{1}{2^n}) \sum_{x=0}^n x \times \binom{n}{x}}{2^n}$ assigned to the true state description (in our example case $\text{info}(h \wedge r \wedge w) = 0.75$) and

a minimum measure of $1 - \frac{(\frac{1}{2^n}) \sum_{x=0}^n x \times \binom{n}{x} + (\frac{1}{n}) \sum_{x=0}^n x \times \binom{n}{x}}{2^n}$ assigned to the false state description (in our example case $\text{info}(\neg h \wedge \neg r \wedge \neg w) = 0.25$).

This discrepancy could also be rectified with the following small change. For any contradiction A , instead of calculating against all of the models that satisfy A , a special subset is selected, namely the set of satisfying models that are *minimally inconsistent*:

Definition An LP-model M is *minimally inconsistent* for a statement A iff it is amongst the models for A with the lowest number of B atomic element valuations.

Example 2.2

$h \wedge \neg h \wedge r \wedge w$ is satisfied in the states $w_{10}, w_{11}, w_{13}, w_{14}$. Out of these, only w_{10} is minimally inconsistent, with one B valuation. Therefore the measure is now:

$$1 - \Delta(h \wedge \neg h \wedge r \wedge w) = 1 - \frac{(\frac{1}{6})}{1} = 0.83$$

□

Thus $r \wedge w$ and $h \wedge \neg h \wedge r \wedge w$ now have the same value. Furthermore, since the set of minimally inconsistent models for a classically satisfiable statement will always consist of only classical states, this method reduces to the standard classical Tichy/Oddie method when it comes to non-contradictions. Table 2.16 lists the new results using this method. As can be seen, the measure ordering has not changed from Table 2.15.

#	Statement (A)	info(A)
1	$h \wedge \neg h \wedge r \wedge w$	0.8333
2	$h \wedge r \wedge \neg(h \wedge r)$	0.6667
3	$h \wedge \neg h \wedge r \wedge \neg r \wedge w$	0.6667
4	$((h \wedge \neg h) \vee (r \wedge \neg r)) \wedge w$	0.6667
5	$(h \wedge \neg h) \vee (r \wedge \neg r \wedge w)$	0.5556
6	$h \wedge \neg h$	0.5
7	$(h \wedge \neg h) \vee (r \wedge \neg r)$	0.5
8	$(h \wedge \neg h) \vee (r \wedge \neg r) \vee (w \wedge \neg w)$	0.5
9	$(h \wedge \neg h) \wedge (r \wedge \neg r)$	0.5
10	$(h \wedge \neg h) \wedge (r \wedge \neg r) \wedge (w \wedge \neg w)$	0.5
11	$(h \wedge \neg h \wedge \neg r) \vee (w \wedge \neg w)$	0.4444
12	$h \wedge \neg h \wedge \neg r$	0.3333
13	$(h \wedge \neg h \wedge \neg r) \wedge (w \wedge \neg w)$	0.3333
14	$\neg h \wedge \neg r \wedge \neg(\neg h \wedge \neg r)$	0.3333
15	$h \wedge \neg h \wedge \neg r \wedge \neg w$	0.1667

Table 2.16: Results for contradictions.

2.4.3 Truthlikeness Adequacy Conditions and Information Conditions

It was seen with the CSI account that if $A \vdash B$ then $\text{cont}(A) \geq \text{cont}(B)$ and $\text{inf}(A) \geq \text{inf}(B)$. This is not going to hold in general for a Tichy/Oddie truthlikeness approach to information quantification.

For example, looking back at Table 2.11, although $\neg h \wedge \neg r \wedge \neg w \vdash \neg h$, $\text{info}(\neg h \wedge \neg r \wedge \neg w) = 0 < \text{info}(\neg h) = 0.334$. However since both of these statements are false, this example is not an issue; given two false statements such as these, the logically stronger one is understandably further away from the truth.

The question of interest is whether or not this property holds when A and B are both instances of information, when they are both true. So the pertinent question is this: among true statements, does information as truthlikeness here covary with logical strength? Put formally, does the following condition hold?:

If A and B are true statements and $A \vdash B$, then $\text{info}(B) \leq \text{info}(A)$

As it turns out, this condition does not hold. To begin with, it can be seen that although $h \vee \neg r \vee \neg w \vdash h \vee \neg h$, $\text{info}(h \vee \neg r \vee \neg w) = 0.48$ whereas $\text{info}(h \vee \neg h) = 0.5$. But leaving aside cases where tautologies are involved, this result also does not hold more generally, in cases where only contingently true statements are involved. For example, although $(h \wedge \neg r) \vee w \vdash h \vee w$, $\text{info}((h \wedge \neg r) \vee w) = 0.6 < \text{info}(h \vee w) = 0.61$. This is not an isolated case either.

Remark In a logical space containing three propositional variables, there are 2187 possible ways to have two true statements A and B such that $A \vdash B$. Out of these, 366 are such that $\text{info}(A) < \text{info}(B)$.

An interesting example arises in a space with four atoms. Take the 4-proposition logical space obtained by adding the propositional variable d to the 3-atom weather framework, with the actual state corresponding to the state description $h \wedge r \wedge w \wedge d$. Whilst it is the case that $\neg h \vee \neg r \vee w \vdash \neg h \vee \neg r \vee w \vee \neg d$, $\text{info}(\neg h \vee \neg r \vee w) = 0.482$ whilst $\text{info}(\neg h \vee \neg r \vee w \vee \neg d) = 0.483$.

This result seems quite counter-intuitive; the addition of a false disjunct to an already true statement slightly increases its information measure. Contrary to this result, it is fair to say that information is proportional to accuracy, and that the addition of disjunctions, particularly when false, decreases accuracy. Also, this result seems contrary to the notion of information loss. If the information $\neg h \vee \neg r \vee w$ stored in a database were corrupted and the salvaged remains were the weakened $\neg h \vee \neg r \vee w \vee \neg d$, this would ordinarily be described as a case of *information loss*. Or if a signal is transmitting the message $\neg h \vee \neg r \vee w$ and it became $\neg h \vee \neg r \vee w \vee \neg d$ due to noise, once again it can be said that there is information loss.

The failure of Tichy/Oddie truthlikeness, hence for our purposes information, to co-vary with logical strength amongst true statements is discussed by Ilkka Niiniluoto, who whilst developing his own account of truthlikeness in [133] surveys how various approaches to truthlikeness fare against a range of adequacy conditions¹¹. Though an interpretation of the Tichy/Oddie method that can serve to explain its failure to satisfy this condition and justify its use will be discussed shortly, in the meantime we will take a brief look at Niiniluoto's preferred approach to truthlikeness, which satisfies all of the adequacy conditions he lists.

2.4.4 Niiniluoto on Truthlikeness

As with the Tichy/Oddie approach to truthlikeness, the task for Niiniluoto is to define some distance function Δ such that $\Delta(A, T) \in [0, 1]$. He looks at six distance functions, which are listed below. Firstly, a recap and establishment of terms to be used:

- $\Delta(w_i, w_j)$ calculates the distance between states w_i and w_j . This is the sum of atomic differences multiplied by the atomic weight ($\frac{1}{n}$).

¹¹These conditions will be looked at in Section 2.5.1

- w_T is the actual state, that corresponds to the true state description T .
- W_A is the set of states in which A is true.
- $\mathbf{B}$ is the set of all states in the logical space.

Here are the distance functions:

- $\Delta_{\min}(A, T) =$ the minimum of the distances $\Delta(w_a, w_T)$ with $w_a \in W_A$.
- $\Delta_{\max}(A, T) =$ the maximum of the distances $\Delta(w_a, w_T)$ with $w_a \in W_A$.¹²
- $\Delta_{\text{sum}}(A, T) =$ the sum of all distances $\Delta(w_a, w_T)$ with $w_a \in W_A$, divided by the sum of all distances $\Delta(w_b, w_T)$ with $w_b \in \mathbf{B}$.
- $\Delta_{\text{av}}(A, T) =$ the sum of the distances $\Delta(w_a, w_T)$ with $w_a \in W_A$ divided by $|W_A|$.
- $\Delta_{\text{mm}}^\gamma(A, T) = \gamma\Delta_{\min}(A, T) + (1 - \gamma)\Delta_{\max}(A, T)$ for some weight γ with $0 \leq \gamma \leq 1$.
- $\Delta_{\text{ms}}^{\gamma\lambda}(A, T) = \gamma\Delta_{\min}(A, T) + \lambda\Delta_{\text{sum}}(A, T)$ for some two weights γ and λ with $0 \leq \gamma \leq 1$ and $0 \leq \lambda \leq 1$.

Δ_{av} is the Tichy/Oddie approach. Niiniluoto's preferred metric for truthlikeness is $\Delta_{\text{ms}}^{\gamma\lambda}$, which he terms the weighted min-sum measure [133, p. 228]. Once again, this distance calculation is then used to calculate truthlikeness: $\text{Tr}(A, T) = 1 - \Delta_{\text{ms}}^{\gamma\lambda}(A, T)$. Table 2.17 lists some results using the min-sum measure, with γ being assigned the value 0.89 and λ being assigned the value 0.44.

2.4.5 An Interpretation of the Tichy/Oddie Measure

Returning to the Tichy/Oddie approach, it is time to investigate further some of its problematic aspects that were touched upon earlier and see what we might make of them in relation to the task of quantifying information. But before doing so, an initial general point to make is that when it comes to quantifying information/misinformation, there is a certain asymmetry between true and false statements. Whilst false statements (misinformation) can be judged to yield some truth (information), true statements are ordinarily judged to just contain information and do not give any misinformation. But like false statements can yield some information, perhaps there is a sense in which true statements can lead to some

¹² $1 - \Delta_{\max}(A, T)$ is arguably an adequate definition of the degree of partial truth of A , a measure of the amount of *information about the truth* it conveys. Cevolani [27] favours this function as a measure of strongly semantic information.

#	Statement (A)	T/F	info(A)
1	$h \wedge r \wedge w$	T	1
2	$h \wedge r$	T	0.96
3	$h \wedge (r \vee w)$	T	0.93
4	$h \wedge (\neg r \vee w)$	T	0.89
5	h	T	0.85
6	$(h \wedge r) \vee w$	T	0.81
7	$(h \wedge \neg r) \vee w$	T	0.78
8	$h \vee r$	T	0.74
9	$h \wedge r \wedge \neg w$	F	0.67
10	$h \vee r \vee w$	T	0.67
11	$h \vee \neg r$	T	0.67
12	$h \vee r \vee \neg w$	T	0.63
13	$h \vee \neg r \vee \neg w$	T	0.6
14	$h \wedge \neg r$	F	0.59
15	$h \vee \neg h$	T	0.56
16	$(h \vee \neg w) \wedge \neg r$	F	0.48
17	$\neg h$	F	0.41
18	$h \wedge \neg r \wedge \neg w$	F	0.33
19	$\neg h \vee \neg r \vee \neg w$	F	0.26
20	$\neg h \wedge \neg r$	F	0.22
21	$\neg h \wedge \neg r \wedge \neg w$	F	0
22	$h \wedge \neg h$	F	N/A

Table 2.17: Results using Niiniluto's min-sum measure.

misinformation. Given an actual state corresponding to the state description $h \wedge r \wedge w$, it is straightforward to say that the false statement $h \wedge \neg r \wedge \neg w$, whilst misinformative on the whole, still yields some information. On the other hand, the statement $h \vee \neg w \vee \neg r$, whilst true, does not give one the complete truth about the domain of inquiry and the majority of its disjuncts are false. As will be shown now, such a statement can be seen as being potentially misinformative in the sense that it can potentially lead to falsity and it is this type of view that can be associated with and justify usage of the Tichy/Oddie approach for certain applications.

In Section 2.4.3 we saw that whilst it is the case that $(h \wedge \neg r) \vee w \vdash h \vee w$, $\text{info}((h \wedge \neg r) \vee w) < \text{info}(h \vee w)$. In light of the interpretation to be now provided, this result finds some support. $h \vee w$ contains no false atoms, whereas $(h \wedge \neg r) \vee w$ contains one, namely $\neg r$. Expanding upon this example to emphasise the point, take a logical space consisting of five atoms ($\{p_n | 1 \leq n \leq 5\}$), such that each is true in the actual state. Whilst it is the case that $(p_1 \wedge \neg p_2 \wedge \neg p_3 \wedge \neg p_4) \vee p_5 \vdash p_1 \vee p_5$, the antecedent statement makes many false assertions whereas the consequent makes none. In this way the Tichy/Oddie approach can be seen to measure not only how much atomic truth a statement contains, but also how much atomic falsity it contains. Also, in light of Section 2.4.2, these types of results can be seen as part

of adopting a uniform method for the measurement of information and its complementary misinformation.

Something along the lines of the idea just outlined can be rigorously captured by viewing the Tichy/Oddie approach in terms of *expected utility*. We begin by assigning each state (or state description) a utility value, where the value for a state w ($\text{val}(w)$) is determined using the following method:

- Let n stand for the number of propositional variables in the logical space.
- Let t stand for the number of true atomic elements, relative to the actual state, in the state w .
- $\text{val}(w) = \frac{t}{n}$.

So in the case of our 3-atom logical space with w_1 being the actual state, each state is valued as follows:

- $\text{val}(w_1) = 1$
- $\text{val}(w_2) = \text{val}(w_3) = \text{val}(w_5) = \frac{2}{3}$
- $\text{val}(w_4) = \text{val}(w_6) = \text{val}(w_7) = \frac{1}{3}$
- $\text{val}(w_8) = 0$

Now, given a statement A that holds in n states, convert it to distributive normal form, which will have n state description disjuncts. Imagine an agent is to choose one and only one of the disjuncts. The value of the selected disjunct's corresponding state determines the utility or informational value of the choice. Using the standard decision theoretic framework, we can say that the estimated utility of A [$\text{eu}(A)$] is:

$$\text{eu}(A) = \sum \text{val}(w_i) \times \text{Pr}(w_i)$$

The sum is over each state w_i that corresponds to a state description disjunct, $\text{val}(w_i)$ is the value of a state and $\text{Pr}(w_i)$ is the probability of it being chosen. Each disjunct has the same probability of being selected as any other, so for each state the probability of it being chosen is $\frac{1}{n}$. This estimated utility value equates to the statement's measure using the Tichy/Oddie approach.

Going back to the sound argument $(h \wedge \neg r) \vee w \vdash h \vee w$, here are both antecedent and consequent in distributive normal form followed by a decision-theoretic-style tabulation of the two statements:

- $(h \wedge \neg r) \vee w \equiv w_1 \vee w_3 \vee w_4 \vee w_5 \vee w_7$
- $h \vee w \equiv w_1 \vee w_2 \vee w_3 \vee w_4 \vee w_5 \vee w_7$

	w ₁	w ₃	w ₄	w ₅	w ₇
$(h \wedge \neg r) \vee w$	1	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$

$$\text{eu}((h \wedge \neg r) \vee w) = (1 \times \frac{1}{5}) + (\frac{2}{3} \times \frac{1}{5}) + (\frac{1}{3} \times \frac{1}{5}) + (\frac{2}{3} \times \frac{1}{5}) + (\frac{1}{3} \times \frac{1}{5}) = 0.6$$

	w ₁	w ₂	w ₃	w ₄	w ₅	w ₇
$h \vee w$	1	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$

$$\text{eu}(h \vee w) = (1 \times \frac{1}{6}) + (\frac{2}{3} \times \frac{1}{6}) + (\frac{2}{3} \times \frac{1}{6}) + (\frac{1}{3} \times \frac{1}{6}) + (\frac{2}{3} \times \frac{1}{6}) + (\frac{1}{3} \times \frac{1}{6}) = 0.61$$

Thus the information measurement of a statement here can be seen in terms of its expected utility, whereby both the positive (truth, information) and the negative (falsity, misinformation) are factored into calculations. Analysing the Tichy/Oddie measure in this way explains its counter-intuitive results. The tautology can be seen as suspension of judgement and sits at 0.5, favouring neither information nor misinformation.

Also, take the example in Section 2.4.3, where $\text{info}(\neg h \vee \neg r \vee w) < \text{info}(\neg h \vee \neg r \vee w \vee \neg d)$. Whilst the addition of a false disjunct to a statement is grounds to think that its information measure should decrease, seen in terms of expected utility, the addition of one false disjunct results in the addition of one more model to the set of satisfying states, namely the state corresponding to the state description $h \wedge r \wedge \neg w \wedge \neg d$. As can be seen, this addition actually results in a greater overall expected utility.

Apparently Tichy himself offered a similar analysis as an argument in defence of his truthlikeness measure [133, p. 238]. Niiniluoto is dismissive of the argument, but whilst his reasons might be legitimate, they pertain specifically to considerations of truthlikeness. At this point it is important to stress that although we are using the notion of truthlikeness to quantify information, this is not to say that they amount to one and the same thing and a consideration in accounting for one need not be a consideration in accounting for the other. Furthermore, the concept of information is generally treated more pluralistically than

truthlikeness. In fact, some of Niiniluoto's claims could be seen as offering support for the use of the Tichy/Oddie approach in a framework for information measurement. For example, he commences his case with the following:

even if it were the case that function M_{av} [the Tichy/Oddie measure] serves to measure the degree of *trustworthiness* (or pragmatic preference for action) of a hypothesis, it would not follow that M_{av} is an adequate measure of *truthlikeness*. These two concepts are clearly distinct. Tichy tends to think that a disjunction of two constituents $h_1 \vee h_2$ is something like a lottery ticket which gives us the alternatives h_1 and h_2 with equal probabilities: if we 'put our trust' to $h_1 \vee h_2$, we have to make a 'blind choice' between h_1 and h_2 . This idea is irrelevant, if we are dealing with answers to cognitive problems - the connection between truthlikeness and practical action is quite another question which has to be studied separately. [133, p. 238]

In closing and without elaboration, it suffices to mention that terms employed here, such as 'trustworthiness' and 'practical action' could be appropriately associated with the notion of information.

Contradictions under this Interpretation

The contradiction-accommodating extension to the Tichy/Oddie approach outlined in Section 2.4.2 can be easily dealt with using this interpretation. A straightforward way is to modify the standard classical way, where the value of a state w becomes:

$$\text{val}(w) = \frac{b}{2n} + \frac{t}{n}$$

where b stands for the number of B-valued atomic elements in the state and t stands for the number of T-valued atomic elements in the state.

Example 2.3

Take the statement $h \wedge \neg h \wedge r \wedge \neg r \wedge w$, which as listed in Table 2.15 as #3 has an information measure of 0.583. This statement holds in states w_{13} and w_{14} of Table 2.14. Now

- $\text{val}(w_{13}) = \frac{2}{3}$

- $\text{val}(w_{14}) = \frac{1}{2}$

so

$$\text{eu}(h \wedge \neg h \wedge r \wedge \neg r \wedge w) = \left(\frac{2}{3} \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{1}{2}\right) = 0.583$$

Using the minimally inconsistent method it simply becomes: $\text{eu}(h \wedge \neg h \wedge r \wedge \neg r \wedge w) = \frac{2}{3}$

□

As mentioned earlier, the adoption of a paraconsistent framework here is for instrumental purposes, with no commitment to the actual obtainability of paraconsistent states. One small issue with the interpretation given in the previous example is that it somewhat entertains the possibility of paraconsistent states. The decision theoretic framework outlined here involves an agent making a choice between possible states, with the possibility that one of the states is the actual state. For cases of contradiction measurement, the states being dealt with are paraconsistent ones. However since the preference here is to remain neutral and refrain from endorsing the possibility or actuality of paraconsistent states, a modified interpretation in terms of classical states is called for.

In the previous example, the contradiction $h \wedge \neg h \wedge r \wedge \neg r \wedge w$ corresponded to a choice between two paraconsistent states, w_{13} and w_{14} . How can this decision-theoretic analysis be made in terms of classical states only? To begin with, take w_{13} , which has the following valuations of atomic elements: $v(h) = B$, $v(r) = B$ and $v(w) = T$. We can say that one of these valuations, namely $v(w)$ is definite, since it results in a classical value. The others though are not definite and result in the paraconsistent valuation of B. In terms of the decision-theoretic analysis, a valuation of B means that the atomic element actually has a value of T or F, but there is no indication as to which one in particular. So it is to be treated as a ‘wildcard valuation’, to be substituted by either T or F. Using this system, substitutions for B in w_{13} result in the following set of classical states of Table 2.10 in Section 2.4: $\{w_1, w_3, w_5 \text{ and } w_7\}$. Let $\text{classical}()$ stand for a function which receives a paraconsistent state and outputs its corresponding classical states in distributive normal form. Since none of the three valuations for w_{14} are definite, substitutions for B in w_{14} result in a set consisting of all the classical states of Table 2.10:

$$\text{classical}(w_{14}) = w_1 \vee w_2 \vee w_3 \vee w_4 \vee w_5 \vee w_6 \vee w_7 \vee w_8.$$

Theorem 2.4.3. Where w is a paraconsistent state, $\text{val}(w) = \text{eu}(\text{classical}(w))$.

Proof.

- Say w has x B valuations.
- x Bs contribute $x \times \frac{1}{2^n}$ to the state's value.
- The set consisting of its corresponding classical states will have a cardinality of 2^x . These states will have $\frac{x2^x}{2}$ extra T valuations in total. That is, the Bs will be replaced by $\frac{x2^x}{2}$ Ts.
- $\frac{x2^x}{2}$ Ts contribute $\frac{x2^x}{2} \times \frac{1}{n}$ to the value of the set of states
- Now to show that $x \times \frac{1}{2^n} = \frac{x2^x}{2} \times \frac{1}{n} \times \frac{1}{2^x}$:

$$\begin{aligned} x \times \frac{1}{2^n} &= \frac{x2^x}{2} \times \frac{1}{n} \times \frac{1}{2^x} \\ \frac{x}{2^n} &= \frac{x2^x}{2 \times n \times 2^x} \\ \frac{x}{2^n} &= \frac{x}{2^n} \end{aligned}$$

□

Given this, we can reframe the situation for an agent choosing amongst the states corresponding to $h \wedge \neg h \wedge r \wedge \neg r \wedge w$ as follows.

Example 2.4

The original calculation was:

$$\text{eu}(h \wedge \neg h \wedge r \wedge \neg r \wedge w) = (\text{val}(w_{13}) \times \frac{1}{2}) + (\text{val}(w_{14}) \times \frac{1}{2})$$

With the modified way, $\text{val}(w_{13})$ and $\text{val}(w_{14})$ are replaced with 'sub' estimated utility calculations, involving the classical states they translate to. Here are the replacements:

- $\text{val}(w_{13}) = \text{eu}(w_1 \vee w_3 \vee w_5 \vee w_7) = \frac{2}{3}$
- $\text{val}(w_{14}) = \text{eu}(w_1 \vee w_2 \vee w_3 \vee w_4 \vee w_5 \vee w_6 \vee w_7 \vee w_8) = \frac{1}{2}$

So that the estimated utility is once again 0.583.

□

2.4.6 Adjusting State Description Utility Values

Given this decision-theoretic interpretation of the Tichy/Oddie method, one possible extension is to adjust the way in which utilities are assigned to state descriptions. As we have seen, the standard method uses a simple, linear utility function; for all levels l , there is a constant difference between l and $l + 1 / l - 1$. As a simple example of this extension, take the following two utility functions, where x stands for the standard linear utility of a state description and y its modified utility value:

$$(1) \quad y = x^2$$

$$(2) \quad y = \sqrt[2]{x}$$

What does a utility function such as (1) say in terms of state description distribution and informativeness? Roughly speaking, it favours high-valued state description efficiency: the more efficiently one statement gets to containing a high valued state description disjunct the better, where more efficiently means doing so with less state description disjuncts.

In the ranking of all possible 255 propositions in the 3-atom logical space, from highest to lowest informativeness, the original method using a standard linear utility and the method with this quadratic utility function agree on the first eight positions, after which differences start to appear. To illustrate the higher premium given to relatively small state collections involving w_1 , whilst the linear way places w_1, w_8 below w_1, w_2, w_8 , the quadratic way places w_1, w_8 above w_1, w_2, w_8 .

Here is another way to look at the effects of such adjustments. Take the formulas $A = h$ and $B = h \wedge r \wedge \neg w$. With the standard utility function, these are both given the same information measure: $\text{info}(A) = \text{info}(B) = 0.67$. In this case, the extra true atom in B is nullified by the false atom.

But suppose that things were to be set up so that one false atom should not detract symmetrically from one true atom, so that two true atoms and one false atom is better than one true atom. This would be captured under the utility function (2): $\text{info}(h \wedge r \wedge \neg w) > \text{info}(h)$.

Alternatively, if the detraction in value due to a false atom should exceed the increase in value due to a true atom, this would be captured under utility function (1): $\text{info}(h) > \text{info}(h \wedge r \wedge \neg w)$

Remark If the rate of increase of the utility function is sufficiently large or set it a certain way, it is even possible to get the covariation with logical strength amongst true statements condition to hold, although at the cost of a skew where a disproportionate number of states containing w_1 are placed at the top of the rankings.

The utility and decision theory literature is rich and this decision theoretic analysis of the Tichy/Oddie method (and in general other similar methods) opens the door up to further experimentation and the application of decision theoretic resources to truthlikeness/information methods.

2.4.7 Another Interpretation of the Tichy/Oddie Measure

In this section I would like to introduce another interpretation/translation of the Tichy/Oddie I have devised, which will be used for some results in the next chapter. It involves using a probability function over the logical space.

Take a statement A that corresponds to states W_A . Given these $|W_A|$ states, for each atom p , p is true in some number x of those states. If one of those $|W_A|$ states had to be chosen randomly, then the logical probability that p would be true in that state is $\frac{x}{|W_A|}$. This leads to the following method.

Given a space with n atoms $\{p_1, p_2, \dots, p_n\}$, $\text{info}_{TO}(A) = \frac{\sum_{i=1}^n \text{Pr}^*(p_i, A)}{n}$, where $\text{Pr}^*(p_i, A)$ is such that:

$$\text{Pr}^*(p_i, A) = \begin{cases} \text{Pr}(p_i|A) & \text{if } v(p_i) = \text{T} \\ \text{Pr}(\neg p_i|A) & \text{if } v(p_i) = \text{F} \end{cases}$$

See Appendix A Theorem A.0.2 for a proof that this function is equivalent to the standard method for the Tichy/Oddie measure.

2.5 Another Method to Quantify Semantic Information

In this section a new, computationally elegant method to measure information along truth-likeness lines is proposed. As will be seen, this *value aggregate* approach differs to both the Tichy/Oddie and Niiniluoto approaches, lying somewhere in between.

As has been well established by now, each statement corresponds to a set of states W in a logical space. Each of the states in this logical space can be valued, with the actual state being the uniquely highest valued. The rough idea behind the value aggregate approach is a simple one, with two factors concerning state distribution in W involved in determining the informativeness of a statement:

1. the greater the highest valued state of W the better for informativeness.
2. given a collection of states with a highest value member w , the fewer other states of lesser value than w the better for informativeness. Also, given other states of lesser value, the closer they are to w the better.

This determination is a two-tier process; first rank a statement according to the first factor and then apply the second. For example, the state description corresponding to the actual state (in our case $h \wedge r \wedge w$) ranks highest relative to the first factor. It then also ranks highest relative to the second factor because there are no other states of lesser value than w_1 in its state collection. Accordingly, its informativeness is maximal.

As another example, the statement $h \vee r$, which corresponds to states $w_1 - w_6$, is more informative than the statement $h \vee \neg r$, which corresponds to states $w_1 - w_4, w_7, w_8$, because although both have the same highest valued state and the same number of lesser valued states, the lesser valued states for $h \vee r$ are closer to the highest valued state.

This informal idea motivates and can be linked to the formal method that will now be detailed. To begin with, each state is once again assigned a value and ranked, where the value for a state w [$\text{val}(w)$] is determined using the following method:

- Let n stand for the number of propositional variables in the logical space
- Let t stand for the number of true atoms in a state w relative to the actual state
- $\text{val}(w) = \frac{t}{n \times 2^n}$

So in the case of our 3-atom logical space with w_1 being the actual state, each state is valued as follows:

- $\text{val}(w_1) = \frac{3}{24}$
- $\text{val}(w_2) = \text{val}(w_3) = \text{val}(w_5) = \frac{2}{24}$
- $\text{val}(w_4) = \text{val}(w_6) = \text{val}(w_7) = \frac{1}{24}$
- $\text{val}(w_8) = 0$

Now, given a statement A , its information measure is calculated using the following algorithm:

1. Determine the set of states W in which A holds ($W = \{w \mid w \models A\}$)
2. Place the members of W into an array X_1 of length $|W|$ and order the members of X_1 from lowest to highest value. This process is represented with a function `arraystates()`, so that $X_1 = \text{arraystates}(W)$.
3. Let X_2 stand for an empty array with 2^n spaces. Next, start by placing the first (lowest) element of X_1 in the first position of X_2 . In general, place the n^{th} element of X_1 in the n^{th} position of X_2 . Unless the statement being measured is a tautology, then $\text{length}(X_1) < \text{length}(X_2)$. So once the last element of X_1 has been reached, use this last element to fill in all the remaining places of X_2 . This process is represented with a function `lineup()`, so that $X_2 = \text{lineup}(X_1)$.
4. Finally, sum up all the values of each element of X_2 to get the information measure

$$\text{info}(A) = \text{sum}(X_2) = \text{sum}(\text{lineup}(X_1)) = \text{sum}(\text{lineup}(\text{arraystates}(W))).$$

With the introduction of this method, for the remainder of this work $\text{info}_{TO}()$ will be used for the Tichy/Oddie measure and $\text{info}_{VA}()$ will be used for the Value Aggregate measure. Following is an example of $\text{info}_{VA}()$.

Example 2.5

Let A be the statement h . Then:

- $W = \{w_1, w_2, w_3, w_4\}$

- $X_1 = [w_4, w_3, w_2, w_1]$
- $X_2 = [w_4, w_3, w_2, w_1, w_1, w_1, w_1, w_1]$
- $\text{info}_{VA}(h) = \text{sum}(X_2) = \text{val}(w_4) + \text{val}(w_3) + \text{val}(w_2) + 5\text{val}(w_1) = 0.8333$

□

Remark With a ranking (highest to lowest) of all 255 possible propositions (state collections), in the 3-atom space, the ordering of statements given by this method largely agrees with Niiniluoto’s min-sum measure (with γ and λ as given in Section 2.4.4) for a significant portion of the first half of ordered statements. In fact it agrees for the first 107 positions, after which telling differences emerge. Also, it agrees more with the Tichy/Oddie approach towards the bottom of the rankings.

Table 2.18 contains some results using the value aggregate method. They seem reasonable and it is evident that their spread is closer to the Niiniluoto spread than the Tichy/Oddie. Also, unlike the Tichy/Oddie approach, where the differences between h , $h \wedge r$ and $h \wedge r \wedge w$ are constant, with the value aggregate approach the first of these differences is greater than the second. This variation is akin to the quadratic nature of Floridi’s degrees of informativeness [65, p. 210].

For a logical space with n atoms, there are 2^{n^n} different propositions (sets of states). Table 2.19 lists an ordering of the different informativeness classes using the Value Aggregate measure for $n = 3$

2.5.1 Adequacy Conditions

As mentioned earlier in Section 2.4.3, Niiniluoto states a number of adequacy conditions “which an explicate of the concept of truthlikeness should satisfy” [133, p. 232]. An investigation into the applicability of these conditions to an account of semantic information quantification (whether all or only some apply) will not be pursued here; apart from perhaps an occasional comment, I do not intend to discuss matter. Suffice it to say, at the least most of them seem applicable.

Nonetheless, in order to initiate this new method I here include a list of these conditions plus a summary of how the value aggregate approach, along with the Tichy/Oddie and

#	Statement (A)	T/F	$\text{info}_{VA}(A)$
1	$h \wedge r \wedge w$	T	1
2	$h \wedge r$	T	0.9583
3	$h \wedge (r \vee w)$	T	0.9167
4	$h \wedge (\neg r \vee w)$	T	0.875
5	h	T	0.8333
6	$(h \wedge r) \vee w$	T	0.7917
7	$(h \wedge \neg r) \vee w$	T	0.75
8	$h \vee r$	T	0.7083
9	$h \wedge r \wedge \neg w$	F	0.6667
10	$h \vee \neg r$	T	0.625
11	$h \vee r \vee w$	T	0.625
12	$h \wedge \neg r$	F	0.625
13	$h \vee r \vee \neg w$	T	0.5833
14	$h \vee \neg r \vee \neg w$	T	0.5417
15	$(h \vee \neg w) \wedge \neg r$	F	0.5417
16	$\neg h$	F	0.5
17	$h \vee \neg h$	T	0.5
18	$\neg h \vee \neg r \vee \neg w$	F	0.4583
19	$h \wedge \neg r \wedge \neg w$	F	0.3333
20	$\neg h \wedge \neg r$	F	0.2917
21	$\neg h \wedge \neg r \wedge \neg w$	F	0
22	$h \wedge \neg h$	F	N/A

Table 2.18: Results using the value aggregate method.

	x	$\frac{x}{24}$	cardinality	# true statements	# false statements
1	24	1	1	1	0
2	23	0.96	3	3	0
3	22	0.92	6	6	0
4	21	0.88	11	11	0
5	20	0.83	15	15	0
6	19	0.79	18	18	0
7	18	0.75	20	20	0
8	17	0.71	18	18	0
9	16	0.67	22	15	7
10	15	0.63	32	11	21
11	14	0.58	34	6	28
12	13	0.54	31	3	28
13	12	0.5	22	1	21
14	11	0.46	7	0	7
15	8	0.33	7	0	7
16	7	0.29	7	0	7
17	0	0	1	0	1

Table 2.19: Sample list of informativeness classes using the Value Aggregate measure

Niiniluoto approaches, fare against them. This summary confirms that the value aggregate approach is similar but not equivalent to either of the other two approaches.

The presentation of these conditions will largely conform to their original form, so bear in mind for our purposes that $\text{Tr}()$ corresponds to $\text{info}()$ and as Δ decreases/increases $\text{info}()$ increases/decreases. Before listing the conditions, some terms need to be (re)established:

- Niiniluoto uses the term *constituent* as a more general term for *state descriptions*.
- A is used to refer to statements in the logical space in general. $\mathbf{I}_A$ is the set of numbers used to index the set of states corresponding to the statement A . For example, in our weather framework, the statement h corresponds to the states $\{w_1, w_2, w_3, w_4\}$, so letting A stand for h we have $\mathbf{I}_A = \{1, 2, 3, 4\}$.
- S_i is used to refer to the constituent (state description) that corresponds to state w_i .
- S_* is reserved for the state description that corresponds to the actual state w_* , so in our case $S_* = S_1$.
- $\mathbf{B}$ stands for the set of mutually exclusive and jointly exhaustive constituents. So in our case $\mathbf{B} = \{w_i \mid 1 \leq i \leq 8\}$. $\mathbf{I}$ is the set of numbers corresponding to $\mathbf{B}$.
- Δ_{ij} stands for $\Delta(w_i, w_j)$.
- An element w_j of $\mathbf{B}$ is called a Δ -complement of w_i , if $\Delta_{ij} = \max \Delta_{ik}$, $k \in \mathbf{I}$. Also, our weather framework example is Δ -complemented; if each $w_i \in \mathbf{B}$ has a unique Δ -complement w_j in $\mathbf{B}$, such that $\Delta_{ij} = 1$, the system $(\mathbf{B}, \Delta)$ is said to be Δ -complemented.

Here are the conditions:

- (M1) (**Range**) $0 \leq \text{Tr}(A, S_*) \leq 1$.
- (M2) (**Target**) $\text{Tr}(A, S_*) = 1$ iff $A = S_*$.
- (M3) (**Non-triviality**) All true statements do not have the same degree of truthlikeness, all false statements do not have the same degree of truthlikeness.
- (M4) (**Truth and logical strength**) Among true statements, truthlikeness covaries with logical strength.
- (a) If A and B are true statements and $A \vdash B$, then $\text{Tr}(B, S_*) \leq \text{Tr}(A, S_*)$.
- (b) If A and B are true statements and $A \vdash B$ and $B \not\vdash A$, then $\text{Tr}(B, S_*) < \text{Tr}(A, S_*)$.

- (M5) (***Falsity and logical strength***) Among false statements, truthlikeness does not covary with logical strength; there are false statements A and B such that $A \vdash B$ but $\text{Tr}(A, S_*) < \text{Tr}(B, S_*)$.
- (M6) (***Similarity***) $\text{Tr}(S_i, S_*) = \text{Tr}(S_j, S_*)$ iff $\Delta_{*i} = \Delta_{*j}$ for all S_i, S_j .
- (M7) (***Truth content***) If A is a false statement, then $\text{Tr}(S_* \vee A, S_*) > \text{Tr}(A, S_*)$.
- (M8) (***Closeness to the truth***) Assume $j \notin \mathbf{I}_A$. Then $\text{Tr}(A \vee S_j, S_*) > \text{Tr}(A, S_*)$ iff $\Delta_{*j} < \Delta_{\min}(A, S_*)$.
- (M9) (***Distance from the truth***) Let $\Delta_{*j} < \Delta_{*i}$. Then $\text{Tr}(S_j \vee S_i, S_*)$ decreases when Δ_{*i} increases.
- (M10) (***Falsity may be better than truth***) Some false statements may be more truthlike than some true statements.
- (M11) (***Thin better than fat***) If $\Delta_{*i} = \Delta_{*j} > 0$, $i \neq j$, then $\text{Tr}(S_i \vee S_j, S_*) < \text{Tr}(S_i, S_*)$.
- (M12) (***Ovate better than obovate***) If $\Delta_{*j} < \Delta_{*i} < \Delta_{*k}$, then $\text{Tr}(S_j \vee S_i \vee S_k, S_*)$ increases when Δ_{*i} decreases.
- (M13) (***Δ -complement***) $\text{Tr}(A, S_*)$ is minimal, if A consists of the Δ -complements of S_* .

Table 2.20 gives a summary of the measures against the adequacy conditions, where $+(\gamma)$ means that the measure satisfies the given condition with some restriction on the value of γ and (λ) ¹³:

The value aggregate method will satisfy those seeking a covariation of information and logical strength. As can be seen, it fails only M11. For example, $\Delta_{*2} = \Delta_{*3} = \Delta_{*5}$, yet $\text{info}(w_2 \vee w_3 \vee w_5) = \text{info}(w_2 \vee w_3) = \text{info}(w_2)$. The value aggregate method, like the Tichy/Oddie method, does not differentiate sets of states such as these. In terms of information, it could be argued that satisfaction of this condition is not essential; removing or adding a state description disjunct in this example does not inform an agent any more or less of the true state.

C-monotonicity

In [28] the following comparative notion of verisimilitude (truthlikeness) for c-theories (i.e. c-statements, Section 2.4.1) is given:

¹³See Appendix A.1 for proofs relating to the value aggregate approach

¹⁴Simply because there is no Δ function involved with this method

Condition	Niiniluoto (ms)	Tichy/Oddie (av)	Value Aggregate
M1	+	+	+
M2	+	+	+
M3	+	+	+
M4a	+	−	+
M4b	+	−	+
M5	$+(\gamma)$	+	+
M6	+	+	N/A ¹⁴
M7	+	+	+
M8	$+(\gamma)$	−	+
M9	+	+	+
M10	$+(\gamma)$	+	+
M11	+	−	−
M12	+	+	+
M13	$+(\gamma)$	+	+

Table 2.20: Measures against adequacy conditions.

Definition Given two c-theories T_1 and T_2 and the true state description C_* , T_2 is *more verisimilar* than T_1 - in symbols, $T_2 >_{vs} T_1$ - iff at least one of the following two conditions holds:

$$(M_t) \quad t(T_2, C_*) \supset t(T_1, C_*) \text{ and } f(T_2, C_*) \subseteq f(T_1, C_*)$$

$$(M_f) \quad t(T_2, C_*) \supseteq t(T_1, C_*) \text{ and } f(T_2, C_*) \subset f(T_1, C_*)$$

where $t()$ and $f()$ measure truth and falsity content respectively as in Section 2.4.1.

A verisimilitude measure is ‘conjunctively monotonic’ (c-monotonic) according to the following definition:

Definition A verisimilitude measure Vs is *c-monotonic* just in case Vs satisfies the following condition:

C-monotonicity. Given two c-theories T_1 and T_2 , if $T_2 >_{vs} T_1$ then $Vs(T_2) > Vs(T_1)$.

Both the basic features measure and the Tichy/Oddie measure are c-monotonic. As pointed out by the authors, a few known verisimilitude measures are not c-monotonic; Niiniluoto’s favoured ‘min-sum’ measure is one of them [28, p. 6].

It turns out that the value aggregate measure is also not c-monotonic:

Example 2.6

Take a 9-proposition logical space, consisting of propositions $p_1 - p_9$. Let:

- $C_* = p_1 \wedge p_2 \wedge p_3 \wedge p_4 \wedge p_5 \wedge p_6 \wedge p_7 \wedge p_8 \wedge p_9$
- $A_1 = p_1$
- $A_2 = p_1 \wedge \neg p_2$
- $t(A_1, C_*) = \{p_1\} \supseteq t(A_2, C_*) = \{p_1\}$
- $f(A_1, C_*) = \{\} \subset f(A_2, C_*) = \{\neg p_2\}$
- So $A_1 >_{V_s} A_2$
- But $\text{info}_{VA}(A_1) = 0.77 < \text{info}_{VA}(A_2) = 0.79167$

□

2.5.2 Adjusting Utilities

As was done with $\text{info}_{TO}()$, the value aggregate method can be extended to incorporate different utility functions for different purposes, in the same way as discussed in Section 2.4.6. As described earlier in Section 2.5, with the value aggregate measure the last element in the array X_1 is the highest valued and is used to fill in the remaining positions of X_2 . In the case of w_1 in particular, this means that a collection of states consisting of w_1 and relatively few other states will jump ahead quite quickly given the continued addition of the highest valued element.

The adoption of a non-linear utility function could be used to regulate this. For example, a simple logarithmic utility function such as $y = 20\log_2(x+1)$ (x being the value of the standard utility) places w_1, w_8 below w_1, w_3, w_5, w_7 , whereas with the original standard linear utility this order is reversed. Also to note, as long as the actual state is assigned the highest utility, the information measure will covary with logical strength for true statements, whatever utility function is used.

2.5.3 Misinformation

How can the value aggregate method be used to obtain a misinformation measure? One idea involves simply complementing the utility values for each state; where $\frac{t}{n \times 2^n}$ is the original utility of a state, its new utility $[\text{val}_m()]$ becomes: $\text{val}_m(w) = \frac{n-t}{n \times 2^n}$. The generation of X_2 is the same as the measure for $\text{info}_{VA}()$. What is different is that the values of $\text{val}_m()$ are used for the sum calculation. Let $\text{sum}_m()$ stand for this version of the sum calculation. Following is an example:

Example 2.7

Let A be the statement $w_1 \vee w_8$. Then:

- $W = \{w_1, w_8\}$
- $X_1 = [w_8, w_1]$
- $X_2 = [w_8, w_1, w_1, w_1, w_1, w_1, w_1, w_1]$
- $\text{misinfo}_{VA}(A) = \text{sum}_m(X_2) = \text{val}_m(w_8) + 7\text{val}_m(w_1) = \frac{3}{24} + (7 \times 0) = 0.125$

□

Like the Tichy/Oddie method, this is equivalent to $\text{misinfo}_{VA}(A) = 1 - \text{info}_{VA}(A)$. According to this approach, $\text{misinfo}(w_1 \vee w_8) = 0.125 < \text{misinfo}(w_4 \vee w_5) = 0.375$. On the other hand, the Tichy/Oddie approach gives $\text{misinfo}(w_1 \vee w_8) = \text{misinfo}(w_4 \vee w_5) = 0.5$. So as can be seen, unlike the Tichy/Oddie method, where there is a perfect symmetry between truth and falsity, the accumulation of values with the aggregate approach results in a skew towards the information measure. As a consequence there are significant differences between the two in terms of misinformation value calculations.

With this approach to calculating misinformation the following hold:

- $\text{info}(A) + \text{misinfo}(A) = 1$
- $\text{info}(A) > \text{info}(B) \Leftrightarrow \text{misinfo}(A) < \text{misinfo}(B)$
- if A and B are true statements and $A \vdash B$, then $\text{misinfo}(A) \leq \text{misinfo}(B)$

But unlike the Tichy/Oddie derived measure for misinformation given in Section 2.4.2, $\text{misinfo}(A) = \text{info}(\text{Rev}(A))$ does not generally hold. For example, $\text{misinfo}(w_2 \vee w_4 \vee w_8) = 1 - 0.5417 = 0.4583 \neq \text{info}(\text{Rev}(w_2 \vee w_4 \vee w_8)) = \text{info}(w_1 \vee w_5 \vee w_7) = 0.875$.

Another option with regards to a measure for misinformation is to complement the utility values for each state as before but replace truth aggregation with *falsity aggregation*, which involves ordering states from lowest to highest $\text{val}_m()$ values.

Example 2.8

Let A be the statement $\neg h \wedge \neg r$. Then:

- $W = \{w_7, w_8\}$
- $\text{val}_m(w_7) = \frac{2}{24}$
- $\text{val}_m(w_8) = \frac{3}{24}$
- $X_1 = [w_7, w_8]$
- $X_2 = [w_7, w_8, w_8, w_8, w_8, w_8, w_8, w_8]$
- $\text{misinfo}(A) = \text{sum}_f(X_2) = \text{val}_m(w_7) + 7\text{val}_m(w_8) = \frac{2}{24} + (7 \times \frac{3}{24}) = 0.9583$

□

This method $[\text{misinfo}_2()]$ differs significantly to the previous one $[\text{misinfo}_1()]$, though does sometimes deliver the same result. For example, let $A = h \wedge r \wedge \neg w$, the state description for w_2 . $\text{misinfo}_1(A) = \frac{1}{3} = \text{misinfo}_2(A)$.

With this approach:

- $\text{info}(A) + \text{misinfo}(A) = 1$ does not generally hold. For example $\text{info}_{VA}(w_2 \vee w_8) + \text{misinfo}_2(w_2 \vee w_8) = 0.5833 + 0.9167 > 1$
- $\text{info}(A) > \text{info}(B) \Leftrightarrow \text{misinfo}(A) < \text{misinfo}_2(B)$ does not generally hold. For example, $\text{info}(w_1 \vee w_8) = \text{misinfo}_2(w_1 \vee w_8) = 0.875 > \text{info}(w_4 \vee w_5) = \text{misinfo}_2(w_4 \vee w_5) = 0.625$.
- $\neg(\forall A)(\forall B)[v(A) = t \wedge v(B) = t \wedge A \vdash B] \Rightarrow \text{misinfo}_2(A) \leq \text{misinfo}_2(B)$

But unlike $\text{misinfo}_1()$, $\text{misinfo}_2(A) = \text{info}(\text{Rev}(A))$ does hold:

Theorem 2.5.1. $\text{info}_{VA}(A) = \text{misinfo}_2(\text{Rev}(A))$

Proof. Let $\text{info}_{VA}(A) = \text{sum}(\text{lineup}(X_1))$. Replacing each element w of $\text{lineup}(X_1)$ with $w' = \text{Rew}(w)$ and applying $\text{val}_m()$ gives $\text{misinfo}_2(\text{Rev}(A))$, since $\text{val}(w) = \text{val}_m(w')$. $\square$

Unlike the other $\text{misinfo}()$ measures, since it is not necessarily the case that $\text{info}_{VA}(A) + \text{misinfo}_2(A) = 1$, there are some particularities. Firstly, $\text{info}(A) + \text{misinfo}(A)$ is no longer uniform across all statements. Where w_1 is the true state, it takes a maximum value when $A = w_1 \vee w_8$.

Also, the independence between $\text{info}_{VA}()$ and $\text{misinfo}_2()$ suggests the possibility of combining them into a weighted measure such as the following:

$$\text{im}(A) = \gamma \text{info}_{VA}(A) - \lambda \text{misinfo}_2(A)$$

where $0 \leq \gamma, \lambda \leq 1$. This weighting is used to regulate the preference between truth and falsity:

- $\text{info}_{VA}(w_1 \vee w_8) = \text{misinfo}_2(w_1 \vee w_8) = 0.875$
- $\text{info}_{VA}(w_4 \vee w_5) = \text{misinfo}_2(w_4 \vee w_5) = 0.625$
- if $\gamma < \lambda$ then $\text{im}(w_1 \vee w_8) < \text{im}(w_4 \vee w_5)$, which indicates that the falsity of $w_1 \vee w_8$ is ‘punished’ more than its truth is ‘rewarded’.
- if $\gamma > \lambda$ then $\text{im}(w_1 \vee w_8) > \text{im}(w_4 \vee w_5)$, which indicates that the truth of $w_1 \vee w_8$ is ‘rewarded’ more than its falsity is ‘punished’.

Remark $\text{im}()$ is c-monotonic

2.6 Estimated Information

Within the literature a distinction is made between the *semantic* and *epistemic* problems of truthlikeness [183, p. 121]:

- *The semantical problem:* “What do we mean if we claim that the theory X is closer to the truth than the theory Y ?”
- *The epistemic problem:* “On what evidence are we to believe that the theory X is closer to the truth than the theory Y ?”

The focus of this chapter has been the semantical problem of information quantification; what do we mean when we say that statement A is more informative than statement B ? The epistemic problem relates to estimating informativeness; given some partial evidence E , are we to estimate that statement A is more or less informative than statement B ?

Of course, an account of the semantical problem is primary, with any account of the epistemic problem being secondary, based on a semantical foundation. Nonetheless a method to estimate informativeness is also very important.

In practice judgements of informativeness are going to be made with limited evidence, without knowledge of the complete truth (the one true state description). If an agent already knows the complete truth in a domain of inquiry then there would be no new information for them to acquire anyway, no statement in the domain of inquiry could be informative for them. In such scenarios, the actual informativeness of any statement is already known, or can be calculated.

In general though, calculations of informativeness are of interest to agents who do not know the complete truth and who are seeking new information. When such agents must choose amongst a set of different statements, their aim is to choose the statement that they estimate will be the most informative (relative to the actual state, which the agent has limited epistemic access to). In making this estimation and choice the agent will often already possess some evidence and this evidence will rule out certain possible states from consideration in the calculations.

The standard formula for estimated utility in decision theory can be used to calculate the expected informativeness of a statement A given prior evidence E ¹⁵:

$$\text{info}_{est}(A|E) = \sum_{i=1}^n \text{info}(A, S_i) \times \text{pr}(S_i|E)$$

n stands for the number of possible states in the logical space and S_i stands for the state description corresponding to state i .

¹⁵This decision theoretic analysis raises the possibility of working with other strategies such as maximax, maximin and minimax regret.

Given a certain piece of information as evidence, the difference between actual and estimated informativeness calculations for a statement can be marked. For example:

- $\text{info}_{TO}(h \wedge r \wedge w) = 1$
- $\text{info}_{TO-est}(h \wedge r \wedge w | h \vee \neg r \vee \neg w) = 0.48$
- $\text{info}_{TO}(\neg h \wedge \neg r \wedge \neg w) = 0$
- $\text{info}_{TO-est}(\neg h \wedge \neg r \wedge \neg w | h \vee \neg r \vee \neg w) = 0.52$

Interestingly, although the completely true $h \wedge r \wedge w$ has a maximum measure of 1 and the completely false $\neg h \wedge \neg r \wedge \neg w$ has a minimum measure of 0, even certain true statements used as evidence in estimation calculations will give $\neg h \wedge \neg r \wedge \neg w$ a higher estimated information measure than $h \wedge r \wedge w$.

Using a Bayesian measure of confirmation such as the following:

$$C(E, A) = \Pr(A|E) - \Pr(A)$$

we can also measure estimated information change:

$$\text{info}_{est}(A|E) - \text{info}_{est}(A|\top)$$

or more generally, with D standing for an agent's prior data, we have

$$\text{info}_{est}(A|E \wedge D) - \text{info}_{est}(A|D)$$

Clearly confirmation and estimated information change do not covary:

Example 2.9

Whilst $w_1 \vee w_7 \vee w_8$ confirms w_1 , it results in a negative estimated information change.

- $C(w_1 \vee w_7 \vee w_8, w_1) = \frac{1}{3} - \frac{1}{8} = 0.2083$
- $\text{info}_{est}(w_1 | w_1 \vee w_7 \vee w_8) - \text{info}_{est}(w_1) = 0.43 - 0.5 = -0.07$

□

Theorem 2.6.1. For any statement A , $\text{info}_{TO-est}(A|\top) = 0.5$

Proof. Given a uniform assignment of logical probability $\frac{1}{n}$ amongst the n states, we can factor out $\text{pr}(S_i|E)$. So it will suffice to show that $\sum_{i=1}^n \text{info}_{TO}(A, S_i) = \frac{n}{2}$.

Take $\text{Rew}()$ and $\text{Rev}()$ to be the as defined in Section 2.4.2. It is easy to see that for all statements A , $\text{info}_{TO}(A) + \text{info}(\text{Rev}(A)_{TO}) = 1$. Since each state w_i (represented by state description S_i) is uniquely paired off with its counterpart state $\text{Rew}(w_i)$ ($\frac{n}{2}$ such pairs), and that $\text{info}_{TO}(S_i) + \text{info}_{TO}(\text{Rev}(S_i)) = 1$, the total is simply $\frac{n}{2}$. □

This result of uniformity does not hold for the value aggregate approach: $\neg(\forall x)(\forall y)(\text{info}_{VA-est}(x|\top) = \text{info}_{VA-est}(y|\top))$

Example 2.10

$$\text{info}_{VA-est}(w_1|\top) = \text{info}_{VA-est}(w_8|\top) = 0.5 < \text{info}_{VA-est}(w_1 \vee w_8|\top) = 0.6875. \quad \square$$

Thus according to info_{VA} , given nothing else, it would be better to commit to the disjunction of two converse states rather than either of the individual states. From this result it follows that for $\text{info}_{VA}()$ the statement with the highest estimated informativeness need not be a state description.

We shall now briefly look at the topic of combining measures before concluding this chapter.

2.7 Combining Measures

One option is to adopt a hybrid system combining CSI and truthlikeness measures. As was seen, the CSI approach works well when its application is confined solely to true statements. So perhaps a system which applies the CSI approach to true statements and a truthlikeness approach to false statements could be used.

Furthermore, apart from such hybrid systems, there is the possibility of combining calculations from both approaches into one metric. For example, an incorporation of the CSI approach into the value aggregate approach could help distinguish between, say w_2 and w_2, w_3 , which as we saw are given the same measure using the value aggregate approach.

This type of combination could also be used to extend the Tichy/Oddie method, as the following example demonstrates.

Example 2.11

Take a situation within the weather example where $h \wedge r$ is evidence upon which estimation calculations are to be based. Is there any difference between opting for $h \wedge r \wedge w$ or $h \wedge r \wedge \neg w$ instead of $h \wedge r$? Or to put it another way, is there any difference between taking a gamble on which of w and $\neg w$ is true or sticking with a conservative policy of choosing neither? This situation is captured in the decision problem shown in Table 2.21. The utilities are the $\text{info}_{TO}()$ values calculated using the listed state as the actual one.

	w_1	w_2
$h \wedge r \wedge w$	1	0.67
$h \wedge r$	0.83	0.83
$h \wedge r \wedge \neg w$	0.67	1
$\text{Pr}()$	0.5	0.5

Table 2.21: A decision problem using info_{TO}

According to estimated information calculations, the estimated utility for each statement is 0.83. Thus there is no difference between opting for $h \wedge r \wedge w$ or $h \wedge r \wedge \neg w$ instead of $h \wedge r$.

In order to capture an agent's preference to either take a gamble on extra true content or to adopt a conservative policy and suspend judgement, one simple option is to construct something like the following measure that is a combination of content and truthlikeness:

$$\text{infocont}(A) = \text{info}(A) + \gamma \text{cont}(A)$$

where

- $0 < \gamma \leq 1$, if gamble on true content is preference
- $-1 \leq \gamma < 0$, if conservative policy is preference

So if $\gamma = 1$ then a gamble on true content is the preference with the resulting decision problem tabulated in Table 2.22. For each formula A and each state w , the utility is the addition of the $\text{info}_{TO}(A)$ value calculated against w and $\text{cont}_{cond}(A|h \wedge r)$ ¹⁶.

	w_1	w_2
$h \wedge r \wedge w$	$1 + \frac{1}{2}$	$\frac{2}{3} + \frac{1}{2}$
$h \wedge r$	$\frac{5}{6} + 0$	$\frac{5}{6} + 0$
$h \wedge r \wedge \neg w$	$\frac{2}{3} + \frac{1}{2}$	$1 + \frac{1}{2}$
$\text{Pr}()$	0.5	0.5

Table 2.22: Decision problem with $\gamma = 1$

Calculations become:

- $\text{infocont}_{est}(h \wedge r \wedge w|h \wedge r) = \text{infocont}_{est}(h \wedge r \wedge \neg w|h \wedge r) = \frac{4}{3}$
- $\text{infocont}_{est}(h \wedge r|h \wedge r) = \frac{5}{6}$

Accordingly, $h \wedge r \wedge w$ and $h \wedge r \wedge \neg w$ have a greater estimated utility than $h \wedge r$.

If on the other hand $\gamma = -1$ then there is an aversion to risk and the resulting decision problem is tabulated in Table 2.23

	w_1	w_2
$h \wedge r \wedge w$	$1 - \frac{1}{2}$	$\frac{2}{3} - \frac{1}{2}$
$h \wedge r$	$\frac{5}{6} - 0$	$\frac{5}{6} - 0$
$h \wedge r \wedge \neg w$	$\frac{2}{3} - \frac{1}{2}$	$1 - \frac{1}{2}$
$\text{Pr}()$	0.5	0.5

Table 2.23: Decision problem with $\gamma = -1$

Calculations become:

- $\text{infocont}_{est}(h \wedge r \wedge w|h \wedge r) = \text{infocont}_{est}(h \wedge r \wedge \neg w|h \wedge r) = \frac{1}{3}$
- $\text{infocont}_{est}(h \wedge r|h \wedge r) = \frac{5}{6}$

¹⁶Refer to Section 2.1.

Accordingly, $h \wedge r \wedge w$ and $h \wedge r \wedge \neg w$ have a lower estimated utility than $h \wedge r$.

□

2.8 Conclusion

The main point of this chapter has been the advocacy and development of quantitative accounts of semantic information based on the notion of truthlikeness. This is in contrast to traditional probabilistic approaches such as the CSI account given by Bar-Hillel and Carnap. CSI measures prove useful in certain applications and effectively capture certain aspects of information, particularly the deep intuition that information is proportional to uncertainty reduction.¹⁷ However as was shown, there are certain features of CSI which are at odds with ordinary senses of information, certain applications of information and certain adequacy criteria that we have established. Instead for our purposes, CSI can be seen as a measure of semantic content.

On the other hand a truthlikeness approach to quantifying semantic information accords with these ordinary senses, meets these application requirements and satisfies these criteria. Furthermore, such an approach naturally accommodates the corresponding notion of misinformation. As was seen, fortunately there is already a significant body of work on truthlikeness to draw from. This approach to truthlikeness has also been extended to first order and even second order logic. One way is to use Hintikka's distributive normal form for first-order logic, which provides a way to define first-order state descriptions/constituents [138, 133]. Once a metric for constituents is provided, truthlikeness measurements can be defined. Thus the extension of information as truthlikeness approaches to richer systems beyond classical propositional logic is one area to explore. As evidenced in this chapter there is still room for further investigation, experimentation and the development of new quantification methods, with a specific focus on information.

¹⁷The ability of CSI to accommodate the notion of *conditional informativeness* is another pro for this approach.

Chapter 3

Agent-Relative Informativeness

In the previous chapter a quantitative account of semantic information based on truthlikeness was proposed; statement A yields more information or is more informative than statement B when A contains more truth or is closer to the whole truth than B . Given a way to measure the information yield of a statement A , an important distinction to make is that between the information yield or *absolute* informativeness of A and its informativeness *relative* to a particular agent. As the simplest of examples, take three true propositions p_1 , p_2 and p_3 . Although the statement $p_1 \wedge p_2$ has a greater quantitative information measure than p_3 , given an agent that already has $p_1 \wedge p_2$ in their database, the statement p_3 is going to be more informative than $p_1 \wedge p_2$ for that agent.

In this chapter, a formal account of agent-relative informativeness is developed. How informative some input statement is for an agent will not only depend upon its (1) information measure, but also on (2) what content the agent already has and (3) what they do with the input if they accept it. In order to deal with (2) and (3), some framework for belief revision is required. Thus agent-relative informativeness here involves a combination of truthlikeness information measures and belief revision. As it so happens, within the last few years there has been some interest in investigating the relationship between the truthlikeness (verisimilitude) and belief revision programs [93, 135]. Continuing on from the previous chapter, in this chapter the approaches to truthlikeness/information focused on are the Tichy/Oddie and Value Aggregate methods. The belief revision methods employed fall under the AGM framework.

Apart from a general outline of the ideas associated with agent-relative informativeness, two further contributions of this chapter are:

- Some results on the behaviour of the belief revision operations of expansion, revision and contraction with regards to truthlikeness information measurements.
- A couple of ways to deal with conflicting sources of input, at least some of which by definition are going to be providing misinformation. This includes (1) combining screened belief revision with estimated information measurement and (2) construction of a paraconsistent approach to belief revision.

3.1 AGM Belief Change

The AGM framework (so called after its three originators: Carlos Alchourron, Peter Gardenfors, and David Makinson [7]) is the dominant framework for belief revision. This research program is quite large, with many extensions and modifications based on Alchourron, Gardenfors and Makinson's seminal work appearing over the last couple of decades. In this section a brief overview of its core aspects will be provided, enough to suffice for the purposes of this chapter ¹. The basic AGM operations will be explained in terms of tools and concepts used in the truthlikeness program.

In the AGM framework, agent belief states are represented using sets of formal statements closed under logical consequence (*belief sets*). An alternative approach uses sets of statements that are not logically closed (*belief bases*) instead. Throughout this work we use the former.

In the literature, $\mathbf{K}$ is generally used to denote a belief set. In this paper, the set of statements corresponding to belief set $\mathbf{K}$ will have a propositional formula representation, the italicised K . For example, $\mathbf{K} = \{h, r \vee w\}$ can be represented as $K = h \wedge (r \vee w)$.

For any set of sentences $\mathbf{K}$, $\text{Cn}(\mathbf{K})$ is the set of classical logical consequences of $\mathbf{K}$ ($B \in \text{Cn}(\mathbf{K})$ is the same as $K \vdash B$). $\mathbf{K}$ is a belief set if and only if $\mathbf{K} = \text{Cn}(\mathbf{K})$. Finally, $\text{Cn}(\phi)$ is the set of tautologies.

There are three core AGM operations, the first two being *additive* and the last one being *eliminative*:

- *Expansion* of $\mathbf{K}$ with statement A , represented as $\mathbf{K} + A$. This involves a statement A being added to $\mathbf{K}$ when A is compatible with the content of $\mathbf{K}$. So in this operation nothing needs to be done to $\mathbf{K}$ prior to the addition of A .
- *Revision* of $\mathbf{K}$ with A , represented as $\mathbf{K} * A$. This involves a statement A being added

¹For more on this topic, consult [96] and [81]

to $\mathbf{K}$ when A conflicts with some of the content of $\mathbf{K}$ ($\neg A \in \mathbf{K}$). So some content from $\mathbf{K}$ must be removed to ensure that the resulting belief set is consistent.

- *Contraction* of $\mathbf{K}$ with A , represented as $\mathbf{K} \div A$. This involves removing A from $\mathbf{K}$, so that $A \notin \mathbf{K} \div A$.

Logic is enough for expansion, the simplest of the operations. It can be seen simply as the conjunction of K and A ($K \wedge A$): $\mathbf{K} + A = \text{Cn}(\mathbf{K} \cup \{A\})$.

The other two operations are more complex and require extra-logical considerations. One guiding principle is that of *informational economy*: the information (as content) lost when giving up beliefs should be kept minimal. Furthermore, collections of rationality postulates or adequacy conditions for contraction and revision are a cornerstone of the AGM framework.

The eight postulates for contraction are:

- *Closure*: $\mathbf{K} \div A = \text{Cn}(\mathbf{K} \div A)$
- *Success*: If $A \notin \text{Cn}(\phi)$, then $A \notin \text{Cn}(\mathbf{K} \div A)$
- *Inclusion*: $\mathbf{K} \div A \subseteq \mathbf{K}$
- *Vacuity*: If $A \notin \text{Cn}(\mathbf{K})$, then $\mathbf{K} \div A = \mathbf{K}$
- *Extensionality*: If $A \leftrightarrow B \in \text{Cn}(\phi)$, then $\mathbf{K} \div A = \mathbf{K} \div B$
- *Recovery*: $\mathbf{K} \subseteq (\mathbf{K} \div A) + A$
- *Conjunctive inclusion*: If $A \notin \mathbf{K} \div (A \wedge B)$, then $\mathbf{K} \div (A \wedge B) \subseteq \mathbf{K} \div A$
- *Conjunctive overlap*: $(\mathbf{K} \div A) \cap (\mathbf{K} \div B) \subseteq \mathbf{K} \div (A \wedge B)$

The eight postulates for revision are:

- *Closure*: $\mathbf{K} * A = \text{Cn}(\mathbf{K} * A)$
- *Success*: $A \in \mathbf{K} * A$
- *Inclusion*: $\mathbf{K} * A \subseteq \mathbf{K} + A$
- *Vacuity*: If $\neg A \notin \mathbf{K}$, then $\mathbf{K} * A = \mathbf{K} + A$
- *Consistency*: $\mathbf{K} * A$ is consistent if A is consistent

- *Extensionality*: If $A \leftrightarrow B \in \text{Cn}(\phi)$, then $\mathbf{K} * A = \mathbf{K} * B$
- *Superexpansion*: $\mathbf{K} * (A \wedge B) \subseteq (\mathbf{K} * A) + B$
- *Subexpansion*: If $\neg B \notin \text{Cn}(\mathbf{K} * A)$, then $(\mathbf{K} * A) + B \subseteq \mathbf{K} * (A \wedge B)$

An operation $\div$ is a *transitively relational partial meet contraction* if and only if it satisfies the contraction postulates above. Likewise, an operation is a *transitively relational partial meet revision* if and only if it satisfies the revision postulates above.

Both contraction and revision can be defined in terms of the other. The *Levi identity* defines revision in terms of contraction:

$$\text{Levi identity: } \mathbf{K} * A = (\mathbf{K} \div \neg A) + A$$

The *Harper identity* defines contraction in terms of revision:

$$\text{Harper identity: } \mathbf{K} \div A = (\mathbf{K} * \neg A) \cap \mathbf{K}$$

Although these postulates form the foundation for the contraction and revision operators, they don't explicitly prescribe constructive models for contraction and revision. *Entrenchment-based contraction/revision* is one way to go about this. The entrenchment-based model for contraction/revision is equivalent to transitively relational partial meet contraction/revision.

Another constructive model in accordance with the AGM postulates is Grove's sphere system [92]. This model can be constructed using tools from the truthlikeness program as follows [135, p. 196]. Any $\mathbf{K}$ corresponds to a set of models $\mathbf{M}_K$ in propositional logic. This set of models corresponds to the centre sphere and can be expressed in distributive normal form as $K \equiv \bigvee_{i \in \mathbf{K}\#} C_i$, where each C_i is a constituent or state description for a particular state (model) in $\mathbf{M}_K$ and $\mathbf{K}\#$ is a set of state indices for the set of states $\mathbf{M}_K$. The distances of other states (their places on the spheres) from the centre states (i.e. models corresponding to $\mathbf{K}$) can be determined with a suitable distance function.

In our case the distance between two states is simply the number of atomic element valuations on which they differ (i.e. Hamming distance). Thus essentially we can use the $\min()$ measure introduced in Section 2.4.4 of Chapter 2. The distance $\text{dis}_s(\mathbf{K})$ of a state s from $\mathbf{K}$ is defined as $\text{dis}_s(\mathbf{K}) = \min(\{\Delta(s, w) | w \in \mathbf{M}_K\})$. Given some input A with models $\mathbf{M}_A$, the set $C_K(A)$ of the closest states to $\mathbf{K}$ entailing A is defined as:

$$C_K(A) = \{s \in \mathbf{M}_A : \text{dis}_s(\mathbf{K}) \leq \text{dis}_w(\mathbf{K})\} \text{ for all } w \in \mathbf{M}_A$$

By using these notions we get the following definitions:

- $\mathbf{K} + A = K \wedge A =_{\text{def}} \bigvee_{i \in K \cap A} C_i$
- $\mathbf{K} * A =_{\text{def}} \bigvee_{i \in C_K(A)} C_i$
- $\mathbf{K} \div A =_{\text{def}} K \vee (K * \neg A)$ [Alternatively, $\mathbf{K} \div A$ can be defined as $\mathbf{M}_K \cup C_K(\neg A)$]

We also have monotonicity for each operator:

- $A \vdash B \Rightarrow A + C \vdash B + C$
- $A \vdash B \Rightarrow A \div C \vdash B \div C$
- $A \vdash B \Rightarrow A * C \vdash B * C$

Remark This construction is equivalent to Dalal's update semantics [41].

Whilst the terms 'belief set' and 'theory' are generally used in the (AGM) belief revision literature, in this work the term 'database' will be opted for. So henceforth $\mathbf{D}$ will be used to represent a database (replacing $\mathbf{K}$) and D will stand for its propositional formula representation. $\mathbf{D}$ and D will be used interchangeably and as determined by context. Now to some examples of revision and contraction using the sphere models system just outlined.

Example 3.1

Take a propositional space with 2 atoms, p and q . $D = p \wedge q$. Since $n = 2$, each atomic element difference between two states contributes $\frac{1}{2}$ to the total difference. Here are the truth table for this space and sphere orderings based on Δ :

	p	q		
w ₁	T	T	Δ	set of states
w ₂	T	F	0 (centre sphere)	w ₁
w ₃	F	T	$\frac{1}{2}$	w ₂ , w ₃
w ₄	F	F	1	w ₄

- If the agent accepts the input $A = \neg(p \wedge q)$, then since $D \wedge \neg(p \wedge q) \vdash \perp$, the agent must perform a revision operation. $C_A(\neg(p \wedge q)) = \{w_2, w_3\}$; these two states are in the sphere closest to the centre sphere and they both satisfy $\neg(p \wedge q)$. Hence $D * \neg(p \wedge q) = (p \wedge \neg q) \vee (\neg p \wedge q)$.
- If an agent performs the operation $D \div (p \wedge q)$, the resulting set of states is $\{w_1\} \cup \{w_2, w_3\}$.

□

Example 3.2

For this example we temporarily depart from the standard distance function used throughout this chapter and see what happens when atomic weight assignments are disuniform (See Section 2.4.2 of Chapter 2).

If p was considered of greater value than q by the agent, then things would turn out differently. If instead p was assigned a weight of $\frac{3}{4}$ and q a weight of $\frac{1}{4}$, the sphere ordering would be:

Δ	set of states
0	w_1
$\frac{1}{4}$	w_2
$\frac{3}{4}$	w_3
1	w_4

- $C_A(\neg(p \wedge q)) = \{w_2\}$, so $\mathbf{D} * \neg(p \wedge q) = p \wedge \neg q$.
- If an agent performs the operation $\mathbf{D} \div (p \wedge q)$, the resulting set of states is $\{w_1, w_2\}$; $p \in \mathbf{D} \div (p \wedge q)$ and $q \notin \mathbf{D} \div (p \wedge q)$.

In this case, p is more entrenched than q

□

3.2 Combining Information Measurement and Belief Revision

The combination of truthlikeness information measurement with belief revision raises new matters to investigate. Let us start with the four possible assignments concerning the truth/falsity of the database D and the input A :

1. D is true and A is true
2. D is false and A is true
3. D is true and A is false
4. D is false and A is false

A first question to ask is whether or not true input accepted via the additive operations is always guaranteed to result in a higher information measure. More generally, for each of these four assignments, do the results of $D + A$, $D * A$ and $D \div A$ have lower or greater information measures compared to D ? For what assignments is

- $\text{info}(D) > \text{info}(D + A) \mid \text{info}(D * A) \mid \text{info}(D \div A)$ ²

and for what assignments is

- $\text{info}(D) < \text{info}(D + A) \mid \text{info}(D * A) \mid \text{info}(D \div A)$

To these questions we now turn.³ Examples will be based on the weather propositional space of Table 3.1 that was introduced in Section 2.1 of Chapter 2, with w_1 once again the actual state.

3.2.1 True Database Content and True Input

Theorem 3.2.1. If D is true and A is true then $\text{info}_{VA}(D + A) > \text{info}_{VA}(D)$.

²‘ $\mid$ ’ just separates alternatives (OR), used here like it is for notation techniques for context-free grammars.

³Trivial cases where the operation does not result in a change of the database contents (e.g. $\text{info}(h \wedge r) + h = \text{info}(h \wedge r)$) are excluded.

State	h	r	w
w ₁	T	T	T
w ₂	T	T	F
w ₃	T	F	T
w ₄	T	F	F
w ₅	F	T	T
w ₆	F	T	F
w ₇	F	F	T
w ₈	F	F	F

Table 3.1: Truth table for example logical space

Proof. Since $D + A \vdash D$ and $D \not\vdash D + A$, this result simply follows from the fact that $\text{info}_{VA}()$ covaries with logical strength amongst true statements: $(v(A) = \text{T}) \wedge (v(B) = \text{T}) \wedge (A \vdash B) \wedge (B \not\vdash A) \Rightarrow \text{info}_{VA}(A) > \text{info}_{VA}(B)$ [Chapter 2, Section 2.5.1]. $\square$

Theorem 3.2.2. If D is true and A is true then $\text{info}_{VA}(D \div A) \leq \text{info}_{VA}(A)$.

Proof. This simply follows from the covariation of $\text{info}_{VA}()$ with logical strength amongst true statements and the basic AGM postulate of Inclusion. $\square$

Theorem 3.2.3. If D is true and A is true then either of (1) $\text{info}_{TO}(D + A) > \text{info}_{TO}(D)$ or (2) $\text{info}_{TO}(D + A) < \text{info}_{TO}(D)$ are possible.

Proof. (1) is obvious. For (2), since covariation with logical strength amongst true statements fails for $\text{info}_{TO}()$ [Chapter 2, Section 2.4.3], there are cases where $\text{info}_{TO}(D + A) < \text{info}_{TO}(D)$. $\square$

Theorem 3.2.4. If D is true and A is true then either of (1) $\text{info}_{TO}(D \div A) < \text{info}_{TO}(D)$ or (2) $\text{info}_{TO}(D \div A) > \text{info}_{TO}(D)$ are possible.

Proof. (1) is obvious. For (2), since covariation of $\text{info}_{TO}()$ with logical strength amongst true statements fails, it is possible that $\text{info}_{TO}(D \div A) > \text{info}_{TO}(D)$. Here is an example: $D = (h \vee w) \wedge (\neg r \vee w)$ and $\text{info}_{TO}(D) = 0.6$. $A = \neg r \vee w$ so $D \div A = h \vee w$ and $\text{info}_{TO}(h \vee w) = 0.61$. Therefore $\text{info}_{TO}(D \div A) > \text{info}_{TO}(D)$. $\square$

3.2.2 False Database Content and True Input

Theorem 3.2.5. For $\text{info}_{VA}()$ and $\text{info}_{TO}()$, if D is false and A is true then either of (1) $\text{info}(D + A) > \text{info}(D)$ or (2) $\text{info}(D + A) < \text{info}(D)$ are possible.

Proof. (1) is obvious. For (2) take the example where $D = \neg h \wedge \neg r$ and $A = h \vee r \vee \neg w$ so that $D + A = \neg h \wedge \neg r \wedge \neg w$.

□

Theorem 3.2.6. If D is false and A is true then either of (1) $\text{info}(D * A) > \text{info}(D)$ or (2) $\text{info}(D * A) < \text{info}(D)$ are possible.

Proof. (1) is obvious. For (2) take the example where $D = h \wedge \neg r \wedge \neg w$ and $A = (h \wedge r \wedge w) \vee (\neg h \wedge \neg r \wedge \neg w)$, so that $D * A = \neg h \wedge \neg r \wedge \neg w$.

□

Theorem 3.2.7. If D is false and A is true then either of (1) $\text{info}(D \div A) < \text{info}(D)$ or (2) $\text{info}(D \div A) > \text{info}(D)$ are possible.

Proof. (1) is obvious. For (2) take the example where $D = h \wedge \neg r \wedge \neg w$ and $A = \neg r \vee w$, so that $D \div A = w_2 \vee w_4$.

- $\text{info}_{VA}(w_2 \vee w_4) = 0.625 > \text{info}_{VA}(h \wedge \neg r \wedge \neg w) = 0.3333$.
- $\text{info}_{TO}(w_2 \vee w_4) = 0.5 > \text{info}_{TO}(h \wedge \neg r \wedge \neg w) = 0.3333$.

□

That the addition of true input to a database does not necessarily lead to an increase of its information measure is not something particular to these two functions. In general, as some of the examples show, given any plausible account of truthlikeness and rational account of belief revision, expansions (+) and revisions (*) of a database with some true input are not guaranteed to increase the database's truthlikeness information yield. We also have seen that the elimination of true input does not necessarily lead to a decrease of information yield. Take the following definitions of $?_1$ and $?_2$ expressed in Backus-Naur Form:

- $\langle ?_1 \rangle ::= "+" \mid "*" \mid "\div"$
- $\langle ?_2 \rangle ::= "<" \mid ">"$

Out of the two assignments looked at so far (D true, A true and D false, A true), it can be seen that universal theorems of the form

$$\text{info}(D \ ?_1 \ A) \ ?_2 \ \text{info}(D)$$

only hold for $\text{info}_{VA}()$ and only for the following two cases:

- (1) $v(D) = v(A) = \text{T}$ and $?_1 = "+"$ and $?_2 = ">"$
- (2) $v(D) = v(A) = \text{T}$ and $?_1 = "\div"$ and $?_2 = "<"$

In fact, none of the remaining true/false assignment combinations for D and A give universal theorems of the form above ⁴.

3.2.3 True Inputs Guaranteed to Increase Information Yield

We have just seen that in general expansion and revision with true input is not guaranteed to increase information yield. This prompts investigation into restricted classes for D and A which are guaranteed to result in an increase.

Regarding the basic feature approach to truthlikeness outlined in Section 2.4.1 of Chapter 2, some theorems are given in [93], which it is suggested “hold for any plausible verisimilitude measure defined on propositional languages” [93, p. 58]. We shall now look at how $\text{info}_{VA}()$ and $\text{info}_{TO}()$ fare against these theorems.

We begin with some requisite definitions, based on earlier definitions and using terminology introduced in Section 2.4.1 of Chapter 2. Let D and A be c-statements in this section. We recall that the set of all the b-claims of a c-statement D is referred to as the basic content (b-content) of D and is denoted D^+ . The set of the negations of the elements of D^+ is denoted by D^- and the set of b-propositions which occur neither in D^+ nor in D^- will be denoted by $D^?$. Given these terms, we define:

⁴For the remaining combinations, proofs by example can be found in Section B.1 of Appendix B

- A_{rD} , the conjunction of the elements $A^+ \cap D^+$, will be called the *redundant part* of A with regards to D .
- A_{cD} , the conjunction of the elements $A^+ \cap D^-$, will be called the *conflicting part* of A with regards to D .
- A_{xD} , the conjunction of the elements $A^+ \cap D^?$, will be called the *extra part* of A with regards to D .

Next, we establish the following facts about $\text{info}_{TO}()$ which will make things easier.

Theorem 3.2.8. Take a logical space with n atoms and a true atom p . We have:

1. $\text{info}_{TO}(D + p) - \text{info}_{TO}(D) = \frac{1}{2n}$
2. $\text{info}_{TO}(D * p) - \text{info}_{TO}(D) = \frac{1}{n}$
3. $\text{info}_{TO}(D + \neg p) - \text{info}_{TO}(D) = -\frac{1}{2n}$
4. $\text{info}_{TO}(D * \neg p) - \text{info}_{TO}(D) = -\frac{1}{n}$

Proof. See Section B.2 of Appendix B

□

Theorem 3.2.9. Suppose that A is true. Then:

1. $\text{info}(D + A) > \text{info}(D)$
2. $\text{info}(D * A) > \text{info}(D)$
3. $\text{info}(D \div A) < \text{info}(D)$

Proof. 1. The expansion of D with a c-statement $A = \pm p_1 \wedge \pm p_2 \wedge \dots \wedge \pm p_k$ can be seen as the successive expansion of D with each conjunct of A , all of which are true.

- $\text{info}_{TO}()$ - it follows from Theorem 3.2.8 that $\text{info}_{TO}(D + A) > \text{info}_{TO}(D)$.
- $\text{info}_{VA}()$ - the expansion of D with A is going to reduce the number of states corresponding to the database, with each successive addition of an atom halving the set of corresponding states. The eliminated states are each going to have a higher valued corresponding state that remains in the collection of states. For each eliminated state, the state corresponding to it will simply be the state that

is the same on every atom except for the atom of the basic claim being added. In the eliminated state it will be false and in the remaining state it will be true. It follows that when a set of states is so halved, the highest valued state will remain. This ensures that $\text{sum}(D + A) = \text{info}_{VA}(D + A) > \text{sum}(D) = \text{info}_{VA}(D)$, since the total value contributed by each of the x eliminated states is replaced by x more instances of the highest valued member of the set of states being reduced.

2. The revision of D with a c-statement $A = \pm p_1 \wedge \pm p_2 \wedge \dots \wedge \pm p_k$ can be seen as the successive revision or expansion of D with each conjunct of A , all of which are true.

- $\text{info}_{TO}()$ - it follows from Theorem 3.2.8 that $\text{info}_{TO}(D * A) > \text{info}_{TO}(D)$.
- $\text{info}_{VA}()$ - the case where $\pm p$ is added via expansion is as above. For cases where $\pm p$ is added via revision, each original state is replaced by its counterpart where $\pm p$ is true. This simply results in higher valued states, since there is one less false atom. So $\text{sum}(D * A) = \text{info}_{VA}(D * A) > \text{sum}(D) = \text{info}_{VA}(D)$, since each replaced state is replaced by one with a higher value.

3. In contracting a true c-statement from a database, things are a little more complex. Unlike the other two operations, a contraction cannot be broken down into successive contractions of its constituent conjunct, as the following example demonstrates.

Example 3.3

	p	q
w ₁	T	T
w ₂	T	F
w ₃	F	T
w ₄	F	F

- $D = p \wedge q$
- $A = p \wedge q$
- $D \div A = w_1 \vee w_2 \vee w_3$
- $D \div p = w_1 \vee w_3$
- $(D \div p) \div q = w_1 \vee w_2 \vee w_3 \vee w_4$

□

So a different approach is required. If there is a conjunct $\pm p$ of A such that $\pm p \notin \mathbf{D}$, then $D \div \pm p = D$. Otherwise, the resulting set of states will increase. Let m stand for the number of conjuncts in the c-statement that represents $\mathbf{D}$. D corresponds to 2^{n-m}

states and A has k conjuncts, with $k \leq n$. The result of contracting A will be a set of states of size $k2^{n-m} + 2^{n-m}$; for each of the k conjuncts there are 2^{n-m} states on the next closest sphere where it fails.

- $\text{info}_{TO}()$ - since D is a c-statement with m conjuncts, it is satisfied in 2^{n-m} states. $\text{info}_{TO}(D) = 1 - (\frac{d}{2^{n-m}} \times \frac{1}{n}) = 1 - \frac{d}{n2^{n-m}}$, where d is the number of differences relative to the actual state.

After contracting with A , the new measure is $\text{info}_{TO}(D \div A) = 1 - (\frac{d_2}{2^{n-m} + k2^{n-m}} \times \frac{1}{n}) = 1 - (\frac{d_2}{n2^{n-m} + nk2^{n-m}})$.

What is the value of d_2 ? Firstly there is the original d . Then, the $2^{n-m} \times k$ extra states further contribute this value of d k times. Finally, each of these $2^{n-m} \times k$ extra states contribute one more difference to the total number of differences relative to the true state.

So $d_2 = d + kd + k2^{n-m}$ and $\text{info}_{TO}(D \div A) = 1 - (\frac{d+kd+k2^{n-m}}{n2^{n-m} + nk2^{n-m}})$

Now to show that $1 - (\frac{d+kd+k2^{n-m}}{n2^{n-m} + nk2^{n-m}}) < 1 - \frac{d}{n2^{n-m}}$:

$$\begin{aligned} 1 - \left(\frac{d+kd+k2^{n-m}}{n2^{n-m} + nk2^{n-m}} \right) &< 1 - \frac{d}{n2^{n-m}} \\ \frac{d+kd+k2^{n-m}}{n2^{n-m} + nk2^{n-m}} &> \frac{d}{n2^{n-m}} \\ \frac{d+kd+k2^{n-m}}{1+k} &> d \\ d + kd + k2^{n-m} &> d + dk \\ kd + k2^{n-m} &> kd \\ k2^{n-m} &> 0 \end{aligned}$$

- $\text{info}_{VA}()$ - since each of the states being added on is of lower value than the original highest, the overall value of $\text{sum}()$ is lower.

□

Theorem 3.2.10. If more than half of the conjuncts for each of A_{cD} and A_{xD} are true, then:

1. $\text{info}_{TO}(D + A) > \text{info}_{TO}(D)$
2. $\text{info}_{VA}(D + A) > \text{info}_{VA}(D)$ does not universally hold
3. $\text{info}_{TO}(D * A) > \text{info}_{TO}(D)$
4. $\text{info}_{VA}(D * A) > \text{info}_{VA}(D)$ does not universally hold

Proof. 1. In this case, since A is compatible with D , A_{cD} is empty, so we just have to show that if more than half of the conjuncts in A_{xD} are true then the theorem holds. For $\text{info}_{TO}()$, as shown in Theorem 3.2.8, adding a true atom the measure increases by

$\frac{1}{2^n}$ and in adding a false atom the measure decreases by $\frac{1}{2^n}$. Since there are more true conjuncts than false conjuncts, the resulting overall change in measure is positive.

2. Interestingly, the same result does not hold for $\text{info}_{VA}()$. Whilst $\text{info}_{VA}(D + A) > \text{info}_{VA}(D)$ is the norm, $\text{info}_{VA}(D + A) < \text{info}(D)$ is possible. Take a 4-proposition logical space, consisting of propositions $p_1 - p_4$, all true.

- $D = p_1$
- $A = p_2 \wedge p_3 \wedge \neg p_4$
- $D + A = p_1 \wedge p_2 \wedge p_3 \wedge \neg p_4$

$$\text{info}_{VA}(D + A) = 0.75 < \text{info}_{VA}(A) = 0.8125$$

3. In this case, A_{xD} is the same as #1 above. For A_{cD} , according to Theorem 3.2.8 each replacement of a false atom with a true atom results in an increase of $\frac{1}{n}$ (and the corresponding decrease for a false atom). Since there are more true conjuncts than false conjuncts, the resulting overall change in measure is positive.
4. Once again, the same result does not hold for $\text{info}_{VA}()$. Take an 8-proposition logical space, consisting of propositions $p_1 - p_8$, all true.

- $D = \neg p_1 \wedge \neg p_2 \wedge p_3$
- $A = p_1 \wedge p_2 \wedge \neg p_3 \wedge p_4 \wedge p_5 \wedge p_6 \wedge \neg p_7 \wedge \neg p_8$
- A_{cD} is $p_1 \wedge p_2 \wedge \neg p_3$
- A_{xD} is $p_4 \wedge p_5 \wedge p_6 \wedge \neg p_7 \wedge \neg p_8$
- $D * A = p_1 \wedge p_2 \wedge \neg p_3 \wedge p_4 \wedge p_5 \wedge p_6 \wedge \neg p_7 \wedge \neg p_8$

$$\text{info}_{VA}(D * A) = 0.625 < \text{info}_{VA}(D) = 0.711$$

□

Theorem 3.2.11. Suppose that A is completely false: Then:

1. $\text{info}_{TO}(D + A) < \text{info}_{TO}(D)$
2. $\text{info}_{VA}(D + A) < \text{info}_{VA}(D)$ does not universally hold
3. $\text{info}_{TO}(D * A) < \text{info}_{TO}(D)$
4. $\text{info}_{VA}(D * A) < \text{info}_{VA}(D)$ does not universally hold
5. $\text{info}_{TO}(D \div A) > \text{info}_{TO}(D)$

6. $\text{info}_{VA}(D \div A) > \text{info}_{VA}(D)$ does not universally hold

Proof. 1. It simply follows from Theorem 3.2.8 that $\text{info}_{TO}(D + A) < \text{info}_{TO}(D)$.

2. Whilst $\text{info}_{VA}(D + A) < \text{info}(D)$ is the norm, there are some exceptions where $\text{info}_{VA}(D + A) > \text{info}(D)$.

The expansion of D by A is going to reduce the number of states corresponding to the database, with each successive addition of an atom halving the set of corresponding states. The eliminated states are each going to have a lower valued corresponding state that remains in the collection of states. For each eliminated state, the state corresponding to it will simply be the state that is the same on every atom except for the atom of the basic claim being added. In the eliminated state it will be true and in the remaining state it will be false.

It follows that when a set of states is so halved, the highest valued state will be removed and replaced with a state that has a value with one less true atom. Despite this, this second highest valued state will be used in the value aggregate sum($D + A$) calculations more times than the highest valued state was in the sum(D) calculations. This difference can sometimes result in a higher overall value for $D + A$.

For example, take a 9-proposition logical space, consisting of propositions $p_1 - p_9$, all true. Let:

- $D = p_1$
- $A = \neg p_2$
- $D + A = p_1 \wedge \neg p_2$
- $\text{info}_{VA}(D) = 0.77 < \text{info}_{VA}(D + A) = 0.79167$

3. It simply follows from Theorem 3.2.8 that $\text{info}_{TO}(D * A) < \text{info}_{TO}(D)$.

4. Whilst $\text{info}_{VA}(D * A) < \text{info}(D)$ is the norm, following on from #2, given a large enough space sometimes $\text{info}(D * A) > \text{info}(D)$. For example, take a space with 17 propositions, $p_1 - p_{17}$, all true. Let:

- $D = p_1$
- $A = \neg p_1 \wedge \neg p_2$
- $D * A = \neg p_1 \wedge \neg p_2$

$$\text{info}_{VA}(D * A) = 0.77 > \text{info}_{VA}(D) = 0.76$$

This appears to be a surprising and somewhat problematic result; two false atoms are more informative than one true atom. However when the statements involved are

represented in terms of their corresponding models, the result becomes more plausible; the information measure is being made in terms of the whole logical space.

5. As seen in Theorem 3.2.9, contraction with A cannot be broken down into the successive contraction of conjuncts. Continuing on from Theorem 3.2.9, after contracting with A , the new measure is $\text{info}_{TO}(D \div A) = 1 - (\frac{d_2}{2^{n-m} + k2^{n-m}} \times \frac{1}{n}) = 1 - (\frac{d_2}{n2^{n-m} + nk2^{n-m}})$.

What is the value of d_2 ? Firstly there is the original d . Then, converse to Theorem 3.2.9 #3, the $2^{n-m} \times k$ extra states contribute $kd - (2^{n-m} \times k)$ differences. So $d_2 = d + (kd - k2^{n-m})$ and $\text{info}_{TO}(D \div A) = 1 - (\frac{d+kd-k2^{n-m}}{n2^{n-m} + nk2^{n-m}})$.

Now to show that $1 - (\frac{d+kd-k2^{n-m}}{n2^{n-m} + nk2^{n-m}}) > 1 - \frac{d}{n2^{n-m}}$:

$$\begin{aligned} 1 - \left(\frac{d+kd-k2^{n-m}}{n2^{n-m} + nk2^{n-m}} \right) &> 1 - \frac{d}{n2^{n-m}} \\ \frac{d+kd-k2^{n-m}}{n2^{n-m} + nk2^{n-m}} &< \frac{d}{n2^{n-m}} \\ \frac{d+kd-k2^{n-m}}{1+k} &< d \\ d + kd - k2^{n-m} &< d + dk \\ kd - k2^{n-m} &< kd \\ 0 &< k2^{n-m} \end{aligned}$$

6. Whilst $\text{info}_{VA}(D \div A) > \text{info}(D)$ is the norm, $\text{info}_{VA}(D \div A) < \text{info}(D)$ is possible. Just take examples corresponding to #2 and #4 above, such as $n = 9$, $D = p_1 \wedge \neg p_2$ and $A = \neg p_2$, so that $D \div A = p_1$.

□

Remark Like $\text{info}_{VA}()$, Niiniluoto's min-sum measure for truthlikeness (Section 2.4.4 of Chapter 2) fails Theorems 3.2.10 and 3.2.11. Whether there is some feature common to both frameworks that can explicate these peculiarities remains to be investigated.

In the last round of theorems, D and A were c-statements. Interestingly, some of the positive results above do not hold when D is extended to the class of all statements in a given logical space. Firstly, the counterpart to Theorem 3.2.9.

Theorem 3.2.12. With D being any statement and A being a true c-statement, $\text{info}(D + A) > \text{info}(D)$ does not universally hold.

Proof. For example, take the following:

- $D = w_4 \vee w_5$
- $A = h$

- $D + A = w_4$
- $\text{info}_{VA}(w_4) = 0.33 < \text{info}_{VA}(w_4 \vee w_5) = 0.625$
- $\text{info}_{TO}(w_4) = 0.33 < \text{info}_{TO}(w_4 \vee w_5) = 0.5$

□

This carries over to revision.

Theorem 3.2.13. With D being any statement and A being a true c-statement, $\text{info}(D * A) > \text{info}(D)$ does not universally hold.

Proof. Take a 6-proposition space with true state description $p_1 \wedge p_2 \wedge p_3 \wedge p_4 \wedge p_5 \wedge p_6$. Let

- $D = \neg p_1 \wedge ((p_2 \wedge \neg p_3 \wedge \neg p_4 \wedge \neg p_5 \wedge \neg p_6) \vee (\neg p_2 \wedge p_3 \wedge p_4 \wedge p_5 \wedge p_6))$.
- $A = p_1 \wedge p_2$
- $D * A = p_1 \wedge p_2 \wedge \neg p_3 \wedge \neg p_4 \wedge \neg p_5 \wedge \neg p_6$.
- $\text{info}_{VA}(D * A) = 0.3333 < \text{info}_{VA}(D) = 0.6589$
- $\text{info}_{TO}(D * A) = 0.3333 < \text{info}_{TO}(D) = 0.4167$

□

The corresponding result still holds for contraction however.

Theorem 3.2.14. With D being any statement and A being a true c-statement, $\text{info}(D \div A) < \text{info}(D)$ does universally hold.

Proof. Firstly, we have:

- D corresponds to $|W_D|$ states ($W_D = \{w_1, w_2, \dots, w_x\}$), which are in the centre sphere.
- These states make up a distance of d atomic element differences to the actual state.
- Let A contain k literals.

Considering cases where $A \in \text{Cn}(\mathbf{D})$, since A is a true c-statement, this means that the states in W_D all agree on the true valuation for all of the k atomic elements corresponding to A .

The next closest sphere from the centre sphere will contain states that differ by one atomic element evaluation from some state in the centre sphere. The number of states at which A fails in this sphere is $k \times |W_D|$; for each of the $|W_D|$ centre states and each of the k atomic elements corresponding to a conjunct of A , flip the element's truth value to get such a state.

- To show that this result holds for $\text{info}_{TO}()$, we will use the equation $\text{info}_{TO}(A) = \frac{\sum_{i=1}^n \text{Pr}^*(p_i, A)}{n}$, as given in Section 2.4.7 of Chapter 2.

For each p_i that occurs as a conjunct in A , $\text{Pr}^*(p_i, D \div A) < \text{Pr}^*(p_i, D)$, since it starts off as 1 and some states are then introduced in which the valuation for p_i differs to the actual true state. From here it will suffice to show that for each p_i that does not occur in A , $\text{Pr}^*(p_i, D \div A)$ is not greater than $\text{Pr}^*(p_i, D)$. In fact, they remain the same:

- Let t stand for the number of true atomic element valuations within W_D corresponding to all p_i not in A .
- The number of states resulting from the contraction is $|W_D| + (k \times |W_D|)$.
- The number of true atomic element valuations within the resulting set of states corresponding to all p_i not in A is $t + kt$. Given this, the following shows that the

$$\begin{aligned} \text{proportion remains equal: } \frac{\frac{t}{|W_D|}}{t} &= \frac{\frac{t+kt}{|W_D|+k|W_D|}}{\frac{t+kt}{|W_D|(1+k)}} \\ &= \frac{t}{t+kt} = \frac{t+kt}{t+kt} \end{aligned}$$

- The case for info_{VA} is simple: each of the added states resulting from $\mathbf{D} \div A$ is going to have a value lower than its corresponding original state, so the aggregate value is lower.

□

Results also fail when D is any statement and A is a c-statement that is completely false. See Section B.2 of Appendix B.

We now draw to the end of our investigation into results concerning the combination of truthlikeness information measurement and belief revision. Beyond these results, finding

ways to explicate increases and decreases of information yield and exploring possible connections between such behaviour and other properties are matters for further research down the

line. As a sample, we once again recall the function $\text{info}_{TO}(A) = \frac{\sum_{i=1}^n \text{Pr}^*(p_i, A)}{n}$, as given in Section 2.4.7 of Chapter 2. Using this interpretation we can get the following descriptions of increase and decrease:

- $\text{info}_{TO}(D + A) > \text{info}_{TO}(D)$ iff the total increase over all $\text{Pr}^*(p_i, D + A)$ is greater than the total decrease over all $\text{Pr}^*(p_i, D + A)$
- $\text{info}_{TO}(D * A) > \text{info}_{TO}(D)$ iff the total increase over all $\text{Pr}^*(p_i, D * A)$ is greater than the total decrease over all $\text{Pr}^*(p_i, D * A)$
- $\text{info}_{TO}(D \div A) < \text{info}_{TO}(D)$ iff the total increase over all $\text{Pr}^*(p_i, D \div A)$ is less than the total decrease over all $\text{Pr}^*(p_i, D \div A)$

These descriptions provide a way to explicate results above. For example, with Theorem 3.2.12 we can see that whilst the addition of true input A increased $\text{Pr}^*(\cdot)$ for some p_i the corresponding decrease for some p_i was greater:

- $D = w_4 \vee w_5$
- $A = h$
- $D + A = w_4$
- $\text{Pr}^*(h|D) = \text{Pr}^*(r|D) = \text{Pr}^*(w|D) = \frac{1}{2}$
- $\text{Pr}^*(h|D + A) = 1 > \text{Pr}^*(h|D)$
- $\text{Pr}^*(r|D + A) = \text{Pr}^*(w|D + A) = 0 < \text{Pr}^*(r|D) = \text{Pr}^*(w|D)$

3.3 Agent-Relative Informativeness

We now come to a specification of agent-relative informativeness. There are three types of scenarios that we are interested in, which we shall sequentially look at in order to develop the key points of this section:

1. The agent has only true database content and receives only true input.

2. The agent has true or false database content and receives only true input.
3. The agent has true or false database content and receives true or false input.

3.3.1 Adding Information to Information

In the first and simplest type of scenario, the database D is true and the input A is true, so the operation of expansion is performed in adding A to D . With this simplest type of scenario we can begin to introduce some definitions.

Definition The *gain* of adding some input A to a database D is the difference between the information measure of the result and the information measure of the original database:

$$\text{gain}(A, D) =_{df} \text{info}(D + A) - \text{info}(D)$$

This definition for information gain can be seen as a version of a general formula for conditional information used in inverse probabilistic approaches (See Section 2.1 of Chapter 2):

$$\text{inf}(A|B) =_{df} \text{inf}(A \wedge B) - \text{inf}(B)$$

The informativeness of a piece of input is judged in terms of the gain resulting from its addition. For example, if an agent's database $D = h \wedge r$, then even though $\text{info}(h \wedge r) > \text{info}(w)$, w is more informative than $h \wedge r$ for the agent, since $\text{gain}(w, D) > \text{gain}(h \wedge r, D) = 0$.

Remark The fact that $(\forall D)\text{gain}(\top, D) = 0$ captures the non-informativeness of tautologies.

Given this definition of information gain, there is another definition of interest based on the gain and the information measure of the input:

Definition The *efficiency* of adding A to D is defined as:

$$\text{efficiency}(A, D) =_{df} \frac{\text{gain}(A, D)}{\text{info}(A)}$$

Put in economic terms, *information efficiency* can be seen as the ‘value for money’ of a piece of information, the resulting gain to investment ratio.

Example 3.4

$D = h \vee \neg r = w_1 \vee w_2 \vee w_3 \vee w_4 \vee w_7 \vee w_8$, so:

1. $\text{info}_{VA}(D) = 0.625$
2. $\text{info}_{TO}(D) = 0.5$

Tables 3.2 and 3.3 list some results for different A :

A	$\text{info}_{TO}(A)$	$D + A$	$\text{info}_{TO}(D + A)$	$\text{gain}(A, D)$	$\text{efficiency}(A, D)$
r	0.67	$h \wedge r$	0.83	0.33	0.493
w	0.67	$(h \vee \neg r) \wedge w$	0.75	0.25	0.373
h	0.67	h	0.67	0.17	0.2537
$r \vee w$	0.61	$(h \vee \neg r) \wedge (r \vee w)$	0.6667	0.1667	0.273
$\neg(w_4 \vee w_7)$	0.5556	$w_1 \vee w_2 \vee w_3 \vee w_8$	0.5833	0.0833	0.15
$\neg(w_3 \vee w_4 \vee w_7)$	0.5333	$w_1 \vee w_2 \vee w_8$	0.5556	0.0556	0.104
$\neg(w_2)$	0.4762	$w_1 \vee w_3 \vee w_4 \vee w_7 \vee w_8$	0.4667	-0.0333	-0.07

Table 3.2: Results for $\text{info}_{TO}()$

A	$\text{info}_{VA}(A)$	$D + A$	$\text{info}_{VA}(D + A)$	$\text{gain}(A, D)$	$\text{efficiency}(A, D)$
r	0.8333	$h \wedge r$	0.9583	0.3333	0.4
w	0.8333	$(h \vee \neg r) \wedge w$	0.875	0.25	0.3
h	0.8333	h	0.8333	0.2083	0.25
$r \vee w$	0.7083	$(h \vee \neg r) \wedge (r \vee w)$	0.8333	0.2083	0.29
$\neg(w_3 \vee w_4 \vee w_7)$	0.7083	$w_1 \vee w_2 \vee w_8$	0.8333	0.2083	0.29
$\neg(w_4 \vee w_7)$	0.6667	$w_1 \vee w_2 \vee w_3 \vee w_8$	0.7917	0.1667	0.25
$\neg(w_2)$	0.5417	$w_1 \vee w_3 \vee w_4 \vee w_7 \vee w_8$	0.6667	0.0417	0.077

Table 3.3: Results for $\text{info}_{VA}()$

□

As can be seen, the gain/efficiency ranking orderings for $\text{info}_{VA}()$ and $\text{info}_{TO}()$ differ. h , r and w all have the same information measure, though given the original contents of the database, r results in the largest gain whilst h results in the smallest gain; the information already contained in the database is most efficiently coupled with the input of r for the greatest return.

As will be covered shortly, the introduction of misinformation into revision scenarios results in general occurrences of negative gains and hence negative efficiency measures. Yet whilst negative gains only become a possibility for $\text{info}_{VA}()$ once misinformation is introduced, as can be seen from Table 3.2 the failure of $\text{info}_{TO}()$ to covary with logical strength

amongst true statements means that in some cases the gain and efficiency measures will be negative even when the original database contains only true content and the input is true.

3.3.2 Adding Information to Information/Misinformation

In the second type of scenario, the received input is true but the database can contain misinformation as well as information. Thus the operation of revision is now brought into consideration and the gain measure is expanded to include revision:

$$\text{gain}(A, D) =_{df} \text{info}(D * A) - \text{info}(D)$$

Whilst AGM is the dominant framework for belief revision, it is worth mentioning that a number of proposals for revision have been developed independently of the AGM framework, particularly within computer science.⁵ Also, the revision construction we have employed is but one of many that adheres to the AGM framework. By using Grove's system of spheres one could get an alternative construction by using a different distance function to measure the distance of states from the centre sphere. For comparative purposes we will now outline such an alternative method in order to emphasise that as well as the information measure used and what content the agent already has, the informativeness of some revision-accepted input really does depend on the revision method employed by the agent.

In this alternative construction the distance $\text{dis}_s(\mathbf{K})$ of a state s from $\mathbf{K}$ is defined as:

$$\text{dis}_s(\mathbf{K}) = \begin{cases} 0 & \text{if } s \in \mathbf{M}_K \\ \text{sum}(\{\Delta(s, w) | w \in \mathbf{M}_K\}) & \text{otherwise} \end{cases}$$

With this distance function, even if an outside state is very close to one state in the centre sphere, it can still be placed on a relatively distant sphere, since what counts is its total distance from all centre states. So if w_1 , w_7 and w_8 constitute the centre sphere, although by the $\text{min}()$ distance function w_2 is in the first closest sphere given its distance of 1 to w_1 , its distance using the $\text{sum}()$ distance function is 6. w_3 on the other hand has the same $\text{min}()$ distance as w_2 but is closer than w_2 in terms of $\text{sum}()$ with a distance of 4.

⁵See [177] for a good sample overview.

Example 3.5

$$D = w_2 \vee w_8$$

Here are the sphere orderings based on the sum distance:

distance	states
0	w_2, w_8
2	w_4, w_6
4	w_1, w_3, w_5, w_7

Using this distance ordering, $D * w_1 \vee w_5 = w_1 \vee w_5$

Here are the sphere orderings based on the min distance:

distance	states
0	w_2, w_8
1	w_1, w_4, w_6, w_7
2	w_3, w_5

Using this distance ordering, $D * w_1 \vee w_5 = w_1$

Thus is this case since $\text{info}(w_1) > \text{info}(w_1 \vee w_5)$ the min distance function gives a better result.

□

As the following example shows, it is possible for two different revision methods $*_1$ and $*_2$ to have $\text{info}(D *_1 A_1) > \text{info}(D *_1 A_2)$ and $\text{info}(D *_2 A_2) > \text{info}(D *_2 A_1)$.

Example 3.6

- Take a 5-proposition space with the true state being $w_1 = p_1 \wedge p_2 \wedge p_3 \wedge p_4 \wedge p_5$.
- $w_2 = p_1 \wedge p_2 \wedge p_3 \wedge p_4 \wedge \neg p_5$
- w_{29}, w_{15}, w_8 and w_{12} each have two true atoms and three false atoms:

$$- w_{29} = \neg p_1 \wedge \neg p_2 \wedge \neg p_3 \wedge p_4 \wedge p_5$$

- $w_{15} = p_1 \wedge \neg p_2 \wedge \neg p_3 \wedge \neg p_4 \wedge p_5$
- $w_8 = p_1 \wedge p_2 \wedge \neg p_3 \wedge \neg p_4 \wedge \neg p_5$
- $w_{12} = p_1 \wedge \neg p_2 \wedge p_3 \wedge \neg p_4 \wedge \neg p_5$
- $w_{32} = \neg p_1 \wedge \neg p_2 \wedge \neg p_3 \wedge \neg p_4 \wedge \neg p_5$
- $w_{13} = p_1 \wedge \neg p_2 \wedge \neg p_3 \wedge p_4 \wedge p_5$
- $D = w_2 \vee w_{29} \vee w_{15} \vee w_8 \vee w_{12}$
- $A_1 = w_1 \vee w_{32}$
- $A_2 = w_1 \vee w_{13}$

Given these details, for the min distance function:

- $\text{dis}_{w_1}(\mathbf{D}) = 1$
- $\text{dis}_{w_{32}}(\mathbf{D}) = 2$
- $\text{dis}_{w_{13}}(\mathbf{D}) = 1$

Given these distances:

- $D * A_1 = w_1$
- $D * A_2 = w_1 \vee w_{13}$

Since $\text{info}(w_1) > \text{info}(w_1 \vee w_{13})$, A_1 is more informative.

For the sum distance function:

- $\text{dis}_{w_1}(\mathbf{D}) = 13$
- $\text{dis}_{w_{32}}(\mathbf{D}) = 12$
- $\text{dis}_{w_{13}}(\mathbf{D}) = 11$

Given these distances:

- $D * A_1 = w_{32}$

- $D * A_2 = w_{13}$

Since $\text{info}(w_{13}) > \text{info}(w_{32})$, A_2 is more informative.

□

Thus, even if two agents a_1 and a_2 have exactly the same data D and the same $\text{info}()$ measure is used, it is possible for A_1 to be more informative than A_2 relative to a_1 and for A_2 to be more informative than A_1 relative to a_2 .

3.3.3 Adding Misinformation/Information to Misinformation/Information

In the third type of scenario, both the received input and the database can contain misinformation or information. So not only is there now the possibility that the input will conflict with the contents of the database, but this can now happen because false input conflicts with true database content.

Example 3.7

$$D = w_1 \vee w_7$$

Here are the sphere orderings based on the sum distance:

distance	states
0	w_1, w_7
2	w_3, w_5
4	w_2, w_4, w_6, w_8

Using this distance ordering, $D * w_4 \vee w_8 = w_4 \vee w_8$

Here are the sphere orderings based on the min distance:

distance	states
0	w_1, w_7
1	w_2, w_3, w_5, w_8
2	w_4, w_6

Using this distance ordering, $D * w_4 \vee w_8 = w_8$

Thus contrary to Example 3.5, in this case since $\text{info}(w_4 \vee w_8) > \text{info}(w_8)$ the sum distance function gives a better result.

□

Obviously the introduction of misinformative input significantly increases the set of possible negative efficiency ratios. It also means that in some instances $\text{efficiency}(D, A)$ is going to be undefined; namely when A is the state description that is completely false and $\text{info}(A) = 0$. The following example gives some results for negative ratios due to misinformative input.

Example 3.8

$D = h \vee \neg r$, so:

1. $\text{info}_{VA}(D) = 0.625$
2. $\text{info}_{TO}(D) = 0.5$

Here are some results for different inputs A :

A	$\text{info}_{TO}(A)$	$D + A$	$\text{info}_{TO}(D + A)$	$\text{gain}(D, A)$	$\text{efficiency}(D, A)$
$\neg h$	0.3333	$\neg h \wedge \neg r$	0.1667	-0.3333	-1
$\neg r$	0.3333	$\neg r$	0.3333	-0.1667	-0.5
$\neg w$	0.3333	$(h \vee \neg r) \wedge \neg w$	0.3333	-0.1667	-0.5
$\neg h \vee \neg w$	0.3889	$\neg r \vee \neg w$	0.3889	-0.1111	-0.2857

Table 3.4: Results for $\text{info}_{TO}()$

A	$\text{info}_{VA}(A)$	$D + A$	$\text{info}_{VA}(D + A)$	$\text{gain}(D, A)$	$\text{efficiency}(D, A)$
$\neg h$	0.5	$\neg h \wedge \neg r$	0.2917	-0.3333	-0.6666
$\neg r$	0.5	$\neg r$	0.5	-0.125	-0.25
$\neg w$	0.5	$(h \vee \neg r) \wedge \neg w$	0.5417	-0.0833	-0.1666
$\neg h \vee \neg w$	0.4583	$\neg r \vee \neg w$	0.4583	-0.1667	-0.3637

Table 3.5: Results for $\text{info}_{VA}()$

□

Apart from the foundational aspects of agent-relative informativeness set forth thus far in this section, the possibility of false input raises new matters to investigate. If a receiving agent cannot be sure that the input they are given is true, say because it is coming from an unreliable source, how might they determine whether to accept or reject the input? Or if receiving multiple mutually inconsistent inputs from multiple sources, how can the agent deal with these conflicting inputs and make a decision?

3.4 Dealing with Uncertainty and Conflicting Input

Relevant to these questions is an extension of the AGM framework called *non-prioritised* belief revision. In the classic AGM model of belief revision, the input data is always accepted by the agent and used to expand or revise accordingly. With non-prioritised belief revision, an agent weighs new input against the data already held in its database and despite the input's novelty, it is given no special priority. "A sentence A that contradicts previous beliefs is accepted only if it has more epistemic value than the original beliefs that contradict it. In that case, enough of the previous sentences are deleted to make the resulting set consistent. Otherwise, the input is itself rejected" [96]. Such a modified revision operator is called a *semi-revision operator* [94].

Screened Revision

One general way to perform semi-revision is as follows:

- Decide whether the input A should be accepted or rejected.
- If A was accepted, revise by A .

David Markinson's *screened revision* [131] is an approach that uses this process. It involves a set of potential core data $\mathbf{C}$ that is immune to revision. The database $\mathbf{D}$ is revised by the input A if A is consistent with the set $\mathbf{C} \cap \mathbf{D}$ of core data held by the agent; so with such a revision the elements of $\mathbf{C} \cap \mathbf{D}$ must remain.

The basic form of a screened revision operation $\#$ is as follows:

$$\begin{aligned} \mathbf{D} \# A &= \mathbf{D} * A \text{ if } A \text{ is consistent with } \mathbf{C} \cap \mathbf{D} \\ &= \mathbf{D} \text{ otherwise} \end{aligned}$$

Yet as demonstrated in the following example, taking $*$ to be the standard AGM revision operator will not do, for it can be the case that whilst the input A is consistent with the core $\mathbf{C} \cap \mathbf{D}$, its addition will eliminate some of the core.

Example 3.9

- $D = q$
- $C = p \vee q$
- $p \vee q \in \mathbf{D}$
- $A = \neg q$, which is consistent with $\mathbf{C} \cap \mathbf{D}$

Using our revision system:

- $D * \neg q = \neg q$
- $p \vee q \notin \mathbf{D} * \neg q$

□

Makinson deals with this issue by first defining a special *protecting contraction*, and defines a *protecting revision* in terms of the protecting contraction using the Levi identity.

A *contraction protecting* $\mathbf{C}$, denoted $\div_C$, is defined as $\mathbf{D} \div_C A = \cap \gamma(\mathbf{D} \perp_C A)$, where $(\mathbf{D} \perp_C A)$ is the set of all maximal subsets of $\mathbf{D}$ that do not imply A but do include $\mathbf{C} \cap \mathbf{D}$, where γ is a selection function, as in standard AGM.

Screened revision then becomes

$$\mathbf{D} *_C A = \text{Cn}((\mathbf{D} \div_C \neg A) \cup \{A\})$$

An operation $\#$ on a database $\mathbf{D}$ is a *screened revision* if and only if:

$$\begin{aligned} \mathbf{D} \# A &= \mathbf{D} *_C A \text{ if } A \text{ is consistent with } \mathbf{C} \cap \mathbf{D} \\ &= \mathbf{D} \text{ otherwise} \end{aligned}$$

Using this, Example 3.9 becomes:

Example 3.10

	p	q
w ₁	T	T
w ₂	T	F
w ₃	F	T
w ₄	F	F

Here are the sphere orderings based on $\Delta_{\min}$:

sphere level	Δ	states
0	0	w ₁ , w ₃
1	$\frac{1}{2}$	w ₂ , w ₄

q fails in w_2 and w_4 , so the result of standard revision using our system would be $\mathbf{D} * A = \{w_1, w_2, w_3, w_4\} \cap \{w_2, w_4\}$, which is $\neg q$. Yet out of w_2 and w_4 , only w_2 satisfies $p \vee q$, so $\mathbf{D} * A = \{w_1, w_2, w_3\} \cap \{w_2, w_4\}$, which is $p \wedge \neg q$.

□

C can be thought of as content that the system takes to be knowledge and it is helpful and appropriate to stipulate that $C \in \mathbf{D}$; all protected content that is classed as knowledge by the system is held by the system.

Combining Screened Revision with Information Estimation

Agents could use a combination of screened revision and information estimation to help them decide whether or not to accept some input and make optimal input selections. This involves using $\mathbf{C}$ as a basis from which to make estimated information calculations and accept an input A if (1) it does not conflict with $\mathbf{C} \cap \mathbf{D}$ and if (2) the estimated informativeness of $\mathbf{D} \# A$ is greater than the estimated informativeness of $\mathbf{D}$. Formally stated, the supplementary condition is:

$$A \in \mathbf{D} \# A \Rightarrow \text{info}_{est}(D \# A | C) > \text{info}_{est}(D | C)$$

Here is an example of this idea:

Example 3.11

Take the following:

- $D = h \wedge (\neg r \vee \neg w)$
- $C = h \vee r \vee w$
- $\text{info}_{TO-est}(D|C) = 0.51$

Consider the inputs $A_1 = \neg h$ and $A_2 = r \wedge w$. Should the agent accept A_1 ? Should the agent accept A_2 ?

- A_1 is compatible with C and $D \# A_1 = \neg h \wedge ((r \wedge \neg w) \vee (\neg r \wedge w))$. So $\text{info}_{TO-est}(D \# A_1|C) = 0.48$.
- A_2 is compatible with C and $D \# A_2 = h \wedge r \wedge w$. So $\text{info}_{TO-est}(D \# A_2|C) = 0.57$.

Thus according to these calculations, the agent should accept A_2 but reject A_1 . Furthermore, this method gives an easy way to compare two inputs. If an agent could select only one of many inputs, then they should choose that input which results in the greatest estimated increase.

□

Accommodative Revision and Resolving Conflicting Input

With screened revision, if A is not consistent with C , then A is rejected. Accommodative revision [63], of which screened revision can be seen as a special case ⁶, is, as the name suggests, a more accommodating form of non-prioritised belief revision. Rather than simply rejecting or accepting the input, the input is first revised with C and the database is then revised with this result. The accommodative revision operator $\otimes$ is defined as:

$$D \otimes A =_{def} D * (A * C)$$

⁶By having $A * C = C$ whenever $\neg A \in \mathbf{C}$, we get screened revision.

The operation of accommodative revision ensures that if an agent has some protected data, then it will remain, whilst still allowing the possibility of extracting other pieces of information from false input which only partially conflicts with the core data.

Remark Alternatively, one could define a similar operator $\otimes_2$ as:

$$D \otimes_2 A =_{def} (D * A) * C$$

which would still ensure that C is an element of the result.

Since $*$ is not associative, these two operators are different, as demonstrated by the following example:

Example 3.12

- $D = w_1$
- $A = w_2 \vee w_5 \vee w_7$
- $C = w_1 \vee w_4 \vee w_8$
- $D * (A * C) = w_1$
- $(D * A) * C = w_1 \vee w_4$

□

3.4.1 Paraconsistent Approaches

We now come to a look at how an inconsistency tolerant paraconsistent framework can be used. This will involve an extension of our system that incorporates paraconsistent models. Once again, the underlying logic will be that of the paraconsistent *Logic of Paradox* (*LP*) [149]. Restall and Slaney [157] pave the way for this type of approach, though the underlying logic they consider is *FDE*, which is both paracomplete and paraconsistent.

Within this new paraconsistent framework, the distance function (Δ) is that introduced in Section 2.4.2 of Chapter 2 and is such that:

- n is the number of propositional variables in the logical space
- add $\frac{1}{n}$ to the total distance for each atomic difference between a T and an F
- add $\frac{1}{2n}$ to the total distance for each atomic difference between a B and a T or F

Division by n is simply a normalisation factor used for truthlikeness information calculations. At times thinking will be facilitated by removing this element and using the distances 1 and $\frac{1}{2}$.

To see this approach at work we again consider the weather framework with the 27 possible states listed in Table 3.6.

State	h	r	w	State	h	r	w	State	h	r	w
w ₁	T	T	T	w ₁₀	B	T	T	w ₁₉	F	T	T
w ₂	T	T	B	w ₁₁	B	T	B	w ₂₀	F	T	B
w ₃	T	T	F	w ₁₂	B	T	F	w ₂₁	F	T	F
w ₄	T	B	T	w ₁₃	B	B	T	w ₂₂	F	B	T
w ₅	T	B	B	w ₁₄	B	B	B	w ₂₃	F	B	B
w ₆	T	B	F	w ₁₅	B	B	F	w ₂₄	F	B	F
w ₇	T	F	T	w ₁₆	B	F	T	w ₂₅	F	F	T
w ₈	T	F	B	w ₁₇	B	F	B	w ₂₆	F	F	B
w ₉	T	F	F	w ₁₈	B	F	F	w ₂₇	F	F	F

Table 3.6: *LP* Truth Table for 3-Proposition Logical Space

As examples, using this distance function $\Delta(w_1, w_2) = \frac{1}{6}$ and $\Delta(w_5, w_{24}) = \frac{1}{2}$.

Theorem 3.4.1. Like its classical counterpart, this distance function satisfies the metric conditions:

1. $\Delta(w_i, w_j) \geq 0$
2. $\Delta(w_i, w_j) = 0$ iff $w_i = w_j$
3. $\Delta(w_i, w_j) = \Delta(w_j, w_i)$
4. $\Delta(w_i, w_k) \leq \Delta(w_i, w_j) + \Delta(w_j, w_k)$

Proof. See Appendix B, Section B.3

□

Paraconsistent Expansion, Contraction, Revision and Consolidation

The belief revision operations behave differently in this context. To begin with, the expansion $\mathbf{D} + A$ can now be performed even when $\neg A \in \mathbf{D}$.

To have a useful contraction, its definition must be modified relative to the classical definition. We saw in Section 3.1 that with classical AGM, contraction can be defined in terms of the Harper identity. We also used the following equation to define contraction:

$$D \div A = D \vee (D * \neg A)$$

In the classical context, performing the contraction $D \div A$ resulted in a set of states that was the union of the original set of states corresponding to D and the set of closest states to D where $\neg A$ was satisfied. However, since in this paraconsistent setting A and $\neg A$ can be satisfied in the same state, this standard definition of contraction can no longer be used. The modified definition of contraction must be:

$$\mathbf{D} \div A = \mathbf{D} \cup \mathbf{X}$$

where $\mathbf{X}$ is the set of closest states at which A fails.

Another operation that is an extension of the three core AGM operations is *consolidation*, which basically involves making an inconsistent database consistent by removing enough conflicting elements. This operation is denoted $D!$. For classical belief bases one way to perform consolidation is to contract falsum (contradiction), so that $D! = D \div \perp$. It does not work for classical inconsistent belief sets however because there is only one inconsistent belief set, that in which the principle of explosion makes everything a consequence. So once a belief set is inconsistent, all distinctions and structure are lost and cannot be restored with consolidation.

Since we are employing a paraconsistent model system here, where we are able to retain structure when the set is inconsistent, we can fashion a method for consolidation whilst still using belief sets. Consolidation here will involve the incremental contraction of elements to restore consistency, as shown with the following example:

Example 3.13

Take the database $\mathbf{D} = \text{Cn}\{h, \neg h, r, w\}$. This corresponds to the states $\{w_{10}, w_{11}, w_{13}, w_{14}\}$ of Table 3.6, all paraconsistent states. One option to restore consistency is to contract with h : $\mathbf{D} \div h$. We start off by working out the sphere ordering based on Δ :

sphere level	Δ	states
0	0	$w_{10}, w_{11}, w_{13}, w_{14}$
1	$\frac{1}{6}$	$w_1, w_2, w_4, w_5, w_{12}, w_{15}, w_{16}, w_{17}, w_{19}, w_{20}, w_{22}, w_{23}$
2	$\frac{2}{6}$	$w_3, w_6, w_7, w_8, w_{18}, w_{21}, w_{24}, w_{25}, w_{26}$
3	$\frac{3}{6}$	w_9, w_{27}

The closest set of states at which h does not hold can be found in sphere 1. This set is $\{w_{19}, w_{20}, w_{22}, w_{23}\}$. The resulting set of states is $\mathbf{D} \cup \mathbf{X} = \{w_{10}, w_{11}, w_{13}, w_{14}\} \cup \{w_{19}, w_{20}, w_{22}, w_{23}\}$.

A practice that can be implemented is to convert the resulting set of states to its equivalent classical set of states. The only classical state amongst $\mathbf{D} \cup \mathbf{X}$ is $\{w_{19}\}$, which is the same as state w_5 of Table 3.1 and as expected corresponds to the statement $\neg h \wedge r \wedge w$.

□

Given these expansion and contraction operations, revision can then be defined using the Levi identity. In fact, the sequence of operations involved in the Levi identity can now be reversed.

Theorem 3.4.2. Given this construction $(D \div A) + \neg A$ is equivalent to $(D + \neg A) \div A$

Proof. See Appendix B, Section B.3

□

This result gives us the *Reverse* Levi Identity:

$$D * A = (D + A) \div \neg A$$

Remark Unlike our construction, the operations $(D \div A) + \neg A$ and $(D + \neg A) \div A$ for belief

bases have been shown to differ in their formal properties [95, p. 46]. I suspect that there are also constructions based on belief sets for which this equivalence fails.

Basic Contradiction Contraction and Weightings

As was seen in Section 3.1, with standard AGM contraction, if p is less entrenched than q ($p < q$) and $p \wedge q \in \mathbf{D}$, then $q \in \mathbf{D} \div p \wedge q$ and $p \notin \mathbf{D} \div p \wedge q$. However, if p is just as entrenched as q , then $q \notin \mathbf{D} \div p \wedge q$, $p \notin \mathbf{D} \div p \wedge q$ and $(p \wedge \neg q) \vee (\neg p \wedge q) \in \mathbf{D} \div p \wedge q$.

In this paraconsistent setting, $\mathbf{D} \div p \wedge \neg p$ results in $p \notin \mathbf{D} \div p \wedge \neg p$, $\neg p \notin \mathbf{D} \div p \wedge \neg p$ and only $p \vee \neg p \in \mathbf{D} \div p \wedge \neg p$ [$(p \wedge \neg \neg p) \vee (\neg p \wedge \neg p) \equiv p \vee \neg p$].

As an aside, it is worth noting that an entrenchment which reflected a priority of p or $\neg p$ would determine that one remains after such a contraction. Such an entrenchment could be based on assignments of probability or credence held by the agent. If an agent has more reason to believe than not belief p , or holds p to be more probable than not, then $\neg p < p$.

In Example 3.2 different weights were assigned to different atomic elements. For present purposes, the required differences can be implemented by making the distance between B and T different to the distance between B and F.

Example 3.14

This example shows how probability distributions can be factored into distance calculations. Ordinarily, $\text{distance}(\mathbf{B}, \mathbf{T}) = \text{distance}(\mathbf{B}, \mathbf{F}) = \frac{1}{2n}$. Let $\Pr(p) = x$ and $\Pr(\neg p) = 1 - x$. Then $\text{distance}(\mathbf{B}, \mathbf{T})$ now becomes $\frac{1}{n} \times (1 - x)$ and $\text{distance}(\mathbf{B}, \mathbf{F})$ now becomes $\frac{1}{n} \times x$.

In this example $n = 2$, $\Pr(p) = \frac{3}{4}$ and $\Pr(\neg p) = \frac{1}{4}$.

	p	q
w ₁	T	T
w ₂	T	B
w ₃	T	F
w ₄	B	T
w ₅	B	B
w ₆	B	F
w ₇	F	T
w ₈	F	B
w ₉	F	F

So now $\text{distance}(B, T) = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8}$ and $\text{distance}(B, F) = \frac{1}{2} \times \frac{3}{4} = \frac{3}{8}$.

$D = p \wedge \neg p \wedge q$. Sphere orderings based on the new Δ are:

Δ	set of states
0	w ₄ , w ₅
$\frac{1}{8}$	w ₁ , w ₂
$\frac{3}{8}$	w ₆ , w ₇ , w ₈
$\frac{1}{2}$	w ₃
$\frac{3}{4}$	w ₉

Since w₁ and w₂ are the closest states in which $p \wedge \neg p$ is not satisfied, $D \div p \wedge \neg p = p \wedge q$, rather than q .

□

Compound Contradictions

The contraction of a contradiction $A \wedge \neg A$, where A is not an atom, can leave more than just a tautology, as demonstrated in the following example.

Example 3.15

The database formula $D = (h \vee r \vee w) \wedge \neg(h \vee r \vee w)$ will serve as a good test case. This statement holds in the following states: w₁₄, w₁₅, w₁₇, w₁₈, w₂₃, w₂₄, w₂₆.

The sphere rankings are:

Δ	states
0	$w_{14}, w_{15}, w_{17}, w_{18}, w_{23}, w_{24}, w_{26}$
$\frac{1}{6}$	$w_5, w_6, w_8, w_9, w_{11}, w_{12}, w_{13}, w_{16}, w_{20}, w_{21}, w_{22}, w_{25}, w_{27}$
$\frac{2}{6}$	$w_2, w_3, w_4, w_7, w_{10}, w_{19}$
$\frac{3}{6}$	w_1

Here are three ways to restore the consistency of this database:

1. $D \div (h \vee r \vee w) = \{w_{14}, w_{15}, w_{17}, w_{18}, w_{23}, w_{24}, w_{26}, w_{27}\}$
2. $D \div \neg(h \vee r \vee w) = \{w_5, w_6, w_8, w_9, w_{11} - w_{18}, w_{20} - w_{26}\}$
3. $D \div (h \vee r \vee w) \wedge \neg(h \vee r \vee w) = \{w_5, w_6, w_8, w_9, w_{11} - w_{18}, w_{20} - w_{27}\}$

Removing paraconsistent states from these sets we get:

1. $D \div (h \vee r \vee w) = \{w_{27}\} = \neg h \wedge \neg r \wedge \neg w$
2. $D \div \neg(h \vee r \vee w) = \{w_9, w_{21}, w_{25}\} = (h \wedge \neg r \wedge \neg w) \vee (\neg h \wedge r \wedge \neg w) \vee (\neg h \wedge \neg r \wedge w)$
3. $D \div (h \vee r \vee w) \wedge \neg(h \vee r \vee w) = \{w_9, w_{21}, w_{25}, w_{27}\}$

□

As can be seen, in this case $\mathbf{D} \div A \wedge \neg A = \{\mathbf{D} \div A\} \cup \{\mathbf{D} \div \neg A\}$. More generally, we have:

Theorem 3.4.3. if $A \wedge \neg A \in \mathbf{D}$ then $\mathbf{D} \div A \wedge \neg A = \{\mathbf{D} \div A\} \cup \{\mathbf{D} \div \neg A\}$

Proof. See Appendix B, Section B.3

□

Theorem 3.4.4. if $A \wedge \neg A \in \mathbf{D}$ and the removal of this contradiction results in a consistent database [paraconsistent states removed from the result], it is never the case that both $\text{info}(D \div A \wedge \neg A) > \text{info}(D \div A)$ and $\text{info}(D \div A \wedge \neg A) > \text{info}(D \div \neg A)$

Proof. • For $\text{info}_{TO}()$, we use a proof by contradiction. Let the distance measure Δ of $D \div A$ be $\frac{d_1}{|W_1|}$ ($\text{info} = 1 - \Delta$) and the measure of $D \div \neg A$ be $\frac{d_2}{|W_2|}$, with W_1 being set of states corresponding to $D \div A$ and W_2 being the set of states corresponding to $D \div \neg A$. By Theorem 3.4.3 $D \div A \wedge \neg A = W_1 \cup W_2$, so the measure for $D \div A \wedge \neg A = \frac{d_1 + d_2}{|W_1 + W_2|}$. For $\frac{d_1}{|W_1|} < \frac{d_1 + d_2}{|W_1 + W_2|}$ it needs to be the case that $d_1 W_2 < d_2 W_1$. For $\frac{d_2}{|W_2|} < \frac{d_1 + d_2}{|W_1 + W_2|}$ it needs to be the case that $d_2 W_1 < d_1 W_2$.

Therefore both cannot be the case.

- For the value aggregate approach, either W_1 or W_2 has the highest valued member. Adding one onto the other will just decrease the $\text{sum}()$ value contribution of this highest valued member.

□

Carrying out consolidation

In order to consolidate a database contradictory elements must be removed. Which contradictory elements must be removed and how exactly to go about this is a matter with multiple solutions. One way to select the elements for removal consists of identifying each contradictory pair $(A, \neg A)$ in $\mathbf{D}$ and making a contraction so that $A \notin \mathbf{D} \vee \neg A \notin \mathbf{D}$. The following describes an implementation of this procedure:

1. Convert the contradictory D to *full* conjunctive normal form.⁷ The resulting formula will have contradictory pairs of the form $(p_1 \vee p_2 \vee \dots \vee p_n, \neg p_1 \wedge \neg p_2 \wedge \dots \wedge \neg p_n)$.
2. For each such contradictory pair, select either $p_1 \vee p_2 \vee \dots \vee p_n$ or a conjunction consisting of one or more of the conjuncts $\neg p_1, \neg p_2, \dots, \neg p_n$. Add the selected element to an array X .
3. Starting with the first element, for each element in $x \in X$, perform the contraction $D \div x$. If the resulting set of states contains at least one classical state, return the resulting set; else move on to the next element.
4. Reduce the final result to its set of classical states.

⁷See Chapter 2, Section A.3

Call this procedure the *Conjunction Removal Consolidation Procedure* (CRCP).

Theorem 3.4.5. The set of states that the Conjunction Removal Consolidation Procedure steps 1-3 result in will contain a subset of classical states.

Proof. See Appendix B, Section B.3

□

Example 3.16

Applying CRCP to Example 3.15, we get the following possibilities:

- $D = (h \vee r \vee w) \wedge \neg(h \vee r \vee w)$
 - $D \div (h \vee r \vee w) = w_{27} = \neg h \wedge \neg r \wedge \neg w$
 - $D \div \neg(h \vee r \vee w) = w_9 \vee w_{21} \vee w_{25} = (h \wedge \neg r \wedge \neg w) \vee (\neg h \wedge r \wedge \neg w) \vee (\neg h \wedge \neg r \wedge w)$
 - $D \div \neg h = w_9 = h \wedge \neg r \wedge \neg w$
 - $D \div \neg r = w_{21} = \neg h \wedge r \wedge \neg w$
 - $D \div \neg w = w_{25} = \neg h \wedge \neg r \wedge w$
 - $D \div \neg h \wedge \neg r = w_9 \vee w_{21} = ((h \wedge \neg r) \vee (\neg h \wedge r)) \wedge \neg w$
 - $D \div \neg h \wedge \neg w = w_9 \vee w_{25} = ((h \wedge \neg w) \vee (\neg h \wedge w)) \wedge \neg r$
 - $D \div \neg r \wedge \neg w = w_{21} \vee w_{25} = ((r \wedge \neg w) \vee (\neg r \wedge w)) \wedge \neg h$

□

It is already evident that a paraconsistent framework affords more flexibility than a classical framework. Consider how both deal with a situation like that presented in Example 3.15, where:

- $D = \neg h \wedge \neg r \wedge \neg w$
- $A = h \vee r \vee w$

The classical revision $D * A$ results in $(h \wedge \neg r \wedge \neg w) \vee (\neg h \wedge r \wedge \neg w) \vee (\neg h \wedge \neg r \wedge w)$. Since we can simulate classical revision by using the Reverse Levi Identity, we could just as easily

carry out $(D + A) \div \neg A$. But as demonstrated in example 3.16, given the expansion $D + A$, there is a range of other results in which the data is made consistent. Thus rather than selectively contracting some content in order to make way for its negation, we can ‘chuck everything in’ first and then selectively remove elements in order to restore consistency.

With CRCP it is important that each element in X is contracted one at a time, rather than contracting their conjunction, as the following example demonstrates:

Example 3.17

$$D = h \wedge r \wedge w \wedge \neg h \wedge \neg r \wedge \neg w$$

- Let $X = [\neg h, \neg r, \neg w]$.
- $D \div (\neg h \wedge \neg r \wedge \neg w)$ results in the set of states $\{w_5, w_{11}, w_{13}, w_{14}\}$.
- There are no classical states in this set and they actually equate to the contradiction $(h \wedge r \wedge \neg r \wedge w \wedge \neg w) \vee (h \wedge \neg h \wedge r \wedge w \wedge \neg w) \vee (h \wedge \neg h \wedge r \wedge \neg w \wedge w)$.

Successive contraction is required. For example, $((D \div \neg h) \div \neg r) \div \neg w = h \wedge r \wedge w$. First perform $D \div \neg h$, place the resulting states in the centre sphere and recalculate the sphere orderings. Then perform $\div \neg r$ on this new ordering and continue so on.

□

Where n denotes the number of contradictory pairs $(A, \neg A)$ in $\mathbf{D}$, it follows that the size of X is n and the number of possible values for X is 2^n . Though CRCP imposes no constraints on which contradictory element is selected, there are cases where the form of an inconsistent database implies that one selection for X should be made over others because it is more economical. This is demonstrated in the following example.

Example 3.18

- $D = (h \vee r) \wedge (r \vee w) \wedge \neg h \wedge \neg r \wedge \neg w$.
- D is satisfied in states $w_{14}, w_{15}, w_{17}, w_{23}, w_{24}$
- The set of contradictory pairs is $\{(h \vee r, \neg h \wedge \neg r), (r \vee w, \neg r \wedge \neg w)\}$

The initial sphere ordering is as follows:

Δ	states
0	$w_{14}, w_{15}, w_{17}, w_{23}, w_{24}$
$\frac{1}{6}$	$w_5, w_6, w_8, w_{11}, w_{12}, w_{13}, w_{16}, w_{18}, w_{20}, w_{21}, w_{22}, w_{26}, w_{27}$
$\frac{2}{6}$	$w_2, w_3, w_4, w_7, w_9, w_{10}, w_{19}, w_{25}$
$\frac{3}{6}$	w_1

Here are the results of various contractions (each result has been reduced to its classical states).

- $D \div \neg r = w_{21} = \neg h \wedge r \wedge \neg w$
- $(D \div \neg h) \div \neg r = w_3 \vee w_{21} = (h \wedge r \wedge \neg w) \vee (\neg h \wedge r \wedge \neg w)$
- $(D \div \neg h) \div \neg w = w_7 = h \wedge \neg r \wedge w$
- $D \div (h \vee r) = w_{27} = \neg h \wedge \neg r \wedge \neg w$
- $D \div (r \vee w) = w_{27} = \neg h \wedge \neg r \wedge \neg w$
- $(D \div (h \vee r) \div \neg r) = w_{21} \vee w_{27}$

Clearly $\neg r$ is common to both contradictions. Given this, an optimal selection would be such that for each pair $\neg r$ is selected. More generally, the key point is that the first element of X is $\neg r$ ($X[0] = \neg r$). Alternatively, selections such that either $X[0] = (h \vee r)$ or $X[0] = (r \vee w)$ would also be optimal selections in that the procedure would reach classical states and a classically satisfiable formula after the first contraction.

□

Another somewhat more direct approach based on a simple examination of the initial sphere ordering involves locating the closest sphere at which there is a set of classical states S_c and selecting a subset of S_c . Call this the *Closest Classical Worlds Procedure* (CCWP).

In Example 3.15 the states in $S_c = \{w_9, w_{21}, w_{25}, w_{27}\}$ are at a distance of $\frac{1}{6}$. Out of the 15 non-empty possible subsets of S_c , these are the 6 ones that cannot be generated using CRCP:

1. $\{w_9, w_{21}, w_{27}\} = ((h \wedge \neg r) \vee (\neg h \wedge r) \vee (\neg h \wedge \neg r)) \wedge \neg w$

2. $\{w_9, w_{25}, w_{27}\} = ((h \wedge \neg w) \vee (\neg h \wedge w) \vee (\neg h \wedge \neg w)) \wedge \neg r$
3. $\{w_{21}, w_{25}, w_{27}\} = ((r \wedge \neg w) \vee (\neg r \wedge w) \vee (\neg r \wedge \neg w)) \wedge \neg h$
4. $\{w_9, w_{27}\} = \neg r \wedge \neg w$
5. $\{w_{21}, w_{27}\} = \neg h \wedge \neg w$
6. $\{w_{25}, w_{27}\} = \neg h \wedge \neg r$

CRCP can technically also generate results that CCWP cannot. Going back to Example 3.18, the three possible results that CCWP can generate are:

- $\{w_{21}, w_{27}\}$
- $\{w_{21}\}$
- $\{w_{27}\}$

As can be seen in Example 3.18, there are results that CRCP can generate such as $(D \div \neg h) \div \neg r = w_3 \vee w_{21}$. But although CCWP cannot generate this result, it is a non-conservative result that does not adhere to a principle of content preservation, going beyond the minimal contraction required to obtain consistency. We can say that CCWP is a more direct route to complete consolidation whereas CRCP is an incremental procedure that is more guided by the contradiction structure and minimal way to get consistency.

As well as complete consolidation, this paraconsistent belief revision system also allows for *localised* consolidation. A standard way to perform consolidation for belief bases is to contract by falsum (contradiction): $A! = A \div \perp$. This means that all inconsistencies are removed. A localised consolidation on the other hand:

consolidates only a compartment of the belief base [set]. Contrary to a full consolidation, this process will not eradicate all inconsistencies. This is a realistic feature, since in real life inconsistencies are often tolerated, and do not propagate to make the whole belief state degenerate. [95, p. 47]

The two consolidation methods given in this section provide broad ways to get from an inconsistent set of data to a consistent one, with little in the way of which possible result it would be optimal to accept. Beyond this, there is room to implement further requirements which constrain possible results. For example, one requirement could be that the result

should be one of maximal content [i.e $\text{cont}()$]. In this case, $D \div \neg w$ would be an acceptable contraction in Example 3.16 whereas $D \div \neg h \wedge \neg r$ would not.

Another feature that could be implemented involves combining a core data *protecting consolidation* with information estimation, similar to what was done in combining screened revision with information estimation in Section 3.4. The consolidation of $\mathbf{D}$ protecting $\mathbf{C}$ can be denoted $\mathbf{D}!_C$. Here is a sketch of how they could work.

Given some core data $\mathbf{C}$, a protecting consolidation based on CCWP could work by going to the closest sphere which has classical states that also satisfy $\mathbf{C}$.

A protecting consolidation based on CRCP would require constructing an X such that carrying out the contraction of each $x \in X$ would not result in $\mathbf{C} \notin \mathbf{D}$. Some checks would have to be taken in constructing X so as not to violate $\mathbf{C}$. For example, take the following:

- $D = h \wedge \neg h \wedge r \wedge \neg r$
- $C = h \vee r$

Whilst neither $D \div h$ nor $D \div r$ violate C , $X = [h, r]$ will result in the removal of C .

Comparing Classical and Paraconsistent Approaches Non-Prioritised Belief Revision

In [95], the decomposition problem for non-prioritised belief revision is discussed. As we have seen, two ways to construct a non-prioritised revision operator are:

Decision + revision

1. Decide whether the input a should be accepted or rejected.
2. If a was accepted, revise by a .

and

Expansion + consolidation

1. Expand by a .

2. Consolidate the belief state.

Screened revision and accommodative revision fall under the former in this biclassification. The protecting consolidation outlined above offers an example of the latter. The question “what is the relation between decision+revision and expansion+consolidation?” [95, p. 47] is an interesting philosophical problem of belief revision. The tools incorporated and methods developed in this section to deal with uncertainty and conflicting input can contribute to the technical side of this investigation. The following comparison is an example.

Example 3.19

Take the following details, where state notation is in the classical context (Table 3.1):

- $w_2 = h \wedge r \wedge \neg w$
- $w_4 = h \wedge \neg r \wedge \neg w$
- $w_7 = \neg h \wedge \neg r \wedge w$
- $w_8 = \neg h \wedge \neg r \wedge \neg w$
- $D = w_2$
- $C = w_2 \vee w_4 \vee w_7$
- $A = w_7 \vee w_8 = \neg h \wedge \neg r$

Using accommodative revision we have:

- $A * C = w_7$
- $D * (A * C) = w_7 = \neg h \wedge \neg r \wedge w$
- $\text{info}_{TO-est}(\neg h \wedge \neg r \wedge w | C) = 0.44$

With a paraconsistent consolidation approach, we start by expanding D with A :

- $D + A = h \wedge r \wedge \neg w \wedge \neg h \wedge \neg r$

Converting this formula to its set of paraconsistent states we have the following distance ordering:

Δ	states
0	w ₁₄ , w ₁₅
$\frac{1}{6}$	w ₅ , w ₆ , w ₁₁ , w ₁₂ , w ₁₃ , w ₁₇ , w ₁₈ , w ₂₃ , w ₂₄
$\frac{2}{6}$	w ₂ , w ₃ , w ₄ , w ₈ , w ₉ , w ₁₀ , w ₁₆ , w ₂₀ , w ₂₁ , w ₂₂ , w ₂₆ , w ₂₇
$\frac{3}{6}$	w ₁ , w ₇ , w ₁₉ , w ₂₅

In a classical setup the maximum value $\text{info}_{TO-est}(X|C)$ takes is when $X = h \wedge \neg r \wedge \neg w$: $\text{info}_{TO-est}(h \wedge \neg r \wedge \neg w|C) = 0.67$. This is a possible result using both the CRCP and CCWP procedures. For the CRCP procedure, $h \wedge \neg r \wedge \neg w$ (w₉ in paraconsistent context notation) is the result of reducing $((D + A) \div \neg h) \div r$ to its classical states. For CCWP, the sphere where $\Delta = \frac{2}{6}$ is the closest sphere at which there is a classical set of states S_c . Since w₉ $\in S_c$, it is a possible result. What is happening here is that according to estimation calculations based on C , it is optimal to accept only part of the input A and reject the rest. Thus unlike the classical accommodative revision a paraconsistent approach affords this selective acceptance and rejection.

On the other hand given the following change to C it turns out that accommodative revision offers the most optimal result according to estimation calculations based on this C :

- w₃ = $h \wedge \neg r \wedge w$
- w₅ = $\neg h \wedge \neg r \wedge w$
- $C = w_2 \vee w_3 \vee w_5 \vee w_7$

In a classical setup there are several values X can take such that $\text{info}_{TO-est}(X|C)$ is maximal. One of them is $\neg h \wedge \neg r \wedge w$, which is the result of using accommodative revision. In this case for every such maximal value X , $X \vdash w$. However, using both the CRCP and CCWP procedures, no such X is obtainable. For CRCP, $D + A \vdash \neg w$ and $D + A \not\vdash w$. For CCWP, each classical state w_c such that w_c $\vdash w$ is not located in the closest sphere in which there is a classical state.

□

3.5 Applications

As we draw to the close of the last two chapters, I would like to go over some possible applications of the ideas that have been developed.

3.5.1 Extension to Other Spaces

There is room to adapt the methods developed over the last two chapters to spaces other than logical ones. To briefly demonstrate this, we recall a motivating example given at the beginning of the previous chapter and show how it can be dealt with.

Example 3.20

A die is rolled and lands on 4. Given the following collection of statements describing the outcome of the roll, which is the most informative? Which is the least informative?

A: The die landed on 1

B: The die landed on 1 or 2 or 4

C: The die landed on 4

D: The die landed on 4 or 5

E: The die did not land on 4

For this example, we divide the problem space into six constituents (i.e. mutually exclusive and exhaustive) possible states: $\{1, 2, 3, 4, 5, 6\}$. For any two constituents x and y , the distance (Δ) between x and y is: $\Delta(x, y) = |x - y|$. Using the average (Tichy/Oddie) metric we get the following ordering, with informativeness inversely related to distance:

C: $\Delta(4, 4) = 0$

D: $\Delta(4 \vee 5, 4) = 0.5$

B: $\Delta(1 \vee 2 \vee 4, 4) = 1.67$

E: $\Delta(1 \vee 2 \vee 3 \vee 5 \vee 6, 4) = 1.8$

A: $\Delta(1, 4) = 3$

Clearly ‘4’ is the most informative response, an intuition that these calculations accord with. Furthermore, according to these calculations ‘1’ is the least informative response.

□

3.5.2 The Value of Information

In this section we test both the CSI and truthlikeness information measures, applying them to a couple of cases involving information quantification.

Example 3.21

A crime is committed and there are four suspects. Furthermore, it is certain that one and only one of the suspects committed the crime. Three witnesses each provide evidence to clear the innocent suspects. On day one of the investigation, Witness 1 produces information to conclusively rule out the possibility that Suspect 1 committed the crime. The following day Witness 2 produces information to conclusively rule out the possibility that Suspect 2 committed the crime. On the final day, Witness 3 produces information to conclusively rule out the possibility that Suspect 3 committed the crime. Consequently, Suspect 4 is convicted and imprisoned.

Letting S_x stand for ‘Suspect x committed the crime’, Table 3.7 lists the four possible states given the situation, with all irrelevant logical possibilities omitted. w_{15} is the actual state.

	S_1	S_2	S_3	S_4	State	Pr()
w_8	T	F	F	F	$S_1 \wedge \neg S_2 \wedge \neg S_3 \wedge \neg S_4$	0.25
w_{12}	F	T	F	F	$\neg S_1 \wedge S_2 \wedge \neg S_3 \wedge \neg S_4$	0.25
w_{14}	F	F	T	F	$\neg S_1 \wedge \neg S_2 \wedge S_3 \wedge \neg S_4$	0.25
w_{15}	F	F	F	T	$\neg S_1 \wedge \neg S_2 \wedge \neg S_3 \wedge S_4$	0.25

Table 3.7: Four possible outcomes of the crime.

Suppose that the investigation team wanted to measure the amount of information they received from each witness, perhaps because they intended to financially reward the witnesses in proportion to the amount of information they provided. Or perhaps at a particular stage of the investigation the team wants to quantify how much information they have. How could they go about this? Since this scenario involves the sequential addition of information,

measures of relative information are called for.

Regarding CSI, we recall four measures of relative information given in Section 2.1 of Chapter 2, two for each of $\text{cont}()$ and $\text{inf}()$:

1. $\text{cont}_{add}(A|B) = \text{cont}(A \wedge B) - \text{cont}(B) = \text{cont}(B \supset A) = \text{cont}(\neg B \vee A)$
2. $\text{cont}_{cond}(A|B) = 1 - \text{Pr}(A|B)$
3. $\text{inf}_{add}(A|B) = \text{inf}(A \wedge B) - \text{inf}(B)$
4. $\text{inf}_{cond}(A|B) = -\log_2(\text{Pr}(A|B))$

Since $\text{inf}_{add}(A|B) = \text{inf}_{cond}(A|B)$, this list essentially reduces to three and we can omit the subscript to get $\text{inf}()$.

Let $D = w_8 \vee w_{12} \vee w_{14} \vee w_{15}$ stand for the agent's starting data. We will omit D from the CSI calculations but use the four states it determines in Table 3.7 as the starting point (this will not affect the correctness of these demonstrations). Here are the results for each measure:

$\text{cont}_{add}()$

- $\text{cont}(\neg S_1) = 0.25$
- $\text{cont}_{add}(\neg S_2|\neg S_1) = \text{cont}(S_1 \vee \neg S_2) = 0.25$
- $\text{cont}_{add}(\neg S_3|\neg S_1 \wedge \neg S_2) = \text{cont}(S_1 \vee S_2 \vee \neg S_3) = 0.25$

$\text{cont}_{cond}()$:

- $\text{cont}_{cond}(\neg S_1) = 1 - \frac{3}{4} = \frac{1}{4}$
- $\text{cont}_{cond}(\neg S_2|\neg S_1) = 1 - \text{Pr}(\neg S_2|\neg S_1) = 1 - \frac{2}{3} = \frac{1}{3}$
- $\text{cont}_{cond}(\neg S_3|\neg S_1 \wedge \neg S_2) = 1 - \frac{1}{2} = \frac{1}{2}$

$\text{inf}()$:

- $\text{inf}(\neg S_1) = 0.415$

- $\inf(\neg S_2 | \neg S_1) = -\log_2(\Pr(\neg S_2 | \neg S_1)) = -\log_2(\frac{0.5}{0.75}) = 0.585$
- $\inf(\neg S_3 | \neg S_1 \wedge \neg S_2) = -\log_2(\Pr(\neg S_3 | \neg S_1 \wedge \neg S_2)) = -\log_2(\frac{0.25}{0.5}) = 1$

As can be seen, $\text{cont}_{add}()$ determines that each witness provides the same amount of information. For the other measures, the closer the deposition was to the end, the more informative it was. It would thus seem fairest to base a reward distribution on the $\text{cont}_{add}()$ measure, particularly since each contribution taken in isolation (unconditionally) provides the same amount of information. But on the other hand, there seems to be something more rewarding about the final contribution of evidence as opposed to the first.

Using the truthlikeness information measures, here are the results:

info_{TO}():

- $\text{info}_{TO}(\neg S_1 | D) = \frac{2}{3} - \frac{5}{8} = \frac{1}{24}$
- $\text{info}_{TO}(\neg S_2 | D \wedge \neg S_1) = \frac{3}{4} - \frac{2}{3} = \frac{1}{12}$
- $\text{info}_{TO}(S_3 | D \wedge \neg S_1 \wedge \neg S_2) = 1 - \frac{3}{4} = \frac{1}{4}$

info_{VA}():

- $\text{info}_{VA}(\neg S_1 | D) = \frac{2}{64} = 0.03125$
- $\text{info}_{VA}(\neg S_2 | D \wedge \neg S_1) = \frac{2}{64} = 0.03125$
- $\text{info}_{VA}(S_3 | D \wedge \neg S_1 \wedge \neg S_2) = \frac{2}{64} = 0.03125$

Interestingly, as can be seen the behaviour of $\text{info}_{TO}()$ with regards to the differences between its calculation results parallels that of $\text{cont}_{cond}()$ and $\inf()$ whilst the behaviour of $\text{info}_{VA}()$ parallels that of $\text{cont}_{add}()$ in that each calculation result has the same value.

If a witness were to directly provide the information that S_4 , then as expected, the measure of this information is inversely related to how much information has already been obtained. That is, for all information measures, generically represented here by $\text{info}()$, the following holds: $\text{info}(S_4) > \text{info}(S_4 | \neg S_1) > \text{info}(S_4 | \neg S_1 \wedge \neg S_2) > \text{info}(S_4 | \neg S_1 \wedge \neg S_2 \wedge \neg S_3)$.

□

Example 3.22

Let us return to the ubiquitous weather theme used throughout the last two chapters. Person 1 is in Melbourne and wants to know what the weather is like in Sydney, where it happens to be hot, rainy and windy.

Person 2 informs Person 1 that it is hot in Sydney, Person 3 informs Person 1 that it is rainy in Sydney and Person 4 informs Person 1 that it is windy in Sydney. What are the amounts of information provided by Person 2, Person 3 and Person 4?

For the CSI measures, we get the following:

cont_{add}():

- $\text{cont}_{add}(h) = 0.5$
- $\text{cont}_{add}(r|h) = \text{cont}(\neg h \vee r) = 0.25$
- $\text{cont}_{add}(w|h \wedge r) = \text{cont}(\neg h \vee \neg r \vee w) = 1 - 0.875 = 0.125$

cont_{cond}():

- $\text{cont}_{cond}(h) = 1 - 0.5 = 0.5$
- $\text{cont}_{cond}(r|h) = 1 - \Pr(r|h) = 1 - 0.5 = 0.5$
- $\text{cont}_{cond}(w|h \wedge r) = 1 - \Pr(w|h \wedge r) = 1 - 0.5 = 0.5$

inf():

- $\text{inf}(h) = 1$
- $\text{inf}(r|h) = -\log_2(\Pr(r|h)) = -\log_2(0.5) = 1$
- $\text{inf}(w|h \wedge r) = -\log_2(\Pr(w|h \wedge r)) = -\log_2(0.5) = 1$

For the truthlikeness information measures, we get the following:

info_{TO}():

- $\text{info}_{TO}(h|D) = \frac{1}{6}$
- $\text{info}_{TO}(r|D \wedge h) = \frac{1}{6}$
- $\text{info}_{TO}(w|D \wedge h \wedge r) = \frac{1}{6}$

$\text{info}_{VA}()$:

- $\text{info}_{VA}(h|D) = 0.3333$
- $\text{info}_{VA}(r|D \wedge h) = 0.125$
- $\text{info}_{VA}(w|D \wedge h \wedge r) = 0.0417$

As can be seen, in this example things are reversed. Whereas now $\text{info}_{TO}()$, $\text{cont}_{cond}()$ and $\text{inf}()$ assign the same information measure to each contribution, $\text{info}_{VA}()$ and $\text{cont}_{add}()$ assign gradually decreasing measures; the greater the information already given, the lower the information measure assigned to new input.

□

As has been shown with these calculations, there is a distinction between the CSI measures that is mirrored in a distinction between the truthlikeness measures.

Dealing with Misinformation

In the examples thus far, each piece of evidence provided has been true. But what about if some piece of evidence is false?

Example 3.23

Take the following instances of communication:

- Person 2 informs Person 1 that it is hot in Sydney
- Person 3 misinforms Person 1 that it is not rainy in Sydney
- Person 4 informs Person 1 that it is windy in Sydney

- Person 5 informed Person 1 that it is rainy in Sydney

What is the informational value of the contributions made by Person 2, Person 3, Person 4 and Person 5?

Starting with the CSI measures, we get the following:

cont_{add}():

- $\text{cont}_{add}(h) = 0.5$
- $\text{cont}_{add}(\neg r|h) = \text{cont}(\neg h \vee \neg r) = 0.25$
- $\text{cont}_{add}(w|h \wedge \neg r) = \text{cont}(\neg h \vee r \vee w) = 1 - 0.875 = 0.125$
- $\text{cont}_{add}(r|h \wedge \neg r \wedge w) = \text{cont}(\neg h \vee r \vee \neg w \vee r) = 0.125$

cont_{cond}():

- $\text{cont}_{cond}(h) = 1 - 0.5 = 0.5$
- $\text{cont}_{cond}(\neg r|h) = 1 - \Pr(\neg r|h) = 1 - 0.5 = 0.5$
- $\text{cont}_{cond}(w|h \wedge \neg r) = 1 - \Pr(w|h \wedge \neg r) = 1 - 0.5 = 0.5$
- $\text{cont}_{cond}(r|h \wedge \neg r \wedge w) = 1 - \Pr(r|h \wedge \neg r \wedge w) = 1$

inf():

- $\text{inf}(h) = 1$
- $\text{inf}(\neg r|h) = -\log_2(\Pr(\neg r|h)) = -\log_2(0.5) = 1$
- $\text{inf}(w|h \wedge \neg r) = -\log_2(\Pr(w|h \wedge \neg r)) = -\log_2(0.5) = 1$
- $\text{inf}(r|h \wedge \neg r \wedge w) = -\log_2(\Pr(r|h \wedge \neg r \wedge w)) = -\log_2(0) = \infty$

As can be seen, the introduction of conflicting evidence is problematic for CSI measures. For $\text{cont}_{add}()$, the contribution of the false $\neg r$ is assigned a higher value than the subsequent contribution of the true r . For $\text{cont}_{cond}()$ and $\text{inf}()$, once $\neg r$ is given, the subsequent input of r is determined to have maximal informativeness, in the case of $\text{inf}()$ this equates to

an incoherent value of ∞ . Things become better for $\text{cont}_{add}()$ if the order that $\neg r$ and r are given in is swapped around. But this would then mean that the misinformation $\neg r$ is determined to have maximal informativeness for $\text{cont}_{cond}()$ and $\text{inf}()$.

On the other hand, a truthlikeness information measure such as $\text{info}_{TO}()$ provides a more agreeable set of results:

- $\text{info}_{TO}(h|D) = \frac{1}{6}$
- $\text{info}_{TO}(\neg r|D \wedge h) = -\frac{1}{6}$
- $\text{info}_{TO}(w|D \wedge h \wedge \neg r) = \frac{1}{6}$
- $\text{info}_{TO}(r|D \wedge h \wedge \neg r \wedge w) = \frac{1}{3}$

As suggested by these results, the contribution of r can be valued twice as highly as the contributions of h and w since it rectifies a piece of misinformation.

□

3.5.3 Lottery-Style Scenarios

Kyburg's well-known lottery paradox [117] arises from the following type of scenario. Consider a fair 1000 ticket lottery that has exactly one winning ticket. Given this information about the lottery it is rational to accept that some ticket will win. It also seems rational to accept a proposition if the probability of its associated event occurring is very high. For any ticket i in this lottery, the probability of it not being the winning ticket is 0.999. Now, a probability greater than or equal to 0.999 is very high, so for any individual ticket i it is rational to accept the proposition that ticket i of the lottery will not win. However, given the closure of rational acceptance and belief under conjunction, this entails that it is rational to accept that every ticket will lose. But since there is the original necessarily rational belief that one ticket will win, this leads to a contradictory state, hence the paradox.

Formally, symbolize propositions about ticket i being a losing ticket as p_i . Symbolize 'a rationally believes' as B_a . The inconsistency is generated as follows:

1. $B_a \neg(p_1 \wedge p_2 \wedge \dots \wedge p_{1000})$ - belief that some ticket will win
2. $B_a p_1 \wedge B_a p_2 \wedge \dots \wedge B_a p_{1000}$ - for each ticket, a belief that it will lose

3. $B_a A \wedge B_a B \supset B_a(A \wedge B)$ - the conjunction principle for rational belief
4. $B_a(p_1 \wedge p_2 \wedge \dots \wedge p_{1000})$ - from (2) and (3)
5. $B_a((p_1 \wedge p_2 \wedge \dots \wedge p_{1000}) \wedge \neg(p_1 \wedge p_2 \wedge \dots \wedge p_{1000}))$ - contradictory conjunction of (1) and (4)

The lottery paradox demonstrates the mutual inconsistency of the following three principles:

1. It is rational to accept a proposition that has a significantly high probability.
2. If it is rational to accept a proposition A and it is rational to accept another proposition B , then it is rational to accept their conjunction $A \wedge B$.
3. It is not rational to accept a contradictory proposition.

Proposed treatments of this paradox abound. Kyburg himself addresses the issue by basically accepting 1 and rejecting 2. Alternatively, there are many who accept 2 and reject 1.

The Lottery Paradox and Our Epistemic Goal

Put simply, we may say that one of our epistemic goals is the maximisation of truth and minimisation of falsity. Might this goal then guide the formation of an acceptance criterion? Instead of suspending belief formation, the aim would be to form a consistent set of beliefs that maximises estimated truth. As indicated by the following quote, something along these lines is suggested by Igor Douven.

Given the same 10-ticket lottery, accept of nine tickets that they will lose, and believe of the remaining one that it will win. Clearly, there is no longer an inconsistency in your beliefs about the lottery. And there is a 90% chance that you have added eight true beliefs and two false ones to your stock of beliefs - still not a bad score (or if you think it is, take a 100-ticket lottery, or ...). Better yet, there is even a 10% chance that you have added nothing but true beliefs. [47, p. 211]

Something like this idea can be formally captured with our information estimation framework. When stuck in such a lottery-style dilemma, sometimes a decision-theoretic approach

will differentiate amongst alternatives and afford a clear way out. In this case, the criterion of rational acceptance would be based on estimated information rather than probability and the right choice would be one that maximises or meets a threshold of estimated information.

The outcome of this approach largely depends on the problem. To facilitate its investigation, we will frame a lottery-style problem in terms of our weather example. Recall that there are eight possible states and the probability distribution is such that for any state w , $\Pr(w) = \frac{1}{8}$.

Now, suppose that one is forming beliefs about which possible state is the actual one. Let the proposition p_x stand for ‘State x will not be the actual state’. So $p_1 \equiv w_2 \vee w_3 \vee w_4 \vee w_5 \vee w_6 \vee w_7 \vee w_8$ and for any $x \in [1, 8]$, $\Pr(p_x) = \frac{7}{8}$. Given this scenario, if the threshold for acceptance were a probability of 0.875⁸, then the lottery paradox would ensue.

What about if one were to adopt an approach guided by the goal of information maximisation instead? If they accept that seven of the states will not be the actual state and believe of the remaining state that it will be the actual state, then there is a 12.5% chance that they have added nothing but true beliefs and an 87.5% chance that they have added six true beliefs and two false ones; it is certain that they will form at least six true beliefs. This obviously contrasts with a suspension of selection, whereby no true or false beliefs are formed.

Notwithstanding these positive odds, a true picture of this approach requires estimation calculations which factor in both information and misinformation. Also, the direction of such a strategy depends on the particular information measure being used.

If $\text{info}_{TO}()$ is used, then going by Theorem 2.6.1 each choice will be given a starting measure of 0.5 given no other information (i.e. given $\top$). So even though some choices would result in more true beliefs than others, this difference would be offset by the cost of potential false beliefs.

$\text{info}_{VA}()$ on the other hand gives a set of disuniform estimated information measures from the start. The following four propositions have the initial highest estimated information measures:

$$\begin{aligned} \text{info}_{VA-est}(w_1 \vee w_8 | \top) &= \text{info}_{VA-est}(w_2 \vee w_7 | \top) = \text{info}_{VA-est}(w_3 \vee w_6 | \top) = \\ &\text{info}_{VA-est}(w_4 \vee w_5 | \top) = 0.6875. \end{aligned}$$

⁸Given the size of the logical space used in these examples the thresholds are relatively low, so bear in mind that they are only for demonstrate purposes and the points could just as well have been made with larger spaces and higher thresholds.

If a strategy for acceptance were based on the selection of one statement amongst the statements ranked highest in terms of estimated informativeness, then selection of any of these four propositions would be right. Alternatively, since these initial results for $\text{info}_{est}()$ do not determine that one particular statement has the highest measure, if an estimated utility threshold criterion of acceptance was employed (with 0.6875 being the threshold for acceptance) and acceptance was closed under conjunction, then lottery-style problems would again arise, because each of these choices is mutually inconsistent.

There are however cases where decision theoretic approaches to a threshold selection problem based on the utility of information avoid the issues of an approach based on probabilistic acceptance criterion. Consider the following possible ways:

- A variation in the informational value assigned to each element as described in Section 2.4.2 of Chapter 2 would produce results whereby the prescribed choice is consistent.
- If parameters are changed or added to reflect a preference for truth acquisition or falsity avoidance (see Sections 2.4.6 and 2.7 of Chapter 2), then there could be a variation in calculations.
- Sometimes enough background data will be collected to rank one and only one choice the highest. For example, using $\text{info}_{VA}()$, suppose that the threshold of acceptance for estimated information was 0.7 and the given evidence was $\neg w_7 \wedge \neg w_8$. From this, the following four have the highest estimated information measures:

$$\begin{aligned} \text{info}_{VA-est}(w_1 \vee w_2 \vee w_5 | \neg w_8 \wedge \neg w_7) &= \text{info}_{VA-est}(w_1 \vee w_2 \vee w_6 | \neg w_8 \wedge \neg w_7) = \\ \text{info}_{VA-est}(w_1 \vee w_2 \vee w_4 | \neg w_8 \wedge \neg w_7) &= \text{info}_{VA-est}(w_1 \vee w_2 \vee w_3 | \neg w_8 \wedge \neg w_7) = 0.7361 \end{aligned}$$

Applying a conjunction principle to the above, we get:

$$B(w_1 \vee w_2 \vee w_5) \wedge B(w_1 \vee w_2 \vee w_6) \wedge B(w_1 \vee w_2 \vee w_4) \wedge B(w_1 \vee w_2 \vee w_3) \vdash B(w_1 \vee w_2).$$

$$\text{with } \text{info}_{VA-est}(w_1 \vee w_2 | \neg w_8 \wedge \neg w_7) = 0.7361$$

Given further evidence $\neg w_6$, we get the unique highest being

$$\text{info}_{VA-est}(w_1 \vee w_2 \vee w_5 | \neg w_8 \wedge \neg w_7 \wedge \neg w_6) = 0.775$$

In these cases, the consequence of following the prescribed selection method is consistent.

In summary, the point of this sketch has been to show that there are some instances of lottery-style problems that can be dealt with by using a decision-theoretic criterion of acceptance instead of a probability threshold. If information is the utility, then our account of information estimation provides the means for such a decision-theoretic approach.

3.5.4 The Conjunction Fallacy

The *conjunction fallacy* describes the judgement that a conjunction of two events is more probable than both of the individual events. The classic example of this fallacy is found in the work of Tversky and Kahneman [173], who presented the following fictitious scenario to subjects:

Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

The subjects were asked which of the following two they thought more probable:

1. Linda is a bank teller.
2. Linda is a bank teller and is active in the feminist movement.

Amongst the subjects there was a tendency to judge 2 as more probable than 1; in fact, out of 142 university students who were simply asked to choose which they judged more probable, 85% of them chose 2.

This behaviour conflicts with the basic conjunction principle of probability theory, according to which the probability of a conjunction is always less than or equal to the probability of each conjunct individually: $\Pr(A \wedge B) \leq \Pr(A)$ and $\Pr(A \wedge B) \leq \Pr(B)$.

A range of responses to this puzzling phenomenon have been proposed. An interesting take on the paradox is presented in a recent paper titled ‘The whole truth about Linda: probability, verisimilitude, and a paradox of conjunction’ [29], in which the authors suggest estimated truthlikeness as an attribute guiding participants’ prevailing responses. Estimated truthlikeness is “an independently motivated and formally definable epistemological notion relying on which many judges would rank ‘feminist bank teller’ over ‘bank teller’ in the Linda problem” [29, p. 612].

So whilst the probability of the conjunction is always less than the probability of each of its conjuncts, the estimated truthlikeness of the conjunction can be higher than the estimated truthlikeness of one or both of its conjuncts. Here is a small formal demonstration of this idea using $\text{info}_{TO}()$.

Example 3.24

Let b stand for ‘Linda is a bank teller’ and f stand for ‘Linda is active in the feminist movement’. Suppose that the subject assigns different subjective probabilities to each of the four possibilities, as represented in Table 3.8, believing it most likely that Linda is both.

State	b	f	$\text{Pr}()$
w_1	T	T	$\frac{6}{16}$
w_2	T	F	$\frac{4}{16}$
w_3	F	T	$\frac{4}{16}$
w_4	F	F	$\frac{2}{16}$

Table 3.8: Linda scenario

Given this we have:

- $\text{Pr}(b) = \text{Pr}(f) = \frac{10}{16}$
- $\text{Pr}(b \wedge f) = \frac{6}{16}$
- $\text{info}_{TO-est}(b|\top) = \text{info}_{TO-est}(f|\top) = \frac{9}{16} = 0.5625$
- $\text{info}_{TO-est}(b \wedge f|\top) = 0.625$
- $\therefore \text{info}_{TO-est}(b \wedge f|\top) > \text{info}_{TO-est}(b|\top) = \text{info}_{TO-est}(f|\top)$

□

3.5.5 Quantifying Epistemic and Doxastic Content

As will be covered in Chapter 5, an account of information that implies truth can be used to develop an informational epistemology; knowledge encapsulates truth because it encapsulates semantic information. A truthlikeness approach to semantic information quantification thus potentially affords a tool to measure one aspect of knowledge.

To demonstrate this we return to the weather example. Let $\text{info}()$ represent an information quantification function in general. Furthermore, let $\text{info}_{Tr}()$ represent a truthlikeness version of such an information function and let $\text{info}_{Pr}()$ represent an inverse-probabilistic-style version of such an information function.

For both $\text{info}_{Tr}()$ and $\text{info}_{Pr}()$, it is the case that $\text{info}(h \wedge r) > \text{info}(h \vee r)$. This result founds and makes numerically precise a sense in which knowledge of $h \wedge r$ is greater than knowledge of $h \vee r$: $K(h \wedge r) > K(h \vee r)$.

It is also fair to suppose that $K(h \vee r) > K(h \vee \neg r)$. $\text{info}_{Tr}()$, according to which the informativeness of $h \vee r$ is greater than the informativeness of $h \vee \neg r$, can accommodate this point. On the other hand, since $\text{info}_{Pr}(h \vee r) = \text{info}_{Pr}(h \vee \neg r)$, $\text{info}_{Pr}()$ cannot be used with information to account for the sense in which $K(h \vee r) > K(h \vee \neg r)$. In this way, quantitative accounts of information based on the notion of truthlikeness can play a part in quantitative accounts of knowledge.

With regards to quantifying beliefs, things are the other way around and it is the inverse probabilistic approach instead which is to be applied. $\text{info}_{Tr}(h \wedge \neg r) < \text{info}_{Tr}(h \wedge r)$ and $\text{info}_{Pr}(h \wedge \neg r) = \text{info}_{Pr}(h \wedge r)$. Yet quantitatively speaking $B(h \wedge \neg r) = B(h \wedge r)$; although one belief is true and the other false, just as much is believed in both cases.

Quite simply, a belief Bp has the semantic content p . $\text{info}_{Pr}()$ is seen in this work as a measure of *semantic content* (meaningful data), rather than a measure of *semantic information* (*truthful*, meaningful data)⁹. Therefore $\text{info}_{Pr}()$ is suitable as a tool for quantifying belief in the same way that $\text{info}_{Tr}()$ works with knowledge.

3.5.6 Agent-Oriented Relevance

Amongst the properties associated with information, relevance is certainly one of them. Like information, it also is a broad, polysemous concept, studied in different fields and applied in a variety of ways. For our purposes the following short background will suffice.

Hitchcock [107] distinguishes two main types of relevance, *causal* and *epistemic*. In short, “causal relevance consists in an item’s ability to help produce an outcome in a situation. Epistemic relevance, a distinct concept, consists in the ability of a piece of information to help achieve an epistemic goal in a situation” [107, p. 1].

⁹Similarly, Floridi writes that “[CONT(s)] does not indicate the quantity of semantic information but, more precisely, the quantity of data in [s]” [76, p. 128]. I assume that data here means semantic content.

Similarly, Borland [21] divides conceptions of relevance into those which are more system-based and those which are more agent-oriented.

System-oriented theories (S-theories) usually analyse relevance in terms of *topicality*, *aboutness* or *matching* (how well some information matches a request), especially in the information retrieval (IR) literature, and various forms of *conditional in/dependence* (how some information can help to produce some outcome), especially in logic, probability theory, philosophy of science and AI. [69, p. 6]

On the other hand

Agent-oriented theories (A-theories), on the other hand, tend to analyse relevance in terms of *conversational implicature* and *cognitive pertinence*, especially in philosophy of language, pragmatics and psychology, and perceived *utility*, *informativeness*, *beneficiality* and other ways of “bearing on the matter at hand” in relation to an agent’s informational needs, especially in IR literature and in epistemology. [69, p. 6]

Linking the two classifications, S-theories can be seen to be interested mainly in causal relevance and A-theories can be seen to be interested mainly in epistemic relevance.

In this section we are interested in accounts of epistemic or agent-oriented relevance which deal with relevance in terms of an agent’s utility and their information requests. Here the relevance of a piece of information is judged relative to a question (how well does an answer satisfy a question), with questions essentially being treated as requests for information. There already exists a small but good body of literature on epistemic relevance.¹⁰ The aim of this section is basically to offer a cursory demonstration of the potential applicability of truthlikeness information measures to accounts of agent-oriented relevance. The basic idea is that if an agent’s question or information request can be translated into its corresponding true logical statement, then the closer an answer is to this true statement the more epistemically relevant it is to the agent.

It is worth noting a parallel between relevance and information quantification. As was discussed in the previous chapter, classical or traditional approaches to semantic information quantification measure the informativeness of a statement in terms of probability or the set of possible states that the statement rules out. On the other hand, approaches to semantic

¹⁰[69], [34] and [118] are some examples

information quantification based on a theory of strongly semantic information are based on truthlikeness measures.

There is a similar bifurcation when it comes to relevance. Firstly, there are traditional accounts of relevance based on conditional probability. Secondly, quantification methods for strongly semantic information can contribute to accounts of agent-oriented relevance. We will start with a brief look at the first type.

Conditional Probability and Relevance

A traditional and simple definition of the relevance relation, involving a probability measure $\Pr()$, is defined as follows:

- B is relevant to A on evidence E iff $P(A|B \wedge E) \neq P(A|E)$
- B is irrelevant to A on evidence E iff $P(A|B \wedge E) = P(A|E)$ ¹¹

This simple definition has a strong intuitive resonance; B is relevant to A means that it makes a difference to the status of A . Despite this the definition does have some shortcomings. A brief history of such work on the logic of relevance is given by Gardenfors, who provides the following definition of relevance accompanied by a rigorous underpinning for it [80, p. 362]:

- B is irrelevant to A on evidence E iff either
 1. $\Pr(A|B \wedge E) = \Pr(A|E)$ and for all sentences S , if $\Pr(A|S \wedge E) = \Pr(A|E)$ and $\Pr(B \wedge S \wedge E) \neq 0$, then $\Pr(A|B \wedge S \wedge E) = \Pr(A|E)$, or
 2. $\Pr(B|E) = 0$
- B is relevant to A on E iff B is not irrelevant to A on E .

Truthlikeness Information Quantification and Agent-Oriented Relevance

If Bob asks Harry whether or not Paris is the capital city of France (p), Harry can do one of three things:

¹¹As Carnap [26, p. 348] pointed out, if one adopts the following slightly modified version then theorems on irrelevance become simpler: B is irrelevant to A on evidence E iff $P(A|B \wedge E) = P(A|E)$ or $B \wedge E$ is logically false.

1. Respond ‘yes’ (p)
2. Remain silent, effectively ‘yes or no’ ($p \vee \neg p$)
3. Respond ‘no’ ($\neg p$)

If Bob is to act on what Harry says, then (1) is the most relevant response, (2) is in-between and (3) is the least relevant. The response $\neg h$ is misinformation which is in a sense negatively epistemically relevant (irrelevant?) to Bob’s information request.

Of course, traditional probabilistic accounts of relevance, in which truth values play no role, do not reflect this. A truthlikeness information metric on the other hand does and could be used to contribute to a relevance metric which captures this sense of relevance, as will now be sketched out.

Since agent-oriented relevance deals with agent requests for information (questions), erotetic logic would have a role to play. To this end we will use Hintikka’s approach to epistemically analysing questions as requests for information in terms of epistemic modal logic [105].

With Hintikka’s imperative epistemic or ‘Make Me Know’ (MMK) approach, the point of a question is to express an epistemic request and the point of an answer is to satisfy this request. Questions are analysed systematically as requests for information and to answer a question is to specify the epistemic state that the questioner wants to be brought about. So the three main ingredients to this analysis are:

1. imperative operator - bring it about that ...
2. presupposition - what the questioner presupposes in asking the question
3. desideratum - the epistemic state of affairs the questioner is asking to be brought about

There are many types of questions (yes-no, whether, who, etc) that can be analysed with Hintikka’s framework. For this sketch, it will suffice to deal with yes-no (polar) questions. Using epistemic logic, we can express the presupposition and desideratum of the above question (for some agent a) as follows:

- Presupposition: $K_a(p \vee \neg p)$
- Desideratum: $K_a p \vee K_a \neg p$.

Given this formal erotetic framework, the other required component is a truthlikeness information measure. For this demonstration, we will use the simple Basic Feature Approach to Truthlikeness given in Section 2.4.1 of Chapter 2.

Now that the required components have been selected, we can construct an example. Using the weather scenario, suppose that an agent wants to know whether or not it is hot and whether or not it is raining. This is a question with the desideratum $(K_a h \vee K_a \neg h) \wedge (K_a r \vee K_a \neg r)$ and the true statement which represents a complete answer to this question is simply $Q = h \wedge r$. The information relevance accuracy of a response R relative to Q is simply:

$$s(R, Q) = \text{cont}_t(R, Q) - \text{cont}_f(R, Q)$$

Obviously the only difference between this measure and the basic features truthlikeness measure is that it makes calculations against Q , which need not be the true state description. Therefore the result will generally not equate to information yield. In this example, $s(h \wedge r, Q) = 1 > \text{info}(h \wedge r) = \frac{2}{3}$.

This has been the barest of outlines on how a truthlikeness information measure can contribute to an account of agent-oriented relevance. In particular, such a measure can contribute to a relevance metric in calculating how adequately some piece of information satisfies some information request. For example, Floridi develops an account of epistemic relevance “based on a counterfactual and metatheoretical analysis of the degree of relevance of some semantic information i to an informee/agent a , as a function of the *accuracy of i understood as an answer to a query q* , given the probability that q might be asked by a ” [69] (*italics mine*). The accuracy of i understood as an answer to a query q is something that is not fleshed out in Floridi’s paper. A truthlikeness information measure would be of use by contributing towards measuring the accuracy of i in answering a query q , precisely capturing why h is more accurate than $\neg h$ and why $h \wedge r$ is more accurate than h .

Beyond this basic outline, here are several other points to consider:

- The contribution here concerns a purely quantitative measure of how adequate a response is in terms of how close it is to the complete truth corresponding to a question. As usual, there is room to add qualitative factors.
- Adapting this idea to richer frameworks that deal with the complete range of statements within a logical space is another task. Functions to measure the distance between two statements (see [138, p. 91] for example) could be of use.

Here are some issues that would arise

- Answers can be more nuanced; if a question corresponding to the truth $h \wedge r$ is asked, how relevant is the answer $h \vee r$?
- Questions can also be more nuanced. Suppose that Bob wants to know whether or not the following inclusive disjunction: ‘Helsinki is the capital of Finland (h) or Oslo is the capital of Sweden (o)’. It would seem that the response $h \vee o$ is completely relevant relative to the question. It also seems right to suppose that in general if $R \vdash Q$, then response R is completely relevant to question Q ¹². This makes sense when we judge the response h , which is completely relevant. But on the other hand, although $o \vdash h \vee o$ the response o is not relevant in the epistemic or agent-oriented sense as it is misinformation.
- Considerations of efficiency can also be brought in. This consists of two parts:
 1. Ranking according to what portion of the response consists of elements necessary to answer the question. For example, if $Q = r \wedge w$, then $R = r \wedge w$ is better than $R = h \wedge r \wedge w$. Some of these ideas are considered in [34, p. 27]
 2. Ranking according to the complexity of the inferences required to infer the requested information. For example, if $Q = w$ then $R = w$ is better than $R = (h \supset r) \wedge (r \supset w) \wedge h$.

3.6 Conclusion

One important aspect of semantic information is that its informativeness is relative to its reception. This chapter has taken the information quantification framework of the last chapter and combined it with a framework for belief revision in order to provide a formal logical account of agent-relative informativeness. Thus whilst there has been some interest in investigating the relationship between the truthlikeness and belief revision programs within the last few years, in the case of information this combination arises out of a need to capture a defined property and formally explicate how the information yield or absolute informativeness of some statement is distinguished from its agent-relative informativeness.

One prominent result of this account is that the insertion of information into a database does not necessarily result in an increase of its information yield. In order to deal with the absence of such guarantees and also with uncertainties in the veracity of data that is received, certain strategies can be employed. As was shown, one general idea involves employing

¹²Leaving aside cases of non-relevant classical inferences dealt with by relevant logic.

non-prioritised belief revision and using trusted data already in the database to guide the acceptance/rejection of incoming data. One method based on this idea that was introduced in this chapter involved a combination of screened belief revision with information estimation. Departing from classical logic, a paraconsistent system that deals with inconsistent data was also introduced and developed. This method also leads to an alternative approach to non-prioritised belief revision and a way to deal with the aforementioned issues concerning guarantees and uncertainties.

Finally, several applications of the information and informativeness framework developed over the last two chapters were outlined. One point that can be drawn from this is that as well as applying to situations calling for semantic information measurement, this framework can explicate or contribute towards accounts of other puzzles and phenomena.

Chapter 4

Environmental Information and Information Flow

In this chapter the focus is shifted to (semantic) environmental information, which was briefly introduced in Section 1.3.3 of Chapter 1. The main task will be to analyse its operation and flow. To begin with, we recall the following general definition of environmental information:

Environmental information: Two systems a and b are coupled in such a way that a 's being (of type, or in state) F is correlated to b being (of type, or in state) G , thus carrying for the information agent the information that b is G . [73]

This definition is quite general, and deliberately so, for environmental information can be found in any type of distributed system which has regularity and correlations between its elements. Here are some rehashed examples involving various systems to get us started:

- The doorbell system of a home is engineered in such a way that when the button is pressed the bell sounds. Thus the sounding of the bell carries the information that the button is being pressed.
- Given the uniqueness of human fingerprints, the presence of fingerprints on an object carries the information that so-and-so handled the object.
- That the output of a square root function is 7 carries the information that the input was 49.

Given this fairly basic and informal understanding of environmental information, we shall now look at some formal accounts. Our starting point is Dretske's account of information, which formed the basis of his informational epistemology [51]. After an exposition of Dretske's account and some commentary, we shall look at some of the debate that has followed from his work and alternative/rival accounts that have been offered. This will lead to some suggestions regarding an account of environmental information, partly in response to this debate, as well as the development of a simple formal framework to model and analyse the relevant alternatives strategy. There are several ways in which one can use the idea of relevant alternatives. My aim is not to argue specifically for one approach over another but rather to provide a common framework for understanding them and investigate their characteristics and implications.

4.1 Dretske on Information

Dretske's *Knowledge and the Flow of Information* [51] is ultimately an attempt to use the notion of information to explicate the notions of knowledge and intentional content. As part of this enterprise, several philosophically interesting issues are tackled. He firstly gives an account of the propositional content of a signal (events, structures, situations) and develops a semantic theory of information. For Dretske, information is an informee independent, objective phenomenon that occurs in a multitude of ways and which existed before the development of agents with the ability to selectively utilise it: "information ... [is] an objective commodity, something whose generation, transmission, and reception do not require or in any way presuppose interpretive processes" [51, p. vii]. Regularity and nomic correlations are crucial to information flow; without a lawfully regular universe, no information would flow.

Dretske uses the Mathematical Theory of Communication (MTC) as a starting point before acknowledging its limitations and departing to "develop a philosophically useful theory of information" [52, p. 55]. Although MTC was outlined in Chapter 1, here are a couple of repeated points particularly relevant to considerations in this section:

- MTC is a syntactic treatment of information and is not, at least directly, concerned with the semantic aspects of information. It is really an account of data quantification and transmission as opposed to a theory of information as it is commonly understood.

Whilst the quantity of data a signal carries can inform us about its content¹, it is

¹Taking an example from Dretske, suppose that a jug will either contain 1 litre of milk or 2 litres of water. Learning that the jug contains only 1 litre of liquid provides information about what type of liquid the jug contains, namely milk.

ultimately an insufficient tool for the analysis of semantic information.

- MTC is concerned with the statistical properties of transmission, with the average amount of information generated by a source. However information as it is commonly understood, and for Dretske's purposes, is something associated with individual events. It is only particular signals that have content which can be propositionally expressed.

So Dretske therefore must adapt the ideas of MTC ideas for his purposes. The formulas for the amount of information generated by a particular event s_a and for the information carried by a particular signal r_a about s_a are easily obtained from MTC formulas.

The amount of information (in bits) generated by s_a is given by:

$$I(s_a) = \log_2\left(\frac{1}{\Pr(s_a)}\right) = -\log_2(\Pr(s_a)) \quad (4.1)$$

The amount of information (in bits) carried by r_a is given by:

$$I_s(r_a) = I(s_a) - E(r_a) \quad (4.2)$$

where $E(r_a)$ stands for the equivocation associated with r_a . Where n is the number of mutually exclusive possible events at the source, the equivocation associated with a specific signal r_a is given by:

$$E(r_a) = -\sum_{i=1}^n \Pr(s_i|r_a) \times \log_2 \Pr(s_i|r_a) \quad (4.3)$$

Following are a couple of demonstrations of these formulas using some examples given by Dretske.

Example 4.1

A boss asks eight of his employees to select amongst themselves one individual to perform some task. Once this person is selected, they will inform the boss of their choice by writing the selected person's name down on a piece of paper and sending it to the boss. Each employer is uniquely named and each has an equal probability of being selected using some random selection process. The employees end up selecting Herman, writing his name down

and passing it on to their boss. Since the probability of Herman being selected (s_H) is $\frac{1}{8}$, $I(s_H) = 3$ bits. Since there is no equivocation in this communication, $I_s(r_H) = I(s_H) = 3$ bits.

□

Example 4.2

A small modification is made to the previous example so that equivocation is greater than zero. Everything is the same except that for some reason, the employees decide that should Barbara be selected as a result of their selection process, they will write Herman's name down on the note instead. So the message 'Herman' would now be used if either Herman or Barbara (s_B) were selected. Herman is once again selected.

The amount of information generated is still the same, so $I(s_H) = 3$ bits. This time however there is some equivocation associated with the received signal:

$$\begin{aligned} E(r_H) &= -[(\Pr(s_H|r_H) \times \log_2 \Pr(s_H|r_H)) + (\Pr(s_B|r_H) \times \log_2 \Pr(s_B|r_H))] \\ E(r_H) &= -[(0.5 \times -1) + (0.5 \times -1)] \\ E(r_H) &= 1 \end{aligned}$$

Therefore $I_s(r_H) = 3 - 1 = 2$ bits.

□

Moving towards a semantic theory of information, Dretske gives three conditions a theory of semantic information should satisfy, as a precursor to his main definition. If a signal carries the information that s is F , it must be the case that:

- (A) the signal carries as much information about s as would be generated by s 's being F .
- (B) s is F

(B) is simply a veridicality requirement; for Dretske, information requires truth.² Regarding (A), in the second example above the signal 'Herman' carried less information (2

²A discussion of the veridicality requirement for information in general, including Dretske's view, are covered in Section 1.4 of Chapter 1.

bits) than the amount generated by Herman being selected (3 bits), so it could not have been the case that ‘Herman’ carried the information that Herman was selected. Also, bear in mind that although figures for $I(s_a)$ or $I_s(r_a)$ might not be feasibly or realistically calculable, it is not necessary to calculate their values; a look back at Formula 4.2 and we see that knowing whether the equivocation is zero or not is sufficient.

These two conditions are necessary but insufficient. Suppose c is a green circle. Suppose further that both c ’s being green and c ’s being a circle carry 1 bit of information each. Now a signal carrying the information that c is a circle carries as much information as is generated by c ’s being green and c is green. Despite this, the signal does not carry the information that c is green. To deal with this a third condition is added:

- (C) The quantity of information the signal carries about s is (or includes) that quantity generated by s ’s being F (and not, say, by s ’s being G).

(A) is termed the *communication condition*. (B) and (C) are termed the *semantic conditions* on information. They are satisfied by the following theoretical definition of a signal’s (structure’s) information content given by Dretske:

A signal r carries the information that s is F = The conditional probability of s ’s being F , given r (and k), is 1 (but, given k alone, less than 1)

Here k is a variable that takes into account how what an agent already knows can determine the information carried (for that agent) by a signal. For example, the conditional probability of A given $A \vee B$ is not 1. However, if an agent already knows that $\neg B$, then for them the signal $A \vee B$ carries the information that A [$\Pr(A|\neg B \wedge A \vee B) = 1$]. Although the k variable which occurs in Dretske’s definition relativises information to what the receiver already knows concerning the possibilities at the source, this relativisation is only meant to accommodate the way information is thought about, that the information one can get from a signal depends on what they already know. It does not undermine the essential objectivity of the information. A simple division resulting in the terms ‘relative information’ and ‘absolute information’ can provide clarification here. If the message $A \vee B$ is contained in a signal and its receiver already knows that $\neg B$, then the relative information of the signal is A whilst the absolute information is $A \vee B$.

There is one issue about Dretske’s definition concerning how an agent’s background knowledge figures in determining the information they can receive from a signal. Whilst it

will be looked at more when knowledge is discussed in the next chapter, it is worth briefly mentioning here via an astute quote from Barwise:

Dretske emphasizes that it is only due to nomic relations between types of situations that one can carry information about the other. But when he defines the basic notion, he relativizes only to what the receiver knows about alternative possibilities. More important is what the receiver knows about the nomic relations. While information *is* out there, it *informs* only those attuned to the relations that allow its flow. [52]

So whilst a smoke alarm's beep carries to you the information that there is smoke, you cannot 'get to this information', cannot be informed that there is smoke, if you don't know that the alarm's beep means smoke. In this way, whilst the information carried by a signal is objective, the information one takes from this signal is subjective.

This aside ³, as an example of how the appeal to probability works here, take the information bearing signals of a clock. The signal from a correctly functioning clock carries the information that the time is such-and-such. Suppose that a digital clock displays 6:30pm. The conditional probability that the time is 6:30pm given the clock's signal of 6:30pm is one, so in accordance with Dretske's definition the clock carries the information that it is 6:30pm.

If the clock were malfunctioning, things would be different. Suppose that the clock stops working at 6:30pm. The next day someone happens to look at the clock when it is 6:30pm. Even though the time indicated by the clock happens to be the actual time, the clock signal here does not carry the information that the time is 6:30pm because the conditional probability that it is 6:30pm given that the clock shows this is less than one. Since there are 1440 minutes in a day, the probability that the time is 6:30pm given the non-functioning clock's signal of 6:30pm is in fact $\frac{1}{1440}$.

Here are three reasons given by Dretske for his stipulation that the value of the conditional probability in his definition of information be one and nothing less:

- One basic principle of information flow to be satisfied is the conjunction principle: if a signal carries the information that s is F and the information that s is G , then it also carries the information that s is F and G . If the conditional probability requirement was relaxed and made lower than one, then a signal could carry the information that s

³The basis for an account of information flow precedes this extra aspect. In the examples that follow it is set that the subjects are attuned to the relevant relations

is F and the information that s is G without carrying the information that s is F and G .

- Another principle of information flow that Dretske maintains is what he terms the ‘Xerox principle’: if A carries the information that B , and B carries the information that C , then A carries the information that C . So information flow is transitive. If the conditional probability threshold is set to anything less than one, then this intuitive principle fails.
- Finally, and as Dretske frankly acknowledges, there is no non-arbitrary figure at which to impose a threshold. If information can be obtained from a signal involving a conditional probability of less than one, then information loses its ‘cognitive punch’. To use an example of Dretske’s, think of a bag with ninety four red balls and six white balls. If one is pulled out at random, you cannot *know* that it was red. Hence why suppose you have the information that it is red?

A main concern against setting the required conditional probability to one is that since there are very few conditional probabilities of one out there, very little information ever flows. Dretske deals with these concerns by introducing the idea of *fixed channel conditions* and *relevant alternatives*. The basic idea is that the conditional probability requirements are made relative to a set of possibilities relevant to the communication channel. This is a more flexible and realistic way of thinking about the conditional probabilities that determine information flow. That calculations are made within a background of stable or fixed circumstances is not to say that circumstances cannot change. Rather it is only to say that for the purposes of determining conditional probabilities, if conditions are normal and there is no significant chance of something happening, such changes are set aside as irrelevant.

For example, take an actual scenario in which a functioning clock is accurately correlated with the time and is thus transmitting information about the time. Now, there are scenarios alternative to this one, such as cases where the clock has flat batteries, in which the clock’s signal is not correlated with the time and it therefore fails to transmit information about the time. If these alternatives are factored into the probability calculations then the probability that it is time x given that the clock shows x is not equal to one. Consequently, in the actual scenario the clock would not meet the requirements of Dretske’s information flow definition and so would be judged as not carrying the information that the time is such-and-such. But the idea is that if the batteries in the clock are in good working order and the clock is functioning correctly, even if there is a minute probability that the new batteries could become defective or a mechanism in the clock could break just before someone looks at the clock, these possibilities are ignored in calculating the information the clock is delivering.

Possible (non-actual) but far-flung or improbable alternatives such as these are deemed irrelevant in considering the set of relevant alternatives against which probability calculations are made. Even if there is technically a non-zero probability that they could occur, if these issues have not actually occurred, then they are excluded from consideration.

Adams [3, p. 233] provides a good discussion of this idea using a metal detector for his example. Al has a metal detector that he has just serviced, which emits a tone when metal is within 10 inches of its detection surface. In order for this tone to carry the information that there is metal present certain channel conditions need to be fixed. The power supply must be providing adequate power, the magnetic field that detects the metal must be in place, the wires through which the detector's signals run must not be broken, etc. These channel conditions must be fixed in order for the tone to carry the information that there is metal, for the relevant inverse conditional probability to be 1. Otherwise, there is a genuine non-zero probability that the detector could emit a tone even when no metal is present. Another way to put this is that possible but non-actual alternatives in which the system is playing up (e.g. the detector is short circuiting) are not relevant to determining the information carried by the detector's tone.

The detector's tone carries the information that metal is present. It does not though carry the information that the metal detector is working properly, that the required channel conditions are set. Rather, the tone carries the information that the metal is present given that the conditions are set. As will be covered further on, in this way the metal detector's tone does not provide *meta-information* that the detector is functioning correctly, that it is carrying the information that there is metal present.

But as Adams points out, given a different set of background conditions, the tone can carry the information that the detector is working:

If we know in advance that there is metal (or is none), we may check to see whether the detector emits a tone (or does not). Thereby, we can use old information about what we already know (there is/is not metal present) to gain information about the channel conditions of the detector. Now, because we *already know* whether or not there is metal present, we can test to see whether or not the detector is working properly. We can then tell that it is working properly and its channel conditions indeed are fixed (or tell that it is broken because its channel conditions are variable). [3, p. 234]

In such a situation, that there is metal present is a piece of information. But although the tone's activation is correlated with the presence of metal, since this presence is a part

of the fixed background conditions (every relevant alternative has metal), this signal does not reduce any uncertainty regarding the presence/absence of metal and therefore does not strictly carry the information that there is metal. It can though tell us whether or not the detector is working⁴.

So the detector's tone can be used to provide the information that there is metal or the information that the detector is functional/dysfunctional. But this is an exclusive disjunction; it cannot be used to simultaneously provide the information that there is metal and the information that the detector is functional/dysfunctional.

There is no determinate method to decide what counts and what does not count as a relevant alternative. In general the selection will depend upon the knowing agent and their environment and will also be a pragmatic decision. Particularly due to this lack of clear determination the notion of relevant alternatives has been a point of philosophical contention and from one point of view its application is seen to be somewhat *ad hoc*. Nonetheless, it is a valuable idea that can serve as a foundation for accounts that afford a way to realistically talk about information (and knowledge). As long as the relevant alternatives template is sound, I do not consider associated vagueness or flexible boundaries in determining a set of relevant alternatives a detraction.

One piece of support for this strategy of employing channel conditions/relevant alternatives can be found in analysing the application of absolute concepts. The concept of information here, like the concept of knowledge to which it will be applied, is absolute; *A* either carries the information that *B* or it doesn't. Now, we legitimately apply absolute concepts all the time, even though at some level they might fail. For example, we might say that an apple box is empty because it contains no apples, even though in some sense it is not empty since it still contains dust and air molecules. Of course, when attributing emptiness to the box, we rightly do not include dust and air molecules in our consideration; they are 'irrelevant alternatives' as it were. In this way, we can invoke contexts and legitimately apply absolute concepts in a flexible way. These ideas of fixed channel conditions and relevant alternatives will be discussed further in Section 4.4.3.

It is important to note that a signal's informational content is not unique. Generally speaking, there is no single piece of information in a signal or structure. For example, anything that carries the information that *x* is the number 7 also carries the information that it is an odd number, that it is a prime number, that it is not an even number and so on. Or to use an instance of the transitivity of information, if the sound reaching one's ears

⁴Though perhaps falsificationist tendencies would lead one to say that given a situation in which it is known that metal is present, the only signal that can carry information about the working condition of the detector is its silence, which would carry the information that it is not working.

carries the information that the radio is on, and the information that the radio is on carries the information that the power is on, then the sound reaching one's ears also carries the information that the power is on; one piece of information is *nested* in another.

This non-uniqueness is one characteristic of information that distinguishes it from meaning; if I say that x is the number 7, I mean just that. Also, whilst the meaning of a signal is independent of its truth, its information is not. Thus information and meaning are two distinct things and it is important that the two are not confused.

This account of information captures a vast range of situation types involving information flow. As noted, a crucial point is that the information of a signal is a result of the regularity within a system and the nomic relations the signal bears to other conditions. These relations can be considered 'counter-factual' supporting relations, expressible with something like the following:

x 's being F carries information about y 's being G = if y were not G , then x would not have been F ⁵

To use one of Dretske's examples [52, p. 58], the reason why the thermometer in Person A's room carries information about the temperature in their room (that it is 72°) but not about the temperature in Person B's room, even though both rooms are at the same temperature, is that given its location the registration of the thermometer is such that it would not read 72° unless Person A's room was at this temperature. This counterfactual is not supported in the case of Person B's room; the reading on A's thermometer is statistically independent of the temperature in Person B's room. Even MTC tells us that there has to be a statistical dependence between a signal and an event if we are to speak of the former carrying information about the latter.

Given this counterfactual nature, one easy place to find examples of information flow is in situations involving causal connections. The pressing of a button causes the doorbell to ring and the doorbell ringing carries the information that the doorbell was pressed. But whilst information flow can be present in cases of causation, causation is neither necessary nor sufficient for the presence of information flow.

A causal connection is insufficient because the system can involve a many-one relationship, whereas the presence of an information connection requires a one-one or one-many relationship between cause and effect(s). Consider a system involving a web server, Server

⁵As we will see in Section 4.3, whilst this definition captures the idea of an information relation, it is doubtful that the information relation can be represented with a true logic of counterfactuals. Rather, this definition should be taken as expressing a contraposition of sorts.

1, and a receiving computer, Computer 1, connected to the internet. The downloading of a page by Computer 1 carries the information that Server 1 served the page. Consider now the addition of another server to the network, Server 2, which is added to ease the load off Server 1. When a request is made to the server system for a web page, a random algorithm selects which of the servers will serve the web page in that instance. If Computer 1 were to now make a request for a web page and Server 1 served the page, the receiving of the web page by Computer 1 would no longer carry the information that Server 1 served the web page. This is because there are now two possible causes for the serving of the web page. The only information that this signal would carry was that either Server 1 or Server 2 served the webpage.

A causal connection is also unnecessary, or at least the presence of information flow between A and B does not require a direct causal connection between A and B . Consider a situation involving Server 1 and two receivers, Computer 1 and Computer 2, where both computers are receiving a live webcast from the server. Although there is no causal connection between Computer 1 and Computer 2, the fact that Computer 1 has just received the webcast carries the information that Computer 2 has also received the webcast. So in cases like this, where A causes B and A causes C , then B carries the information that C and C carries the information that B .

This cannot be extended to the generalisation that if B carries the information that A because A causes B and C carries the information that A because A causes C , then B carries the information that C and vice versa. If A is such that it can either just cause B , just cause C or cause both simultaneously, then the occurrences of B and C are not correlated and they do not carry information about each other.

How far Dretske goes and how far one could go in defining the boundary for cases of information flow is an open matter. Dretske seems to insist on the necessity of nomic correlations for information flow, as indicated by the following passage:

The same point is, perhaps, easier to appreciate if one thinks of two communication systems: A - B and C - D . A transmits to B and C transmits to D . Quite by chance, and at exactly the same time, A pecks out exactly the same message (sequence of dots and dashes) to B that C does to D . Assuming that no other messages are ever transmitted, and assuming that the channels are perfectly reliable, there is a perfect correlation (over all time) between what A transmits and what D receives. Yet, despite this correlation, D receives *no* information from A . To make this even more vivid, we could imagine system A - B separated by hundreds of light-years from system C - D . Physically, there is no basis for com-

munication. Still, the correlation between what transpires at A and D is perfect. The conclusion is inescapable that perfect correlation does not suffice for the transmission of information. [51, p. 74]

So according to Dretske, it is the *nomic dependence* between A and B that determines information flow. Since there is no nomic dependence between A and D , there is no information flow. I am not sure so about this. With a more accommodating account I have in mind, perfect and persistent correlation can suffice for information flow. Not only need there not be an appeal to causation for the presence of information flow, there need not be a direct *nomic dependence* at all; sufficient covariation is enough.

It is not hard to imagine two separate events, with different causes, which happen to always occur at the same time. Someone starts a stopwatch (Stopwatch 1) at time t to time an event. It also happens that nearby someone else starts another stopwatch (Stopwatch 2) at time t to time a completely separate event. Both of these events last for five minutes. The signal that Stopwatch 1 has kept time for 1 minute also carries the information that Stopwatch 2 has kept time for 1 minute. Whilst there is no causal link or nomic dependence involved between the two stopwatches, the laws of the regular physical system they are within ensure a perfect and persistent correlation between them.

Of course, to ‘tap into’ this information relation one would need to be attuned to the correlation. But the existence of this objective correlation/information is independent of any potential informee and thus independent of any common element other than the correlation between the watches. To put it simply, we can say that A carries the information that B if and only if every time A happens B happens. Just what ‘every time’ means depends on the set of relevant alternatives it ranges over. In the stopwatch example above, yes it could have been the case that one of the watches was never started, or that they were started at different times. But given that they were started at the same time, if someone is aware of the situation, we can fairly say that Stopwatch 1 will consistently and reliably tell them something about Stopwatch 2; there is an information connection between Stopwatch 1 and Stopwatch 2. We can satisfy the required technical details for this by circumscribing the set of relevant alternatives to include all and only those moments that follow for the next five minutes from the start. This ensures that for each time x , $\Pr(\text{Stopwatch 2 shows } x | \text{Stopwatch 1 shows } x) = 1$.

Concerns related to this limited purview of Dretske’s definition are raised by Foley [78]. If perfectly correlated events which are neither nomologically nor logically related are to not have an information relation between them, then according to Dretske’s definition they should not have a probability of 1 with respect to one another. If this is the case, then what kind of probability will generate these results for Dretske [78, p. 180]?

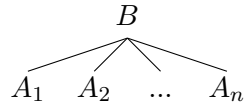


Figure 4.1: Depiction of information flow relationship structure

Another consideration concerns the arrow of information flow. Given the specific use of inverse conditional probabilities, it seems that with accounts such as Dretske's, if A carries the information that B , then B is prior to or concurrent with A . Is it possible for A to carry the information that B if A precedes B ? Whilst for the purposes of this work it will suffice to focus on inverse information relations, expanding the domain to include cases where A (present) carries the information that B (future) is something to consider. Some discussion on the matter can be found in Appendix C Section C.2.

In closing this section, it is worth mentioning a way to think about information flow that will facilitate the investigation that follows. A_1 carries the information that B if and only if there is a one-one relationship between A_1 and B or a one-many relationship between $\{A_1, A_2, \dots, A_n\}$ and B , as depicted in Figure 4.1.

If this unique traceability from A_1 to B is lost, then so is (at least some of) the information. To take a small but good example from logic, consider a negation operator with the following truth table:

A	$\neg A$
1	0
*	1
0	1

Given a proposition p with a truth value of 1, the result of applying the $\neg$ operator is a proposition with a truth value of 0. In this case, the information that the input is 1 has flowed through the operation, the evaluation of $v(\neg p) = 0$ carries the information that $v(p) = 1$. If however, p has a truth value of 0, the result of applying $\neg$ results in a loss of information. At the start we have the information that the input is 0. After the operation, we have the information that the input was 0 or *. This is an example of how information can be lost in transmission or after an operation. It is also an example of the fact that information is lost when the set of representing signals (in this case 0 and 1) becomes less than the set of possible outcomes (in this case 1, 0 and *).

Sometimes a similar thing happens when signals have a semantic significance and result in the recording of some semantic information which is subsequently corrupted. Suppose that you are looking at a correctly functioning digital clock which is displaying '8:00 AM'.

The clock is thus carrying the environmental information that it is 8 a.m.. You use this information to store in a computer the time at which a particular event E occurred. This datum in the computer subsequently gets corrupted so that ‘8:00 AM’ becomes ‘8:00’. In the language of 12hr clocks, the ‘AM’ is crucial and the information that E occurred at 8 a.m. is consequently lost; the datum does though still give the information that E occurred at 8 a.m. or 8 p.m.. However, in a 24hr system the information that E occurred at 8 a.m. would not be lost. In a 24hr system, ‘AM’ is either redundant (in the case that it is *ante meridiem*) or invalid (in the case that it is *post meridiem*).

4.1.1 Dretske’s Account and Properties of Information Flow

It is now time to introduce a logical vocabulary for speaking about information flow and list some of the properties which we will consider throughout our investigation. This is just a vocabulary and will be used in general when discussing the specific proposals that follow. It consists of the following:

- Symbols $A, B, C, D, \dots$ stand for information bearing structures (events / situations / facts / signals).
- The formula $A \sqsupset B$ stands for ‘ A carries the information that B ’.
- $\vdash$ is a consequence relation that relates such formulas
- Whilst in some of the proposals we look at it would be possible to nest information statements, in this vocabulary $\sqsupset$ is treated as a meta-connective. The symbols $A, B, C, D, \dots$ are taken to be statements in the propositional calculus and do not contain $\sqsupset$.

To demonstrate this terminology, take the truism that if A carries the information that B , then A carries the information that B , which is represented as:

$$A \sqsupset B \vdash A \sqsupset B$$

The main properties of information flow which we will look at are:

1. **Conjunction:** - $(A \sqsupset B) \wedge (A \sqsupset C) \vdash A \sqsupset (B \wedge C)$
2. **Reverse-Conjunction:** - $A \sqsupset (B \wedge C) \vdash (A \sqsupset B) \wedge (A \sqsupset C)$

3. **Disjunction:** - $(A \sqsupset C) \wedge (B \sqsupset C) \vdash (A \vee B) \sqsupset C$
4. **Reverse-Disjunction:** - $(A \vee B) \sqsupset C \vdash (A \sqsupset C) \wedge (B \sqsupset C)$
5. **Transitivity:** - $(A \sqsupset B) \wedge (B \sqsupset C) \vdash (A \sqsupset C)$
6. **Contraposition:** - $A \sqsupset B \vdash \neg B \sqsupset \neg A$
7. **Monotonicity:** - $A \sqsupset B \vdash A \wedge C \sqsupset B$
8. **Veridicality:** - $A, A \sqsupset B \vdash B$

Let us take a simpler version of Dretske's definition of information by removing the elements of anti-vacuity and the agent's background knowledge, to get A carries the information that B if and only if $\Pr(B|A) = 1$. It is this simpler, absolute form of correlation that will generally be used when investigating these properties for accounts of information flow. This will allow us to investigate these properties against the core logic of an account, rather than being affected by removed extra factors. For as will be shown a little further on, some of these principles, which hold using this 'pure definition', would strictly fail to hold using Dretske's full definition (due to anti-vacuity or k).

This stripped down definition of information flow satisfies all of the above conditions.⁶ Here are demonstrations for the Conjunction Principle and Transitivity.

Theorem 4.1.1. Conjunction Principle: if $\Pr(B|A) = 1$ and $\Pr(C|A) = 1$ then $\Pr(B \wedge C|A) = 1$

Proof. Since, $\Pr(B|A) = 1$, $\Pr(B \wedge A) = \Pr(A)$. Similarly, $\Pr(C \wedge A) = \Pr(A)$.

We need to show that $\Pr(B \wedge C \wedge A) = \Pr(A)$.

Firstly rule out $\Pr(B \wedge C \wedge A) > \Pr(A)$, which is not possible according to the laws of probability.

Secondly, rule out $\Pr(B \wedge C \wedge A) < \Pr(A)$:

$\Pr(A) = \Pr(B \wedge A) = \Pr(C \wedge A)$, so $\Pr(B \wedge C \wedge A) < \Pr(A)$ implies that $\Pr(B \wedge C \wedge A) < \Pr(B \wedge A)$ and $\Pr(B \wedge C \wedge A) < \Pr(C \wedge A)$. But as the following demonstrates, either of $\Pr(B \wedge C \wedge A) < \Pr(B \wedge A)$ or $\Pr(B \wedge C \wedge A) < \Pr(C \wedge A)$ result in a contradiction:

⁶Technically there are some exceptions involving formulas with a probability of zero and hence undefined calculations. With monotonicity, $\Pr(p|p) = 1$ [$p \sqsupset p$] but $\Pr(p|p \wedge \neg p)$ is undefined [not $(p \wedge \neg p) \sqsupset p$]. With Reverse Disjunction, $\Pr(p|p \vee (q \wedge \neg q)) = 1$ [$(p \vee (q \wedge \neg q)) \sqsupset p$] but $\Pr(p|q \wedge \neg q)$ is undefined [not $(q \wedge \neg q) \sqsupset p$].

$$\begin{aligned}
\Pr(B \wedge C \wedge A) &< \Pr(C \wedge A) \\
\Pr(B \wedge A) + \Pr(C) - \Pr((B \wedge A) \vee C) &< \Pr(C \wedge A) \\
\Pr(A) + \Pr(C) - \Pr((B \wedge A) \vee C) &< \Pr(C \wedge A) \\
\Pr(A \vee C) &< \Pr((B \wedge A) \vee C) \\
\Pr(A \vee C) &< \Pr((B \vee C) \wedge (A \vee C)) \\
&\text{but} \\
\Pr(A \vee C) &\geq \Pr((B \vee C) \wedge (A \vee C)) \\
&\therefore \perp
\end{aligned}$$

□

Theorem 4.1.2. Transitivity: if $\Pr(B|A) = 1$ and $\Pr(C|B) = 1$ then $\Pr(C|A) = 1$

Proof. In terms of sets:

- since $\Pr(B|A) = 1$, $A \subseteq B$
- since $\Pr(C|B) = 1$, $B \subseteq C$
- $\therefore A \subseteq C$ and $\Pr(C|A) = 1$

□

4.2 Probabilistic Information

In *Information Without Truth* [162], Scarantino and Piccinini reject the Veridicality Thesis for both natural (environmental) and non-natural (semantic) information⁷. The Veridicality Thesis for Natural Information (VT_N) is stated as:

(VT_N) If a signal s being F carries natural information about an object o being G , then o is G .

Contrary to this, Scarantino and Piccinini endorse a Probability Raising Thesis for Natural Information (PRT_N), where the transmission of natural information involves nothing more than the truth of the following probabilistic claim:

⁷Some of their discussion on the latter was covered in Section 1.4 of Chapter 1.

(PRT_N) If a signal s being F carries natural information about an object o being G , then

$$\Pr(Go|Fs) > \Pr(Go|\neg Fs) \text{ }^8$$

It is perhaps clearer to reformulate this claim as the equivalent:⁹

If a signal s being F carries natural information about an object o being G , then

$$\Pr(Go|Fs) > \Pr(Go)$$

According to this definition, if A carries information about B then it need not be the case that B . What are the repercussions of such a claim? It is firstly important to note how this definition is worded, with *about* being an operative word. Although PRT_N supports the claim that information can be transmitted without truth, this position does not really threaten the Dretskean sense in which information entails truth. As Scarantino and Piccinini mention:

Even though Dretske's theory of information focuses exclusively on all-or-nothing information, Dretske was not oblivious to the fact that natural information can come in degrees. He wrote that "[i]nformation about s comes in degrees. But the information that s is F does not come in degrees. It is an all or nothing affair". Thus, although Dretske's emphasis on all-or-nothing information may appear to commit him to the Veridicality Thesis for Natural Information (VTN), he actually endorsed only the weaker thesis that if a signal s being in state F carries *the* information *that* o is G , then o is G . We too endorse this weaker thesis. [162, p. 318]

So in order for a signal to carry the information *that* o is G , it must be the case that o is G . But it need not be the case that o is G in order for a signal to carry information *about* o being G . The motivations behind this position are understandable. With probabilistic theories of information, signals carry natural information about anything they reliably correlate with. Yet these correlations are seldom perfect and a typical signal might sometimes indicate something even though that something has not occurred. But whilst a typical signal might

⁸It should be noted that they footnote this definition with the following: "one of us has argued that signal s being F can also carry (negative) natural information about o being G by lowering the probability that o is G (Scarantino unpublished). We disregard this complication in what follows". Presumably a more general definition for PRT_N would look something like:

If a signal s being F carries natural information about an object o being G , then $\Pr(Go) \neq \Pr(Go|Fs)$

⁹See Appendix C.1 for a proof of this equivalence

not guarantee the occurrence of that thing which it indicates, it does make its occurrence more likely.

Unlike all-or-nothing information, probabilistic natural information comes in degrees. For instance, if o being G is much more probable given the signal constituted by s being F than it would have been without the signal, s being F carries a lot of probabilistic natural information about o being G . If, instead, the probability of o being G increases only marginally given the signal, s being F carries a limited amount of probabilistic natural information about o being G . [162, p. 318]

Whilst I have no conclusive position on whether or not probability raising relations suffice to support genuine environmental information transmission, I do have some thoughts on the matter. As noted elsewhere, “in ordinary information-theoretic terms, if the occurrence of a signal raises or lowers the probability of a proposition, that is enough for the signal to count as bearing information” [179, p. 4]. But if we are concerned about propositional informational content and the flow of environmental/natural information rather than the sense of information dealt with by MTC, we might say that whilst PRT_N deals with statistical interrelations between source and signal, for environmental/natural information to flow between source and signal, that with which the signal is correlated must be true at the source. Recall a similar requirement for Gricean natural meaning, discussed in Section 1.3.1 of Chapter 1: ‘ x means that p ’ entail the truth of p . The point here is that there might be confusion between information as a change in probabilities and propositional information about something. There could be a case for saying that whilst a probability raising signal might carry information as a statistical correlation in the technical sense of MTC, there is a category of environmental/natural information that it does not qualify for. Unlike Dretske’s definition, this definition does not give a form of propositional information about something.

One first thing to note is that on a standard reading, the phrase ‘ x carries natural information about y ’, could be construed so that x is a signal, y is an object and x carries the information that $B_1, B_2, \dots, B_n$, where each B is a property of y . So such cases would reduce to the signal providing a collection of *information-that* about the source.

Granted this is not how Scarantino and Piccinini are using the term, let us further test their account with the following example. There are four equiprobable outcomes A, B, C and D , associated with the codes ‘00’, ‘01’, ‘10’ and ‘11’ respectively. You receive the message ‘1’, thereby receiving 1 bit of information in accordance with MTC. In terms of Dretske’s information-that, you also receive the information that $C \vee D$. But do you receive natural information about C ? Even if D is the actual outcome, I would say that in a sense you

do receive informational content about C . But although this information about C is not tied to or dependent upon the occurrence of C , it might be a mistake to appeal to C 's non-occurrence as a relevant absence of truth. For when we receive the message '1', perhaps all we can say regarding natural informational content concerning C are things like 'the signal '1' carries the environmental information that the probability of C is 0.5'. Or if we had a situation where $\Pr(A) = \Pr(C)$, $\Pr(A) = 2\Pr(B)$ and $\Pr(C) = 2\Pr(D)$, we would be able to say that 'the signal '1' carries the environmental information that the outcomes is probably C '. So ultimately, there is some truth that such information-about claims are linked to and dependent upon.

In light of this, the 'information about' associated with PRT_N could be reduced to 'information that' and PRT_N could be translated to the following claim:

If a signal s being F carries natural information about an object o being G , then s being F carries the information *that* o being G is more probable.

In fact, this idea can be extracted from a quote due to Scarantino and Piccinini:

Consider the signal that the doorbell is ringing. In order for the ringing doorbell to carry the natural information that a visitor is at the door, a visitor must be at the door. This seems right. But it does not follow that a ringing doorbell carries no natural information about there being a visitor at the door in the absence of a visitor at the door. If the ringing doorbell is a standard one in standard conditions, *it carries the natural information that a visitor is probably at the door*. It is on account of this information, the only natural information a standard doorbell manages to carry about visitors, that we reach for the door and open it. [162, p. 318] (italics mine)

Irrespective of whether or not this conception of *information about* constitutes a genuine type of environmental information carriage, this notion of information will not suffice for the purpose of developing an informational basis for knowledge, for reasons outlined in Section 4.1. As can be gathered from that outline, the information without truth is not the problem. If it were, it could be dealt with by simply adding a truth condition to the definition:

If a signal s being F carries natural information about an object o being G , then

$$(\Pr(Go|Fs) > \Pr(Go)) \wedge Go$$

This could lead to accounts where a conditional probability above some threshold extremely close to yet below one still might suffice to generate information with knowledge creation potential. As long as the resulting belief was true, if it was formed by a signal which almost guaranteed something was true, then it could count as knowledge. But the main issue is that probability don't satisfy certain sought after properties, as noted also by Dretske when he suggests this possibility [51, p. 66].

In developing an informational account of knowledge, we want to be able to say things like if one knows that A and they know that B , then they know that $A \wedge B$. Such a probabilistic account fails most of the information flow conditions we have looked at. To begin with, the conjunction principle fails.

Example 4.3

Take a domain consisting of three predicates $\{P, Q, R\}$ and five objects $\{a, b, c, d, e\}$, with the following assignments:

- $Pa \wedge Qa \wedge \neg Ra$
- $Pb \wedge \neg Qb \wedge Rb$
- $\neg Pc \wedge \neg Qc \wedge \neg Rc$
- $\neg Pd \wedge Qd \wedge Rd$
- $\neg Pe \wedge \neg Qe \wedge \neg Re$

Given this:

- $\Pr(Qx|Px) = \frac{1}{2} > \Pr(Qx) = \frac{2}{5}$
- $\Pr(Rx|Px) = \frac{1}{2} > \Pr(Rx) = \frac{2}{5}$
- $\Pr(Qx \wedge Rx|Px) = 0 < \Pr(Qx \wedge Rx) = \frac{1}{5}$

□

So even though Px would carry information about Qx and Px would carry information about Rx according to this definition, it does not carry information about $Qx \wedge Rx$, which

is itself impossible given Px . In fact, Px carries information about $\neg(Qx \wedge Rx)$ instead: $\Pr(\neg(Qx \wedge Rx)|Px) = 1 > \Pr(\neg(Qx \wedge Rx)) = \frac{4}{5}$. It is a somewhat peculiar feature of this account of natural information carriage that the following three statements are not inconsistent:

1. A carries information about B
2. A carries information about C
3. A carries information about $\neg(B \wedge C)$

The following conditions also all fail to hold for this probabilistic relation:

- Reverse Conjunction
- Disjunction
- Reverse Disjunction
- Transitivity
- Monotonicity

Interestingly however, Contraposition is one principle that is satisfied, as demonstrated with the following.

Theorem 4.2.1. $\Pr(B|A) > \Pr(B) \Rightarrow \Pr(\neg A|\neg B) > \Pr(\neg A)$

Proof. Showing that $\Pr(\neg A|\neg B) > \Pr(\neg A)$ is equivalent to showing that $\Pr(\neg B \wedge \neg A) > \Pr(\neg B)\Pr(\neg A)$, which we now proceed to do. We also require the following fact: $\Pr(B|A) > \Pr(B)$ so $\Pr(A \wedge B) > \Pr(A)\Pr(B)$ (Fact 1).

$$\begin{array}{llll}
 \Pr(\neg B \wedge \neg A) & > & \Pr(\neg B)\Pr(\neg A) & \Leftrightarrow \\
 \Pr(\neg(B \vee A)) & > & (1 - \Pr(A))(1 - \Pr(B)) & \Leftrightarrow \\
 1 - \Pr(B \vee A) & > & (1 - \Pr(A))(1 - \Pr(B)) & \Leftrightarrow \\
 1 - \Pr(B \vee A) & > & 1 - \Pr(B) - \Pr(A) + \Pr(A)\Pr(B) & \Leftrightarrow \\
 \Pr(B) + \Pr(A) - \Pr(B \vee A) & > & \Pr(A)\Pr(B) & \Leftrightarrow \\
 \Pr(A \wedge B) & > & \Pr(A)\Pr(B) & \text{Fact 1}
 \end{array}$$

□

One of the apparent advantages of a general probabilistic account of information over Dretske's is its flexibility. As we saw, the strictness of Dretske's requirement that information flow depends on a conditional probability of one seems to extremely minimise the amount of information flow in the world, since such perfect correlations are rare. Instead of confining the threshold to one, by using a larger portion of the probability scale we could reflect the imperfections and account for the lack of complete certainty in the world.

But conditional probabilities of one are only rare in an absolute sense. As discussed earlier and will be explored further on, we can judge a conditional probability of one relative to a fixed set of channel conditions and relevant alternatives. How exactly a set of relevant alternatives is determined is itself not a straightforward manner. But what is important is that at least this idea provides a flexible framework that affords gradation and a variable range.

This preference for the variability of the relevant alternatives framework (RAF) rather than the variability of a general probabilistic account has other important consequences and advantages. Unlike RAF and as we have seen, a general probabilistic account of information fails to satisfy certain properties we are interested in.

This point can be seen as part of the general issue concerning probability and topics such as epistemology and justification, of which the lottery paradox is a classic example. When basic probability is used, many standard properties of logic break down. Whilst a wealth of literature exists on the topic of developing accounts of probabilistic criteria that satisfy certain logical principles, these attempts are generally confined to a specific domain. At any rate, the failure of such a general probabilistic account of information to satisfy such a large range of conditions which intuitively hold for knowledge imply that such an account of information is inadequate to serve as a bedrock for the informational epistemology we aim for. Thus we shall continue with our investigation into accounts that are logically stronger and in which information flow implies truth.

4.3 A Counterfactual Theory of Information

In a collection of commentaries on *Knowledge and the Flow of Information*, Barry Loewer [126] points out that none of the usual interpretations of probability supports Dretske's usage of inverse conditional probability in his definition of information. A subjective degree of belief interpretation will not do, since Dretske's aim is to analyse belief and knowledge in terms of objective, physicalistic notions of informational content. Since frequency interpretations posit that the probability of an event is its relative frequency of occurrence over a number

of repeats, they don't seem suitable, because Dretske deals with events that are not the outcome of repeats. In any case, Dretske rules out a relative frequency interpretation [51, p. 245]. Loewer suggests the possibility of employing a propensity interpretation, which addresses the issue of probability assignments to non-repeatable events. However, the inverse probability $\Pr(s|r)$ is problematic on a propensity interpretation, since there will be no propensity $\Pr(s)$.¹⁰ Loewer then offers a reformulation of Dretske's idea without reference to probability:

r 's being R carries the information that s is F iff r is R , and if r is R then s must have been (or must be) F .

He goes on to write:

“If r is R then s must have been F ” is what Lewis (1979) calls a “backtracking conditional”. Truth conditions for these, as for other conditionals, are (approximately) “there are laws L , conditions C which are *cotenable* with $R(r)$ such that $L \& C \& R(r)$ imply $F(s)$ ”. This account dovetails with Dretske's view that the information carried by a signal is relative to “channel conditions”. These are conditions that are thought of as fixed in a given situation. My suggestion is that the conditions cotenable with $R(r)$ are the channel conditions. [126, p. 76]

Whilst this nomic account does seem like a reasonable alternative approach to Dretske's, its appeal ultimately depends on how exactly it is formally realised.

Cohen and Meskin (C&M) [36] relate to some of Loewer's considerations. They offer an alternative theory of information that differs to Dretske's and similar traditional accounts in two main respects. Firstly, it explains information in terms of counterfactuals rather than conditional probabilities. Secondly, the definition they provide does not make reference to an agent's background knowledge and bypasses doxastic states. For C&M:

(S*) Information relations are constituted by the non-vacuous truth of counterfactuals connecting the informational relata. Thus, x 's being F carries information about y 's being G if and only if the counterfactual conditional ‘if y were not G , then x would not have been F ’ is non-vacuously true.^{11 12}

¹⁰For an outline of interpretations of probability, including a discussion of ‘Humphreys paradox’, concerning propensities and probability, see [108].

¹¹This use of ‘about’ is not to be confused with that of the probabilistic account in the previous section.

¹²They also consider the following weaker claim, which “expresses the view that certain counterfactuals can

Like the ‘but, given k alone, less than 1’ component in Dretske’s definition, the non-vacuity component in (S*) rules out cases where y ’s being G is necessary.

In a footnote C&M mention that x ’s being F and y ’s being G are construed as actual events, so one event carries information about a second only if they are actual. Given the significance of this addendum, it is surprising that they relegated it to a footnote. With this qualification given, we can formally state C&M’s definition as follows, where $>$ stands for the counterfactual conditional:

$$Fx \sqsubset Gy =_{df} Fx \wedge Gy \wedge (\neg Gy > \neg Fx)$$

Whilst the requirement that the events be actual ensures veridicality, it also means that the power to express the range of information relations for potential events is lost. For example, according to this definition the statement ‘Sally’s presence in Germany carries the information that Sally is in Europe’ does not hold given that Sally is in Australia. As we will show later on in this section, removing the requirement that the events be actual creates problems of its own.

A first thing to note is that since this counterfactual account of information makes no reference to probability, it avoids the interpretation issues a probabilistic account such as Dretske’s faces. Perhaps this is an advantage, but not a resounding one. Demir [43, p. 58], who endorses a probabilistic approach, supports the prospect that some account of probability will provide a suitable interpretation and offers some suggestions.

He also defends Shannon’s MTC against the claim made by C&M that it also runs into the probability interpretation problem [43, p. 49]. Shannon defines the mutual information between s and r as:¹³

$$I(s, r) = - \sum_{i=1}^n \Pr(s_i) \log_2 \Pr(s_i) + \sum_{j=1}^n \Pr(s_i | r_j) \log_2 \Pr(s_i | r_j) \quad (4.4)$$

What C&M fail to realise though is that mutual information is commutative:

provide a handy diagnostic test for the existence of an informational relation”:

(W*) x ’s being F carries information about y ’s being G if the counterfactual conditional ‘if y were not G , then x would not have been F ’ is non-vacuously true.

The difference between (W*) and (S*) prompts consideration of the possibility that information flow might result from more than one type of relationship; so not ‘information flow if and only if relationship holds’, but ‘if relationship1 holds then information flow, if relationship2 holds then information flow, ...’.

¹³This is another way of saying that information received = information sent - equivocation.

$$I(s, r) = I(r, s) = - \sum_{j=1}^n \Pr(r_j) \log_2 \Pr(r_j) + \sum_{i=1}^n \Pr(r_j | s_i) \log_2 \Pr(r_j | s_i) \quad (4.5)$$

So the value of $I(s, r)$ can be calculated without using inverse conditional probabilities, something that C&M subsequently concede [37, p. 637].

Whilst I can see no evidence of such a suggestion in the literature, since it is generally possible to calculate inverse conditional probabilities from conditional probabilities and vice versa, cannot Dretske's definition also be formulated in terms of forward conditional probabilities? Would not the following do the job?

$$\Pr(s|r) = \frac{\Pr(r|s)\Pr(s)}{\Pr(r)}$$

At any rate, as will be discussed later on, I think that an account functionally equivalent to Dretske's account, which does not make use of probabilities, can be given.

C&M discuss several other points in comparing their counterfactual approach with Dretske's account. It is claimed that whilst Dretske's account relies on natural laws to explain informational relations, their counterfactual approach does not and therefore has a less expensive ontology. I am not so sure that this is an outright advantage. Dretske does indicate that the conditional probability of one in his definition is a result of nomic dependence:

In saying that the conditional probability (given r) of s 's being F is 1, I mean to be saying that there is a nomic (lawful) regularity between these event types, a regularity which *nominally precludes* r 's occurrence when s is not F . [51, p. 245]

But there is no clear reason not to have an account of information which appeals to natural laws or nomic dependencies. Furthermore, whilst Dretske chooses to invoke natural laws or nomic dependencies, the formal apparatus of his account remains viable independently of this interpretation.

Regarding properties of information flow, a difference between the counterfactual account and Dretske's is that whilst the latter validates the Xerox principle, the former does not, given that transitivity is not a counterfactual validity; information flow is neither transitive nor intransitive but non-transitive. But whilst C&M suggest that "it is in fact an advantage of the counterfactual account that it fails to licence the Xerox principle" [36, p. 7], they do nothing to show how the failure of transitivity for counterfactuals translates to a plausible account of failure for information transitivity.

Another difference between the two accounts which it is claimed is an advantage for the counterfactual theory concerns reference to doxastic states. The counterfactual theory contains no doxastic references whilst Dretske's probabilistic definition contains reference to an agent's background knowledge via the k variable. Since the former account does not make essential reference to doxastic states of subjects, it "allows for the sort of objective, reductive explanations of notions in epistemology and philosophy of mind that many have wanted from an account of information" [36, p. 333]. With Dretske's account on the other hand, there is the concern that the element of background knowledge makes his definition of information non-naturalistic and his definition of knowledge circular. I deem this a non-issue. As Dretske suggests, reference to k can be eliminated by backwards recursion in these recursive definitions of information and knowledge.¹⁴

Scarantino [161] questions C&M's attempt to offer a *non-doxastic* theory of information. He argues that eliminating reference to an agent's background knowledge results in instances of information transmission wrongly being classed as not being instances of information transmission and a sacrificing of the explanatory power of the information concepts.¹⁵ Take Dretske's Shell Game example:

[S]uppose that there are four shells and a peanut is located under one of them. In attempting to find under which shell the peanut is located, I [Person A] turn over shells 1 and 2 and discover them to be empty. At this point you [Person B] arrive on the scene and join the investigation. You are *not* told about my previous discoveries. We turn over shell 3 and find it empty. How much information do you receive from this observation? How much do I receive? Do I receive information that you do not receive? [51, p. 78]

Let K_a and K_b stand for the knowledge states of the recipients A and B prior to the turning over of shell 3. K_a includes the knowledge that shells 1 and 2 are empty and that the peanut is either under shell 3 or shell 4. K_b has no knowledge beyond the starting fact that the peanut could be under any one of the four shells. The signal in question is the emptiness of shell 3 upon its turning over. What information does this signal carry?

With Dretske's account, relative to K_a the signal carries the information that the peanut is under shell 4. Relative to K_b , the signal carries the information that the peanut is under shell 1, shell 2 or shell 4. The signal also carries the information that the peanut is not under shell 3 relative to both K_a and K_b .

¹⁴The case for knowledge will be covered in the next chapter. Demir [43, p. 55] discusses the definition of information.

¹⁵The importance of this explanatory affordance was not lost on Grice either: "Those spots didn't mean anything to me, but to the doctor they meant measles." [90, p. 377]

With C&M's account, things are different. Using the definition they provided, we can substitute the signal of shell 3's emptiness to say that the emptiness of shell 3 carries information about any event 'y is G' such that 'if y were not G, shell 3 would not be empty' is non-vacuously true. Given this yardstick, the signal continues to carry information that the peanut is not under shell 3. But does the signal carry the information that the peanut is under shell 4? It would be so if the counterfactual:

If the peanut had not been under shell 4, shell 3 would not have been empty.

was true. Whilst C&M provide no clear way to work this out, it would seem that under a legitimate account of counterfactuals this conditional is false. So the emptiness of shell 3 does not carry information about the peanut's not being under shell 4. All it carries is the information that the peanut is not under shell 3; not much at all. This is counterintuitive, and discords with an ordinary conception of information, where the information a signal carries for some agent can be identified with what the agent can learn from it.

However, apart from this discordance, Scarantino raises another reason why "informational semanticists should be suspicious of non-doxastic accounts of information" [161, p. 632]. C&M suggest that the relativisation of information to background knowledge is at best a reflection of our 'ordinary ways of thinking about information'. On the other hand, a 'technical notion of information' which is to be used for reductive aspirations in a naturalised epistemology, need not accommodate this.

However this may be, what is ultimately at stake is the extent to which a technical notion of information can play a causal role in the explanation of behaviour. As Scarantino writes, if Person A bets a vast sum of money on the peanut being under shell 4 but Person B bets refrains from betting, Dretske's account can explain this by appealing to the difference in information received from the emptiness of shell 3. This explanatory option does not seem possible with C&M's theory, according to which the emptiness of shell 3 carries the same information for Person A and Person B.

I do not think this doxastic versus non-doxastic issue a significant one. Although information is something that is out there in the world, our treatment of the notion of information needs to incorporate the fact that the information an agent can get from a signal depends on their informational/doxastic/epistemic states. This relativisation does not undermine the essential objectivity of the information and as mentioned earlier a simple division resulting in the terms 'relative information' and 'absolute information' can provide clarification here. In the shell game scenario, the absolute information from the signal is that shell 3 is empty, whilst the relative information is that shell 3 is empty and shell 4 has the peanut. To em-

phasise the essential objectivity of information one could even substitute the reference to k with something else. For example, consider again the shell game scenario. One could say that it is not the signal of an empty shell 3 plus k which provides the information that the peanut is under shell 4. Rather, it is the composite signal of ‘empty shell 1 *and* empty shell 2 *and* empty shell 3’ that carries the information that the peanut is under shell 4. So this collection of signals here would be treated as one entity, and it is this temporal-spanning entity that is the complete signal. This is one way of thinking about things that eschews reference to an agent’s background knowledge. One could even remove the k element from Dretske’s definition if they wanted to exclude reference to doxastic states of subjects and still have a functional definition. So ultimately I don’t think that Dretske’s reference to doxastic states of subjects is a problematic factor, but simply a way of accommodating an important aspect of the notion of information.

Aside from comparisons between counterfactual and probabilistic approaches, counterfactuals have their own philosophical issues, some of which are discussed by C&M [36, p. 15].

Furthermore, some exploration reveals that using the standard logic of counterfactuals, C&M’s definition gives some results that disagree with what it seems are fairly straightforward properties of information flow. For example, take the following: $A \sqsupset B \vdash A \sqsupset B \vee C$. According to C&M’s definition we have:

- $A \sqsupset B =_{df} A \wedge B \wedge (\neg B > \neg A)$
- $A \sqsupset (B \vee C) =_{df} A \wedge (B \vee C) \wedge (\neg(B \vee C) > \neg A)$

But since the monotonic inference $\neg B > \neg A \vdash (\neg B \wedge \neg C) > \neg A$ is not counterfactually valid, neither is the inference from $A \sqsupset B$ to $A \sqsupset (B \vee C)$.

Also, Demir [44] recently pointed a problematic consequence of the counterfactual theory of information that arguably makes it untenable. He shows that given the standard possible worlds account of counterfactuals, according to C&M’s definition “‘A carries information that B’ necessarily implies ‘A carries information that B and C’ for any C such that the closest not-C world is more remote than the closest not-B world”. So given that ‘Joe is in France’ (A) carries the information that ‘Joe is in Europe’ (B), it follows that ‘Joe is in France’ carries the information that ‘Joe is in Europe and Paris is the capital city of France’ ($B \wedge C$). The reasoning is as follows:

- $\neg B > \neg A$ (A carries the information that B. The closest $\neg B$ world is a $\neg A$ world)

- $\neg B \vee \neg C > \neg A$ (since $\neg C$ [Paris not being France's capital] is much more unlikely than Joe not being in Europe, the closest $\neg B$ world is a $\neg B \vee \neg C$ world)
- $\neg(B \wedge C) > \neg A$ (A carries the information that $B \wedge C$)

So we end up with the problematic $A \sqsupset (B \wedge C)$. Unless the Reverse-Conjunction principle fails, which would itself be highly problematic, it then follows that $A \sqsupset C$. But it is just wrong to say that Joe being in France carries the information that Paris is the capital city of France. Any way you look at it, these are indeed peculiar and problematic results. Such issues aside, it is now time to investigate the counterfactual arrow as a definition of the information-carriage relation.

4.3.1 The Logic of Counterfactuals and Information Flow Properties

If one is to adopt a counterfactual theory of information, then they are tasked with the selection of a logic to model counterfactuals. Where $>$ stands for the counterfactual conditional, in this section I would like to investigate the following definition:

$$A \text{ carries the information that } B =_{df} A > B$$

Since there are several approaches to and a variety of logics for the counterfactual conditional, the specific choice is not a straightforward one. In this investigation, as in the previous section, we employ the standard possible worlds approach, which is based on the idea that $A > B$ is true when B is true in the A -worlds (worlds where A is true) that are closest to the actual world in terms of some similarity relation. The systems fall under the group of *minimal change theories* [12].

Referring back to the logical properties listed in Section 4.1.1, we start by establishing that the Conjunction Principle is valid in counterfactual logics: $(A > B) \wedge (A > C) \vdash A > (B \wedge C)$. Even the basic normal conditional logic **CK** given by Chellas [32, 33] has the following simple axiomatisation in which it is an axiom:

- Axiomatisation of Propositional Calculus

- RCEA:
$$\frac{A \equiv B}{(A > C) \equiv (B > C)}$$

- RCEC:
$$\frac{A \equiv B}{(C > A) \equiv (C > B)}$$

- CM: $(A > (B \wedge C)) \supset ((A > B) \wedge (A > C))$
- CC: $((A > B) \wedge (A > C)) \supset (A > (B \wedge C))$
- CN: $A > \top$

Chellas gives this logic a possible worlds semantics in terms of standard conditional models and a selection function. One way to look at this construction is by associating a necessity operator $\Box_A$ for every formula A and defining $A > B$ as $\Box_A B$ over Kripke frames. The counterfactual logics based on sphere systems and similarity relations are proper extensions of this logic.¹⁶

The Converse Conjunction Principle is also valid in this logic, as is a host of other properties, including $A > (B \vee C) \wedge A > \neg B \vdash A > C$ and $A > (A > C) \vdash A > C$.

By imposing conditions on conditional frames, one can further strengthen the logic to get desired properties of information flow such as:

- $A > A$
- $A \wedge A > B \vdash A$
- $(A \vee B) > C \vdash (A > C) \wedge (B > C)$
- $(A > C) \wedge (B > C) \vdash (A \vee B) > C$ ¹⁷

As can be seen, given its possible worlds basis, this framework for conditional logic provides a fair bit of flexibility with regards to selecting valid properties. Of course, moving up the scale from lower to stronger systems, including those based on sphere models, increases the set of validities. But at a certain level, issues can arise. Although it is an open matter what the correct counterfactual logic is, perhaps it is a telling sign that there are unnegotiable issues with perhaps the two most prominent systems **C2** and **VC**, associated with counterfactual pioneers Stalnaker and Lewis respectively. As a result of their philosophical foundations, the systems developed by these founders of conditional logic contain certain idiosyncratic validities, which whilst provoking interesting philosophical debate on the nature of counterfactuals, are unequivocally unacceptable as principles of information flow.

Firstly, take Lewis' system, which validates the following:

¹⁶See [148], Chapter 5. for details.

¹⁷Demir [43, p. 52] incorrectly claims that C&M's account fails the Conjunction Principle. As we saw above, in the form of $(A > B) \wedge (A > C) \vdash A > (B \wedge C)$ the Conjunction Principle holds. In terms of C&M's definition (S*), the conjunction principle is $(\neg B > \neg A) \wedge (\neg C > \neg A) \vdash \neg(B \wedge C) > \neg A$. This validity is just Reverse-Disjunction, which is valid in standard counterfactual logic.

$$A \wedge B \vdash A > B$$

This would certainly be unacceptable as a principle of information flow. A coincidental or contingent co-occurrence of two events A and B in the actual world certainly does not imply that A carries the information that B . If my television and radio happen to be on at the same time, it does not follow that my television being on carries the information that my radio is on.

Secondly, take Stalnaker's system, which validates the following:

$$\vdash (A > B) \vee (A > \neg B)$$

This also would be unacceptable as a principle of information flow. It is not necessarily the case that A must carry either the information that B or the information that $\neg B$. My television being on carries neither the information that my radio is on nor the information that my radio is not on.

If this is true counterfactual territory, then counterfactual theories of information are in trouble.

4.3.2 Transitivity, Monotonicity and Contraposition

Whichever the logic used to model counterfactuals, failure of the following three conditions are generally taken to be essential to the nature of the counterfactual conditional.

- **Transitivity** - $(A > B) \wedge (B > C) \not\vdash (A > C)$
- **Contraposition** - $A > B \not\vdash \neg B > \neg A$
- **Monotonicity** - $A > B \not\vdash A \wedge C > B$

As was seen earlier, these three properties hold with Dretske's account. Should they hold for an account of information flow? I think that it is not a matter of simply determining whether or not these properties universally hold and selecting an account based on this determination. A richer account of information flow will provide a way to accommodate the validity of these properties in one sense and the possible invalidity of these properties in another sense. For now, I would like to discuss how the nature of counterfactuals, with regards to their invalidation of these three properties, makes them unsuitable to capture the idea of information flow we are working with.

Non-Transitivity

Fitting the counterfactual analysis of information with standard examples given in the literature can be a bit tricky. In the case of transitivity, the standard examples given of how transitivity can fail for counterfactuals do not clearly translate to a reading in terms of information flow; the types of premises given do not express proper information-that relations. Regarding transitivity, take Lewis' classic example from [123]:

- If J Edgar Hoover had been born a Russian, then he would have been a communist. (r)
- If J Edgar Hoover had been a communist, then he would have been a traitor. (c)
- Therefore, if J Edgar Hoover had been born a Russian, then he would have been a traitor. (t)

This is not a good argument, for if Hoover had been born a Russian then he would have been a patriotic communist. In terms of counterfactuals we can see why this is the case:

- if $r > c$ is true then the nearest r -world is also a c -world.
- if $c > t$ is true then the nearest c -world is also a t -world.
- it does not follow from these premises that the nearest r -world is a t -world ($r > t$), for the nearest r -world could differ from the nearest c -world.

What about this example from the perspective of information flow? Can we have the following?

- Hoover having been born a Russian carries the information that he would have been a communist
- Hoover having been a communist carries the information that he would have been a traitor
- It is not the case that Hoover having been born a Russian carries the information that he would have been a traitor

It is the context equivocation in this argument that eliminates the possibility of transitive information flow. Firstly, making some simplifications and barring some political exceptions,

we can fix a context and say that $r \sqsupset c$. The first problematic premise in terms of expressing an information relation is $c \sqsupset t$; someone being a communist is insufficient to carry the information that they are a traitor, as that relation is also dependent upon their country of origin. The same type of thing carries over to $r \sqsupset t$, thus it is no surprise that it does not hold, for the premises do not represent an information-that carrying chain.

Obviously here there are certain background conditions which are involved in the premises being asserted. And it is precisely this type of thing that counterfactual conditions are made to deal with. But whilst we can see why counterfactuals might operate like this, their failure to distinguish between invalidity across contexts and validity within the one context renders them insufficient to fully capture information carriage.

There is some flexibility with a counterfactual modelling of information flow, and one can see how it suggests the variability associated with different sets of relevant alternatives. But it is a limitation of $>$ that it cannot validate restricted forms of transitivity.¹⁸ As mentioned earlier, whilst C&M suggest that “it is in fact an advantage of the counterfactual account that it fails to licence the Xerox principle” [36, p. 7], they fail to provide any insight into how the Xerox principle might fail.¹⁹

Contrary to the supposed advantage in counterfactuals failing to licence the Xerox principle, I will show that the Xerox principle does capture a valid form of transitivity whilst at the same time showing how transitivity for information flow relations fails in a more general sense.²⁰ But although information-carrying relations need not be absolutely transitive, the non-transitivity of counterfactuals does not represent this sense of transitivity failure.

Non-Monotonicity

Perhaps the most telling invalidity is that of non-monotonicity. The following standard counter-example to monotonicity for counterfactuals will serve to illustrate what is going on here. Let b stand for ‘Tux is a bird’, f stand for ‘Tux flies’ and p stand for ‘Tux is a penguin’. Clearly the following monotonic reasoning is incorrect:

- $b > f$
- Therefore $(b \wedge p) > f$

¹⁸Apparently counterfactual accounts of causation are one place where people have tried to get transitivity for counterfactuals.

¹⁹In terms of C&M’s definition, the Xerox principle ($A \sqsupset B, B \sqsupset C \vdash A \sqsupset C$) fails since $A, B, C, \neg C > \neg B, \neg B > \neg A \not\vdash A, C, \neg C > \neg A$

²⁰Note that whilst the non-transitivity of counterfactuals fails to licence the Xerox principle, another validity related to the Xerox principle does hold for the counterfactual: $A, A > B, B > C \vdash C$

Framing this in terms of information carriage, we have:

- Tux being a bird carries the information that Tux flies
- Tux being a bird and Tux being a penguin does not carry the information that Tux flies (Tux being a penguin carries the information that Tux does not fly)

If the information-carrying relation were to be classed as monotonic, how would the above case be explained? One response would be to say that the premise is false ($b \sqsupset f$ does not hold); in general, if A carries the information that B , then A should suffice for the obtaining of B . In this case, there is insufficient information to imply that Tux does or does not fly. If supposedly $A \sqsupset B$ but not $A \wedge C \sqsupset B$, then A provides insufficient information to assert $A \sqsupset B$ in the first place.

This suggests that if the appropriate qualifications were made, then an assertion of information carriage could be made. For example:

- Tux being a penguin carries the information that Tux does not fly
- Tux being an eagle carries the information that Tux does fly

So information flow would be monotonic if A carries the information that B implies that A is sufficient for B . But being a bird is not a sufficient condition for flying, so there is no information-that relation here.²¹ However, one could take this further.

- Tux being an eagle carries the information that Tux flies
- Tux being an eagle and Tux having clipped wings does not carry the information that Tux flies

An adherence to the monotonicity of information flow would force one to deny the first point here. Clearly with this line of reasoning one could just keep on adding extra clauses to invalidate an apparent information carrying relation. When is A sufficient to carry the information that B ?

The key here is to realise that an information relation $A \sqsupset B$ exists relative to a fixed set of background conditions, a set of relevant alternatives. Considering, for example, a context

²¹Being a bird does carry information about the probability of flying though

where all that is being considered are standard pet shops with healthy birds, then something being a bird (A) carries the information that it flies (B). Within this relevancy set, there are no possibilities where something is a bird and is of a certain type or has a certain condition which precludes it from flying (C). Therefore $A \sqsupset B$ holds relative to this set. If however the set is changed to include all animals, then clearly $A \wedge C$ is possible and would imply $\neg B$. Overall, we can say that information flow is monotonic across one context and that monotonicity does not generally hold across contexts.

Whilst monotonicity does not hold for counterfactuals, information flow is monotonic under C&M's account, given that the added premise C refers to an actual event:

- $A \sqsupset B =_{df} A \wedge B \wedge (\neg B > \neg A)$
- $\neg B > \neg A \vdash \neg B > (\neg A \vee \neg C)$
- $\neg B > \neg A \vdash \neg B > \neg(A \wedge C)$
- $A \wedge C \sqsupset B =_{df} A \wedge B \wedge C \wedge (\neg B > \neg(A \wedge C))$

Non-Contraposition

The failure of contraposition is perhaps the most problematic. Continuing with the bird theme, take the following:

1. If Tweety is a bird then Tweety would fly ($b > f$)
2. Therefore if Tweety were not to fly then Tweety would not be a bird ($\neg f > \neg b$)

Counterfactually speaking, it could be the case that (1) is true whilst (2) is false. For example, in the closest $\neg f$ worlds, it could be the case that Tweety is an emu, in which case he is still a bird. Framing this in terms of information carriage, we have:

1. Tweety being a bird carries the information that Tweety flies ($b \sqsupset f$)
2. Tweety being flightless carries the information that Tweety is not a bird ($\neg f \sqsupset \neg b$)

Once again we have a situation where both of these statements hold relative to certain contexts. Given a background condition of animals excluding flightless birds, then $b \sqsupset f$ holds and $\neg f \sqsupset \neg b$ follows. On the other hand, given the absence of this background

condition, both $b \sqsupset f$ and $\neg f \sqsupset \neg b$ do not hold. So an information relation $A \sqsupset B$ and its contraposition $\neg B \sqsupset \neg A$ only hold and imply each other relative to a fixed set of relevant alternatives.

It really does seem like contraposition has some part to play in an account of information flow, so the lack of a means to capture the cases where contraposition is valid is problematic.²² There needs to be a validity corresponding to the following intuition: if every time A happens then B happens, then B not happening means that A has not happened.

The failure of contraposition for counterfactuals also enforces the arguably stringent requirement in C&M's definition that the events being talked about are actual events, lest the possibility that veridicality fails. For if this requirement were removed, then since contraposition is not valid in counterfactual logic ($\neg Gy > \neg Fx \not\vdash Fx > Gy$), this would mean that it would be possible for Fx to carry information about Gy whilst $Fx > Gy$ is false, which means that $Fx \wedge \neg Gy$ could not be ruled out.

4.4 A Modal Logical Account of Information Flow

So as the cycle of philosophy would have it, Dretske's account faces its share of criticism. As we have just seen, amongst the critics there are those who have advocated a counterfactual theory of information. But as I argued, such an account is itself problematic, with one reason being that properties of a counterfactual conditional discord too much with properties commonly associated with information flow. If an account of information is to abandon the universality of these properties, it could at least offer some accommodation by delineating the restricted applicability of these properties and account for why they are linked to an ordinary conception of information flow.

In this section I show how Dretske's account can essentially be translated into a modal logic framework, synthesising ideas and observations from the previous sections along the way. As we will see, this framework actually turns out to be a more flexible and powerful way to express information flow.

Let us start with the following 'template' definition of information flow:

A carries the information that B relative to some system S if and only if S determines that
in every state where A is the case B is the case.

²²Recall that it even holds for the probabilistic account of Section 4.2

We can formally express this definition simply by using a strict conditional: $\Box(A \supset B)$. Thus we get the following:

$$A \text{ carries the information that } B =_{df} \Box(A \supset B)$$

Conversely:

$$A \text{ does not carry the information that } B =_{df} \Diamond(A \wedge \neg B)$$

For now, let the specific modal logic for $\Box$ be the base normal modal logic **K**. For those not happy with talk of probabilities, this approach offers a viable alternative and can also be made to fit with Loewer's definition discussed in Section 4.3. It also stands to reason that since Dretske's account requires a conditional probability of 1 and does not make use of the probability value range $[0,1]$, probability can be replaced with logic. Clearly there is a correspondence between this modal definition and a conditional probability of 1:

$$\Pr(B|A) = 1 \Leftrightarrow \Box(A \supset B)$$

Here is an informal sketch of this:

1. $\Pr(B|A) = 1 \Leftrightarrow A \subseteq B$ (The states with A are a subset of the states with B)
2. $\Box(A \supset B) \Leftrightarrow A \subseteq B$ (Where $\Box$ ranges over states)
3. $\therefore \Pr(B|A) = 1 \Leftrightarrow \Box(A \supset B)$

The sets of worlds W for a model equate to the set of possible states or alternatives. So the accessibility relation for $\Box$ ranges over possible states. Given the actual state w , the set of all states accessed from w ($\{w' | wRw'\}$) can be treated as a set of relevant alternatives.

This modal formulation simply captures the core connection involved in A carrying the information that B ; every time A occurs, B occurs. On top of this, a couple of additions can be made to reflect certain conditions associated with information flow. As suggested via Dretske's definition, one standard condition is that a signal A can only carry the information that B if B is a contingent truth. The addition of $\neg\Box B$ would be used to express the fact that if B is a necessary/tautological truth, then A does not carry the information that B . A condition $\Diamond A$ can also be added to rule out contradictions trivially carrying information. This gives us the following full definition:

A carries the information that $B =_{df} \neg \Box B \wedge \Diamond A \wedge \Box(A \supset B)$

As with Dretske's definition, an agent's background knowledge can also be incorporated into this definition. This will be covered later.

One extra feature of this approach is that it affords a way to express iterated information flow. For example, that A carries the information that B carries the information that C could be represented with:

$$\Box(A \supset \Box(B \supset C))$$

How this could be done using conditional probabilities I do not know.

The employment of a strict conditional emphasises that truth preservation is different to information carriage. Whilst counterparts of the following are valid in classical logic, they do not hold for this account of information flow:

- $\Box((A \wedge B) \supset C) \not\models \Box(A \supset C) \vee \Box(B \supset C)$
- $\Box(A \supset B) \wedge \Box(C \supset D) \not\models \Box(A \supset D) \vee \Box(C \supset B)$

Also, clearly neither of the following two counterfactual oddities from Section 4.3 is valid:

- $A \wedge B \not\models \Box(A \supset B)$
- $\not\models \Box(A \supset B) \vee \Box(A \supset \neg B)$

As with Dretske's account and as expected, the core definition $\Box(A \supset B)$ validates all of the following properties: Conjunction, Reverse Conjunction, Disjunction, Reverse Disjunction, Transitivity, Contraposition and Monotonicity.

When things like the condition to rule out vacuousness ($\neg \Box B$) and the agent's prior information/knowledge are factored in, then some of these conditions fail. For example, Reverse-Conjunction:

$$\Box(A \supset (B \wedge C)) \wedge \neg \Box(B \wedge C) \not\models \Box(A \supset C) \wedge \neg \Box C$$

as the following counter-model demonstrates:

- $w_1 R w_2$
- $w_1 R w_3$
- $w_1 R w_4$
- $v_{w_2}(A) = 1, v_{w_3}(A) = 1, v_{w_4}(A) = 0$
- $v_{w_2}(B) = 1, v_{w_3}(B) = 1, v_{w_4}(B) = 0$
- $v_{w_2}(C) = 1, v_{w_3}(C) = 1, v_{w_4}(C) = 1$

Finally, there is at least one important addition that needs to be made to this base logic. To ensure veridicality ($A \wedge A \sqsupset B \vdash B$), we add the axiom $\Box A \supset A$ and get reflexive frames. Therefore, the system we are working with is the modal logic **T** and the actual state is necessarily one of the alternatives that is relevant.

The reflexivity property also gives us other desirable validities which would not hold without it. For example, if A carries the information that B carries the information that C , then A and B together carry the information that C :

$$\Box(A \supset \Box(B \supset C)) \vdash \Box((A \wedge B) \supset C)$$

As is made clear in this formalisation, veridicality does not mean that for $A \sqsupset B$ to hold both A and B must be true or actual. It means that if $A \sqsupset B$ does hold, then the actuality/truth of A entails the actuality/truth of B .

The metal detector's beep (A) carries the information that there is metal (B) even if there is actually no metal and therefore actually no beep. In the relevant alternative that is the actual world $\neg A$ and $\neg B$ so $A \supset B$ holds. In all relevant alternatives where A is true so is B , in which case $A \supset B$ would hold.

4.4.1 Variable Information Flow Contexts

Whilst the validation of our test set of conditions is essentially correct, there are cases where it seems reasonable to talk of information flow in a way that violates one or more of the conditions. For example,

violations of transitivity occur when the signal type A of the first channel is inconsistent with the background conditions on the channel from B to C . For example, let A be ‘the doorbell is short circuited’, B be ‘the doorbell is ringing’, and C be ‘someone is at the door’; that B carries the information C is one of Dretske’s standard examples. Then, contrary to Barwise and Seligman, it seems intuitively reasonable to accept that A carries the information B , B carries the information C , but A does not carry the information C . Thus transitivity should be rejected by any theory of information flow. [178, p. 105]

The task here now is to devise a system that accommodates the form of transitivity embodied in the Xerox principle whilst rejecting transitivity in general. Let us begin by laying out this example:

1. the doorbell is short circuited *carries* the information that the doorbell is ringing
2. the doorbell is ringing *carries* the information that someone is at the door
3. the doorbell is short circuited *does not carry* the information that someone is at the door

What is going on here is that strictly speaking relation 2 only holds if the doorbell is not short circuited. If the doorbell is ringing because it is short circuited, then it plainly does not carry the information that someone is at the door. So with the possibility that the doorbell is short circuited in such a scenario, the doorbell ringing does not carry the information that someone is at the door. Given this, one option is to deny relation 2.

But there still seems to be a sense in which we want to maintain the validity of both 1 and 2, in which case there is a failure of transitivity. How does this work? Well, if relation 2 above holds, it will involve the exclusion of all alternatives in which the doorbell is short-circuited. So there are alternatives associated with relation 1 that will not be associated with relation 2. Thus both relation 1 and relation 2 hold relative to a certain set of alternatives. Since the set of accessed alternatives for a state is captured using $\Box$, the employment of a multi-modal logic will easily allow us to capture this idea. So now let $\Box_1, \Box_2, \dots, \Box_n$ stand for a group of operators, each one with its own accessibility relation $R_1, R_2, \dots, R_n$. A relation such as wR_1w' means that w' is an accessed alternative for $\Box_1$. This can be used to express the idea that according to the context or channel conditions associated with $\Box_1$, w' is a relevant alternative.

We can now capture both the Xerox principle and the rejection of universal transitivity. The Xerox principle applies when the set of relevant alternatives is the same for both $\Box(A \supset$

B) and $\Box(B \supset C)$:

$$\Box_x(A \supset B) \wedge \Box_x(B \supset C) \vdash \Box_x(A \supset C) \text{ (Xerox Principle)}$$

On the other hand, if $\Box(A \supset B)$ and $\Box(B \supset C)$ do not share the same set of alternatives, then transitivity need not hold:

$$\Box_x(A \supset B) \wedge \Box_y(B \supset C) \not\vdash \Box_z(A \supset C) \text{ (Non-Transitivity)}$$

where $\neg(x = y = z)$.

The addition of $\Box_2 A \supset \Box_1 A$ as an axiom can be used to define a frame condition which ensures that the relevant alternatives for $\Box_1$ form a subset of the relevant alternatives for $\Box_2$ ($w_1 R_1 w_2 \Rightarrow w_1 R_2 w_2$).

Let us apply all of this to the doorbell example. We start off with the following:

- $\Box_1(A \supset B)$ = the doorbell is short circuited *carries* the information that the doorbell is ringing
- $\Box_2(B \supset C)$ = the doorbell is ringing *carries* the information that someone is at the door

Failures of transitivity would be represented with invalidities such as the following:

- $\Box_1(A \supset B) \wedge \Box_2(B \supset C) \not\vdash \Box_1(A \supset C)$
- $\Box_1(A \supset B) \wedge \Box_2(B \supset C) \not\vdash \Box_2(A \supset C)$

In the doorbell example neither $\Box_1 A \supset \Box_2 A$ nor $\Box_2 A \supset \Box_1 A$ are applicable as axioms. It is certainly not the case that $\Box_2 A \supset \Box_1 A$ is valid, for there are $\Box_1$ alternatives where the doorbell is ringing and there is no one at the door ($B \wedge \neg C$).

It is also not the case that $\Box_1 A \supset \Box_2 A$ is valid. For all $\Box_2$ alternatives, if someone is not at the door then the doorbell is not ringing ($\Box_2(B \supset C) \Leftrightarrow \Box_2(\neg C \supset \neg B) \Leftrightarrow \neg \Diamond_2(\neg C \wedge B)$). Yet this is exactly something that would happen in an alternative corresponding to $\Box_1$. So:

- it is not the case that every $\Box_1$ alternative is a $\Box_2$ alternative. If $\Box_2 A \supset \Box_1 A$ were to be implemented as an axiom, then the unacceptable $\Box_1(A \supset C)$ would be derivable.
- it is not the case that every $\Box_2$ alternative is a $\Box_1$ alternative. If $\Box_1 A \supset \Box_2 A$ were to be implemented as an axiom, then the unacceptable $\Box_2(A \supset C)$ would be derivable.

With this system in place we have a way to capture the sense in which properties of information flow that hold on one level fail to universally hold. For example, we can easily see why monotonicity fails:

$$\Box_2(B \supset C) \not\vdash \Box_1((B \wedge A) \supset C) \text{ (Non-Monotonicity)}$$

That the doorbell is ringing carries the information that someone is at the door relative to a certain set of relevant alternatives in which it is not the case that the doorbell is short-circuited. However it does not follow from this that relative to a different set of alternatives in which it is possible that the doorbell is short-circuited, the doorbell ringing and being short-circuited carries the information that someone is at the door.

This is not the whole story though and there are some final pieces to put in place. These pieces will deal with problematic validities such as the following: $\Box_2(B \supset C) \vdash \Box_2((B \wedge A) \supset C)$.

To continue this development, we will switch to another example which will better serve to highlight the issues. Take the following predicates:

- $Pa = a$ is a penguin
- $Ba = a$ is a bird
- $Fa = a$ flies

and the following information relations:

1. $\Box_1(Pa \supset Ba)$
2. $\Box_2(Ba \supset Fa)$
3. $\Box_1(Fa \supset \neg Pa)$

These relations all seem fine. Relations 1 and 3 are zoological truths that are true in any state of the world and therefore we can set them to range over the same set of alternatives. The states associated with relation 2 form a subset of the states associated with 1 and 3, one which excludes states with flightless birds.

As it is this setup has some clearly unacceptable consequences. Firstly, if it is actually the case that Pa then by implication it is also the case that $\neg Pa$. Secondly, since $\Box_2$ ranges over a subset of $\Box_1$, it would follow that $\Box_1 A \supset \Box_2 A$ and therefore $\Box_2(Pa \supset \neg Pa)$: a being a penguin carries the information that a is not a penguin. Other clearly problematic results follow from further additions. For example, adding the zoological truth $\Box_1(Pa \supset \neg Fa)$ leads to the problematic $\Box_2(Pa \supset \neg Ba)$.

Framing the information carriage relation in terms of the counterfactual $>$ fares little better. Take the following:

1. $Pa > Ba$
2. $Ba > Fa$
3. $Fa > \neg Pa$

Since transitivity is not valid, the problematic $Pa > \neg Pa$ is not derivable. But as above, if these information relations are combined with Pa , then $Pa \wedge \neg Pa$ is derivable.

What has gone wrong here is a lack of required qualifications. If $\Box_2(Ba \supset Fa)$ is to hold, then it must not be the case that $\Box_2$ ranges over alternatives in which Pa is the case. The trick is to realise that $\Box_2(Ba \supset Fa)$ holds or fails to hold relative to certain conditions. If a being a bird carries the information that a flies, then Pa is not amongst the relevant alternatives being considered.

This can be formally expressed as

$$\Box_2(Ba \supset Fa) \supset \Box_2 \neg Pa$$

which is perhaps more clearly expressed with the equivalent

$$\Diamond_2 Pa \supset \neg \Box_2(Ba \supset Fa)$$

With all this in place, our starting three relations become:

1. $\Box_1(Pa \supset Ba)$
2. $\Box_2(Ba \supset Fa) \supset \Box_2\neg Pa$
3. $\Box_1(Fa \supset \neg Pa)$

From these $Pa \supset \neg\Box_2(Ba \supset Fa)$ is derivable whereas $Pa \supset \neg Pa$ or $\Box_2(Pa \supset \neg Pa)$ are not derivable.

If $\Box_2(Ba \supset Fa)$ is given, then the following two can be derived:

- $\Box_2\neg Pa$
- $\Box_2(Pa \supset \neg Pa)$

Whilst the second of these might seem problematic, with a requirement that the antecedent of an information relation be possible, Pa fails to carry the information that $\neg Pa$.

We can also deal with monotonicity and contraposition and show how they fail. For monotonicity, suppose that $\Box_2(Ba \supset Fa)$ holds. Since it follows from this that $\neg\Diamond_2 Pa$, then given a definition of information with a qualification ruling out impossible signals, it is not the case that $(Ba \wedge Pa) \supset Fa$. Regarding $\Box_1$, $\Box_1(Ba \supset Fa)$ is illegitimate and leads to the incorrect $\Box_1(Pa \supset \neg Pa)$. $\Box_1(Ba \wedge Pa \supset \neg Fa)$ is derivable in the system so the addition of $\Box_1(Ba \wedge Pa \supset Fa)$ would lead to the unacceptable $\neg\Diamond_1(Ba \wedge Pa) \vee \Diamond_1(Fa \wedge \neg Fa)$. Hence $\Box_1(Ba \wedge Pa \supset Fa)$ is also illegitimate.

For contraposition, $\Box_2(Ba \supset Fa)$ implies $\Box_2(\neg Fa \supset \neg Ba)$; this is unproblematic. But $\Box_2(Ba \supset Fa)$ does not imply $\Box_1(\neg Fa \supset \neg Ba)$, which is as it should be.

This multi-modal framework provides a simple but effective way to talk about information flow, in which we can easily make use of the proof procedures available for modal logic to explore valid information relations. As we have seen, it requires that the conditions which preclude the presence of an information relation from A to B be spelled out. Where $C_1, C_2, \dots, C_n$ are such conditions, statements of information carriage are of the form: $\Box(A \supset B) \supset \Box\neg(C_1 \vee C_2 \vee \dots \vee C_n)$. In Dretskean parlance, $\neg C_1, \neg C_2, \dots, \neg C_n$ can be described as channel conditions. In this way, since $\neg C_1, \neg C_2, \dots, \neg C_n$ hold over the given set of relevant alternatives, information-carrying signals cannot tell us anything new about these channel conditions relative to this set. Whilst theoretically or philosophically the list of conditions can be open-ended and indefinite, in practical applications with a finite, manageable list of

predicates, an appropriate listing of conditions will be possible.²³

Going back to the metal detector example in Section 4.1, we can now formalise the fact that a metal detector's active tone can be used to provide the information that there is metal or the information that the detector is functional/dysfunctional, but not both:

- T - metal detector sounds
- M - metal is present
- F - metal detector is functional
- $\Box_x(T \supset M) \supset \Box_x F \wedge \neg \Box_x M$
- $\Box_x(T \supset F) \supset \neg \Box_x F$
- $\Box_x(T \supset \neg F) \supset \Box_x \neg M \wedge \neg \Box_x \neg F$
- $T \wedge \Box_x(T \supset M) \wedge [\Box_x(T \supset F) \vee \Box_x(T \supset \neg F)] \vdash \perp$

I would like to close this section with a good example that really shows the importance and applicability of this relevant alternatives framework in making sense of realistic ways in which we treat signals as carrying information. Navigation systems based on the Global Positioning System (GPS) are now commonplace. Take a standard scenario in which a lost driver, who is located at a particular address x , consults their fully functioning GPS navigation device. The device indicates that they are at address x and we can say that the device's signal carries the information that they are at address x ; since the GPS is fully functioning, every time the device indicates address x one is at address x .

But in attributing such information carriage to this signal we are implicitly confining the alternatives against which we make this judgement to locations on Earth. For the GPS system uses a trilateration method such that given its calculations, the location of the GPS device can be narrowed down to two points. One is the point corresponding to address x and the other is a point somewhere off in space. Given that the latter is irrelevant with regards to determining geo-positioning, the GPS device can automatically eliminate it.

So a signal of position x from the GPS system carries the information that you are at position x on Earth given this set of alternatives. If the set of alternatives were expanded to include all points in the Solar System, then the signal would fail to carry this information relative to this greater set. At best it would carry the information that one is at position x

²³Perhaps in less practical discourse, a catch-all clause C could act as a stand-in of sorts, such that $\Box(A \supset B) \supset \Box \neg C$.

on earth or at its corresponding position y in space. Given that the actual state is one in which the person actually is at x , a standard context can be fixed whereby the possibility that one is floating in space is deemed irrelevant.

4.4.2 Agent-Relative Information Flow

As we have seen, an agent's background knowledge figures in determining what information a signal carries for them. This agent-relativeness is part and parcel of the way we use the notion of information. It motivated Chapter 3, where a formal framework was constructed in which the informativeness of a piece of semantic information depends on the receiving agent's database.

For information flow Dretske incorporates the variable k (representing the agent's background knowledge) into his definition, making, as he puts it, "a minor concession to the way we think and talk about information" [52, p. 58].

The modal logic framework given in 4.4 can be easily extended to incorporate an agent's background knowledge, so that information flow is relativised to the given agent. Letting the formula K stand for agent a 's background knowledge base, we get the following:

$$A \sqsupset B =_{df} \Diamond A \wedge \Box(A \wedge K \supset B) \wedge \Diamond(K \wedge \neg B)$$

As the simplest of examples, let $K = p \vee q$, $A = \neg q$ and $B = p$. Then A carries the information that B for a :

- $\Diamond \neg q$
- $\Diamond((p \vee q) \wedge \neg p)$
- $\Box(\neg q \wedge (p \vee q) \supset p)$

A consequence of Dretske's account is that logical, mathematical or analytic truths generate zero information. It is also interesting that nomic necessities, such as the identity $\text{H}_2\text{O} = \text{water}$ generate zero information, since these things have a probability of one. This is an informational version of Frege's puzzle; both $a = a$ and $a = b$ generate no information. Consequently, no signal can carry the information that $a = b$. Whilst treating logical, mathematical and analytic truths this way is okay given the purpose of this account, the

consequence for nomic necessities is particularly problematic: “It certainly seems strange to treat nomic necessities in the same way one treats mathematical or logical truths.” [3, p. 232]

Given an upgrade to first-order modal logic with identity, the modal logic framework can simply deal with this issue by incorporating a form of contingent rather than necessary identity. Another possibility is to change the underlying logic of the system so that it becomes possible for certain pieces of information to be carried. For example, if the system of classical logic were to be replaced by intuitionistic logic, then a signal could carry the information that $p \vee \neg p$.

4.4.3 Information Closure

One point of contention in epistemology is the Epistemic Closure Principle [127, 39], according to which the set of things one knows is closed under entailment. There are many ways to formulate the closure principle. For our purposes, we will settle on the following broad formulation, taken from [169, p. 2]:

If S knows that p and knows that p implies q , then S is in a position to know that q .²⁴²⁵

In this section, our focal point will be the following informational versions of the closure principle:

- If S holds the information that p and S holds the information that p implies q , then S holds the information that q .
- If S holds the information that p and p carries the information that q , then S holds the information that q .

The main task consists of addressing the following issues:

- Dretske rejects closure for knowledge. Although the foundations for this rejection preceded his informational account of knowledge, if closure for knowledge is to fail in

²⁴Another suitable formulation that captures the same gist is the following:

If S knows that p , and comes to believe that q by correctly deducing it from her belief that p , then S knows that q . [39]

²⁵ As will be covered in more detail later, in epistemic logic the closure principle can be formalised as:

$$Kp \wedge K(p \supset q) \vdash Kq$$

this account, then closure for information must also fail. Problematically for Dretske though, given the definition he uses in his account of information, it seems like closure for information does hold. I will show that this issue rests on a limitation with the technical foundations of Dretske's definition before showing how the failure of closure for information he has in mind can be captured using the multi-modal framework given in the previous section.

- One important distinction is that between contextualist relevant alternative accounts that accept closure and relevant alternative accounts such as Dretske's that deny closure. As will be shown, this distinction can be easily captured using the multi-modal logic framework. Whilst I do not exactly intend to take sides here and argue for one position over the other, I aim to explicate both positions and tease out their logic, particularly the logic of Dretske's account.

Whilst the closure principle does seem quite sensible ²⁶, it is not without reason that some epistemologists such as Dretske and Nozick deny that closure always holds. There is a variety of reasons why the closure principle is rejected. Dretske (and Nozick) essentially denies closure because under his analysis of knowledge, knowledge is not closed. This is the argument from the analysis of knowledge:

given the correct analysis, knowledge is not closed, so it isn't. For example, if the correct analysis includes a tracking condition, then closure fails. [127]

So rather than reject the analysis based on the idea of relevant alternatives, which he believes to be correct, reject the principle. For Dretske, one can know that p because they have ruled out all relevant alternatives to p and even if they know that $p \supset q$, since they cannot rule out all the alternatives relevant to q (some of which will be irrelevant to p) they are not in a position to know q :

Suppose that the alternative that this is a Siberian grebe is irrelevant to my knowing that it is a Gadwall duck. Notice too that the negation of the former proposition is a necessary consequence of the latter proposition if this is a Gadwall duck, then it is not a Siberian grebe. Dretske claims that I can know that this is a Gadwall duck even though I don't know that it's not a Siberian grebe. Thus, Dretske holds that the closure principle is false. [39]

²⁶[It] is both plausible and explanatorily valuable. For one thing, it helps to explain how we come to know things via deduction. I know, for example, that tomorrow is Saturday. I know this because I know that today is Friday and that if today is Friday then tomorrow is Saturday. The closure principle helps to account for this knowledge, and the fact that I come to know things via deduction and in accordance with the closure principle renders that principle both plausible and desirable [39].

It is important to note that this reason adduced by the likes of Dretske for rejecting closure is but one amongst several reasons for rejecting closure of some sort. For example, as highlighted by issues in epistemic logic, there is the issue of logical omniscience and the acceptable reason for wanting to avoid simpler formulations of closure such as:

If S knows that p and knows that p implies q , then S knows that q .

Of course, knowing that p and that p implies q does not necessarily mean that one knows q , since one might not have performed the required deductive inference to know q . But when Dretske denies closure, this is not the issue. It seems that this has been a point of some confusion in the literature. For example, in a recent collection on Floridi and The Philosophy of Information, Fred Adams [1] challenges Floridi's endorsement of information closure, which is part of the latter's logic of being informed. But whilst Adams is basing this challenge on the Dretskean line, Floridi's commentary in response to Adams' paper suggests something different. This disagreement seems to be in large part due to each of them having a different conception of being informed. Consider the following quote from Floridi:

In the left pocket of your jacket you hold the information that, if it is Sunday, then the supermarket is closed. Your watch indicates that today is Sunday. Do you hold the information that the supermarket is closed today? The unexciting answer is maybe. Perhaps, as a matter of fact, you do not, so Adams is right. You might fail to make the note in the pocket and the date on the watch "click." Nevertheless, I would like to argue that you should, that is, that as a matter of normative analysis, you did have the information that the supermarket was closed. So much so that you will feel silly when you are in front of its closed doors and realise that, if you had been more careful, you had all the information necessary to save you the trip. [70, p. 410]

Floridi's example case is correct. As will be consolidated at the end of this section, even on a Dretskean analysis one would have access to the information that the supermarket was closed in this example, even if they did not explicitly make that inference. What this suggests is that being informed, unlike knowledge, can be considered as 'implicit'²⁷; an agent does not need to actively/consciously infer the information to have it.²⁸ This is also why the

²⁷Recalling certain distinctions between explicit belief/knowledge and implicit belief/knowledge [121], the notion of being informed here is aligned with the latter.

²⁸When talking about the information held by an agent, we need not restrict talk to the content they have inferred or are aware of. We can talk about all the true content closed under consequence, which the agent has the potential to logically infer. If an agent knows that p and knows that $p \supset q$, then they will only come

definition of information closure at the start of this section, in contrast to the definition of knowledge closure, does not contain the qualification ‘in a position’.

But Floridi closes the discussion with the following:

If Adams’s thesis is that information closure is at best only a matter of normative logic and certainly not an empirical fact, I am convinced. [70, p. 411]

As I understand it this is not Adams’ thesis. According to Adams (and Dretske), information closure is not even a matter of normative logic; there are some cases where the agent will be informed that p , be informed that $p \supset q$ and make the inference that q , yet still not be informed that q . This is because even if you infer q , the conditions might not be right for it to count as information.

I would also like to briefly address another argument against closure before focusing on the argument from the analysis of information/knowledge. This is the argument from nonclosure of knowledge modes:

since the modes of gaining, preserving or extending knowledge, such as perception, testimony, proof, memory, indication, and information are not individually closed, neither is knowledge. [127]²⁹

Dretske has argued that since one of the modes of gaining, preserving or extending knowledge are individually closed we should expect closure to fail. As he puts it, “how is one supposed to get closure on something when every way of getting, extending and preserving it is open?” [127]

Adams mentions something similar in response to Floridi, who after introducing the distributive thesis for his logic of being informed comments that:

although this is entirely uncontroversial, it is less trivial. Not all ‘cognitive’ relations are distributive. ‘Knowing’, ‘believing’ and ‘being informed’ are, as well as ‘remembering’ and ‘recalling.’ ... However, ‘seeing’, and other experiential relations, for example, are not; if an agent a sees (in a non-metaphorical sense)

to know that q once they have made that inference. On the other hand, if an agent has the information that p and the information that $p \supset q$, then they have the information that q , independently of the inference being made by them. Restricted to this sense, a modal logic of being informed can be normal and is not troubled by the logical omniscience problem.

²⁹We could replace ‘knowledge’ with ‘information’ and remove ‘information’ from the list of modes.

or hears or experiences or perceives that $p \rightarrow q$, it may still be false that, if a sees (hears etc.) p , a then also sees (hears etc.) q . [67, p. 440]

Adams notes that:

Floridi's choices of exceptions to the principle are telling. For tracking theorists of empirical knowledge, how can the knowledge resting upon perceptual states obey this principle, when the very experiential states themselves do not obey the principle? Knowledge acquires its properties from its informational base. *If the experiences upon which one's perceptual knowledge rests do not obey this informational principle, then empirical knowledge itself may also fail to obey the principle.* [1, p. 339] (italics mine)

Upon reading this passage it struck me as missing something crucial. Although one particular empirical method of gaining information, such as seeing, may not be distributive/closed, that doesn't mean that one or more methods cannot be combined to infer some final piece of final information. For example, one might hear on the television that there is a football game being played on the 6th of August ($p \supset q$) and see that it is the 6th (p) of August by looking at their calendar. Via these experiential states they come to acquire the information that $p \wedge (p \supset q)$. Although they do not see that q or hear that q , they can simply use the method of deductive inference to gain the information that q .

With this detour discussion of other arguments against closure out of the way, we now turn to focus on the argument from the analysis of knowledge. The idea that closure is not universally valid because of the analysis of knowledge was suggested early by Dretske and can be found amongst such pieces as 'Epistemic Operators' [49] and 'Conclusive Reasons' [50]. Whilst perhaps counterintuitive, some thought on the matter can see a reason for this position.

According to Dretske, knowing that p requires having a conclusive reason for p . Following is a definition of this concept:

R is a conclusive reason for $p =_{df}$ If p were false R would be false (or: If p were not the case R would not be the case). [169, p. 2]

Basically, the idea is that R is a conclusive reason for p if it is the case that amongst all the relevant alternatives for p , if not- p holds then R does not hold. Given that knowledge

requires conclusive reasons, sometimes knowledge will not be closed under known implication. Take the following classic Zebra example from Dretske [49].

You are at the zoo and go to the zebra section. Your visual experience of a zebra is a conclusive reason for, and so provides you with the knowledge that the animal before you is a zebra. The zebra visual experience is a conclusive reason because in all the relevant alternatives where it is false that a zebra is before you in the enclosure there is no zebra-like visual experience. Say this set of relevant alternatives consists of worlds with standard zoo scenarios in which other animals such as tigers or camels are in the enclosure.³⁰

Now, something being a zebra implies that it is not a cleverly disguised mule painted to look like a zebra. This is an analytic truth which you know. Does it therefore follow that you know that the animal before you is not a cleverly painted mule? According to the conclusive reasons analysis, the answer is no. This is because you do not have a conclusive reason for it not being the case that the animal is a disguised mule. To have a conclusive reason, you would need to be able to rule out relevant alternatives where the animal is a disguised mule. But since your visual experience would be the same for a zebra and a disguised mule, it is not the case that if the animal was a painted mule, then you would not have the zebra-like visual experience. That is, the visual experience is not a conclusive reason for not-mule.

The catch here is that “the negation of a proposition p is automatically a relevant alternative to p (no matter how bizarre or remote not- p might be) but often not a relevant alternative to things that imply p [so] we can know something p only if we can rule out not- p but we can know things that entail p even if we cannot rule out not- p [127].

Framing this idea of closure failure another way, we can say that knowledge of ordinary things generally implies knowledge of what Dretske terms *heavyweight propositions*. One can read a newspaper and therefore come to know the results of last night’s football game. That the team you support won implies the following heavyweight propositions:

- The physical world exists
- The past is real and not subject to a Russellian-scenario, where the world was only created a few minutes ago

Despite these implications, the methods by which you gain knowledge of some ordinary proposition will not suffice to give you knowledge of heavyweight propositions it implies. Whilst the information that (some ordinary proposition) A implies the truth of (some heavyweight proposition) B , the method by which A was acquired cannot inform you whether the

³⁰This example is sometimes talked about in terms of nearby possible worlds.

implied heavyweight proposition is true. So whilst you can gain empirical knowledge via things like perception and testimony, these things don't give one conclusive reasons for ruling out these heavyweight propositions. To put it another way,

in general terms, instruments can be reliable indicators of certain features of reality. But they cannot be indicators of their own reliability, no matter how reliable they may in fact be. [169, p. 3]

One important thing to take from this discussion is that the rejection of closure is a qualified one; generally closure holds. It fails in cases involving an ordinary proposition and a heavyweight proposition as just explained.

Dretske's strategy offers a way of dealing with the following trilemma [169, p. 3]:

1. We have lots of knowledge of ordinary things
2. There is no way of knowing the heavyweight implications of our knowledge of ordinary things
3. Closure

It is this trilemma which the skeptic uses. Once again, take the Zebra example. Firstly, the following seems fair:

- It is not the case that you know that the animal before you is not a painted mule. For a zebra-like visual impression would be given by a zebra or a painted-mule, so you cannot rule out not-mule

But furthermore, it seems reasonable to make the following claims:

- You know that the animal before you is a zebra
- You know that if it is a zebra, then it is not a painted mule
- Therefore you know that the animal is not a painted mule

The result is a contradiction: you both know and do not know that the animal is not a painted mule.

Rather than rejecting 1 or 2, Dretske opts to reject 3. In this way he can hold on to the claim that we have knowledge of ordinary things (it is a zebra) whilst acknowledging that we cannot know the implied true negation of some skeptical proposition (it is not a disguised mule).

As an interesting historical footnote to wrap up this portion of the discussion with, it is worth noting that Dretske

wasn't led to deny closure because it represented a way around skepticism. [He] was led to it because it was a result of what [he] took to be a plausible condition on evidence (justification, reasons) required for knowledge. [57, p. 43]

By 1981 with *Knowledge and the Flow of Information* Dretske framed all of this in terms of his information-theoretic account of knowledge. Since Dretske bases his definition of knowledge on information, it should follow that information is also not closed under entailment. There has recently been some new discussion in the literature regarding the matter, with Jäger [111] suggesting that given Dretske's definition of information, information, hence knowledge, are in fact closed under known entailment. For if, as it is claimed, under Dretske's account the closure principle for information holds, then since his account of knowledge is based on information this spells a problem and would force him to accept the epistemic closure principle.

Ultimately Jäger's argument does not really trouble Dretske's account, and Dretske has no problems himself providing a defence in [58] on similar terms to my own response. But, as I will show, ultimately some clarifications and minor modifications to Dretske's formalisation of his account are required. Let us start with Jäger's technical argument, which shows that information closure follows from Dretske's probabilistic definition of information.

Jäger approaches the issue by examining whether information is closed under known entailment using the following formulation:

PIC : If r carries, relative to the subject K , the information that p , and K knows that p entails q , then r carries, relative to K , the information that q . [111, p. 192]

Let e stand for some appropriate empirical proposition (such as 'this is a zebra'), h stand for some heavyweight proposition and k stand for the agent's background knowledge. According to Dretske's definition of the information, the veridicality of knowledge and the probability calculus, if r carries the information that e we get:

$$(1) \Pr(e|r \wedge k) = 1 \ \& \ \Pr(e|k) < 1$$

$$(2) e \rightarrow \neg h$$

Therefore:

$$(3) \Pr(\neg h|r \wedge k) = 1$$

In keeping with Dretske's condition that the prior probability of the proposition in question is less than one, it only remains to be shown that $\Pr(\neg h|k) < 1$. Jager does this by simply noting that since h is a skeptical hypothesis, we are not entitled to be certain that it is false; there is at least a minute, non-zero probability that it is true. Therefore $\Pr(\neg h|k) < 1$.

Given all of this, information is apparently closed according to Dretske's definition. Shackel [165] and Baumann [17] offer some manoeuvres which would provide a way out of Jager's problem. Even if Jager's example was legitimate, Shackel demonstrates that there are cases where PIC would fail. Start by re-formulating PIC as:

PIC Dretske: If $\Pr(p|r \wedge k) = 1$ and $\Pr(p|k) < 1$ and you know that p entails q , then $\Pr(q|r \wedge k) = 1$ and $\Pr(q|k) < 1$

As the following counterexample demonstrates, this is false:

Let p be E and q be $E \vee F$, and let $E \vee F$ be a conjunct of k . Clearly p entails q and we assume the subject knows it. Let $\Pr(p|r \wedge k) = 1$ and $\Pr(p|k) < 1$. When the antecedent of *PIC Dretske* is thus satisfied, its consequent is false because the second conjunct is false. $\Pr(q|k) = 1$ because q , being a conjunct of k , is entailed by k . So *PIC Dretske* is necessarily false. [165, p. 394]

Similarly, Baumann claims that $\Pr(\neg h|k) < 1$ should be replaced by $\Pr(\neg h|k) = 1$. He claims that since "ordinary propositions entail the denial of skeptical hypotheses, we can say that" [17, p. 407]:

$$\bullet k \rightarrow \neg h$$

which together with the probability calculus gives us

- $\Pr(\neg h|k) = 1$

So basically, since $\neg h$ has a probability of 1 prior to the signal, the signal r does not raise its probability to 1. “Even though the subject receives information that p and knows that $p \rightarrow \neg h$, she still does not receive information that $\neg h$ ” [17, p. 407].

Whilst Shackel and Baumann offer some manoeuvres which would provide a way out of Jager’s problem, their perhaps somewhat awkward defences are not very helpful for Dretske and not in line with his overall picture. The issue they are looking to resolve has not been formulated in the right way. Even though it is the case that the signal r need not carry the information that $\neg h$, according to their arguments this is because the agent in question already possess that information. But Dretske is going to want a different form of closure failure, according to which one can be informed that e and informed that $e \supset \neg h$ without at all being in a position to be informed that $\neg h$ due to the signal r not carrying this information at all because of the relevant alternatives analysis (not just that the signal does not inform the agent that $\neg h$ because they already have this information).

This is further evidenced by a subsequent case that Shackel makes, according to which, Dretske is committed to closure for a modified version of closure, a general closure principle for signalled information:

Closure for Dretske Information: If a signal, r , gives you Dretske information that p , you know that p entails q and q is available for signalling, then r gives you Dretske information that q .

Formally, q being available for signalling simply means that $\Pr(\neg h|k) < 1$ is added as a premise. According to Shackel, this is Dretske’s real issue, as there are cases where this principle will give closure for a skeptical hypothesis, which is available for signalling, even though Dretske doesn’t want this. Consequently, Dretske would be committed to closure for some cases of knowledge where he is against closure.

Dretske responds in turn to Jager’s (and Shackel’s) objection. Whilst Shackel and Baumann deal with Jager’s technical argument, Dretske argues that they miss the point. It is because of Dretske’s system of relevant alternatives that information is not closed. It is because of heavyweight propositions and the fact that our methods of gaining the information required for knowledge generally do not provide the information about their own information carriage.

Dretske provides some good examples in explaining why the closure of information does not always hold:

... tree rings carry information about the age of a tree *given that the past is real*. They don't provide information about the reality of the past. The reality of the past is, if you will, a condition in which things (newspapers, photos, fossils, tree rings) carry information about the past. It isn't itself a condition about which these things can carry information - at least not when they are carrying information about details of the past (the age of this tree, what you had for breakfast this morning, what Suzy looked like when she was a baby). That is why you can't read the newspaper to find out whether the past is real although the newspapers supply you with a lot of information about details of the past, details that imply there was a past. That is why a voltmeter's pointer pointing at the numeral "5" provides information that the voltage is 5 without providing information - logically implied by the information it does provide - that the meter isn't misrepresenting a voltage of 4 as 5. [58, p. 411]

So information-bearing signals need not be self-authenticating; a signal will not carry the meta-information that it is carrying some information x given that it does not provide information about the status of the conditions required for x .

If the motivations and foundations for Dretske's account are accepted (which of course, they need not be), then I think that Dretske's response to this debate is successful and shows why Jäger's objections miss the mark. But whilst Dretske is able to defend his scheme, his formal definition of information flow is not powerful enough to adequately serve it. Ultimately, his probability function is not enough to capture the rich idea of relevant alternatives. So rather than upheaving things and developing "a theory of perceptual information that differs considerably from the one Dretske proposes" [111, p. 195], what is actually needed is a replacement of Dretske's formal apparatus with something such as the multi-modal framework, which can adequately do the job.

A consequence of this analysis is that at some level transitivity must fail. As was mentioned earlier, Dretske endorses the Xerox principle: $A \sqsubset B \wedge B \sqsubset C \vdash A \sqsubset C$. This is the only reference he makes to transitivity and would seem to imply that transitivity universally holds. But as I have established, there is a difference between the Xerox principle and transitivity in general, with the former being a special case of the latter. With this point, we can associate the failure of information closure with the failure of information transitivity.

Let us investigate all of this by returning to the zebra example, beginning with the following proposition assignments:

- V = zebra-like visual signal
- Z = is a zebra
- M = painted mule

Now, in a Dretskean logic of being informed, the failure of information closure would be represented with:

$$Iz \wedge I(z \supset \neg m) \not\vdash I\neg m$$

This invalidity goes hand in hand with an invalidity of the following form in a logic for information flow:

$$(V \sqsubset Z) \wedge (Z \sqsubset \neg M) \not\vdash V \sqsubset \neg M$$

Here is why. The visual signal I am receiving of a black and white striped equid carries the information that the creature before me is a zebra. Something being a zebra carries the information that it is not a mule painted to resemble a zebra. Does it follow from this that the visual signal I am receiving carries the information that the creature before me is not a mule painted to resemble a zebra? Dretske would want to say no; in effect he would have to deny the transitivity of information flow in this case. For in the zebra example, if it were to be the case that $V \sqsubset \neg M$ holds, then when V is received it would carry with it the information that $\neg M$. But if an agent could get the information $\neg M$ from a signal V , then they could come to know $\neg M$ by forming the appropriate belief, something Dretske can't have.

But with Dretske's probabilistic definition of information, we have:

- $\Pr(Z|V) = 1$
- $\Pr(\neg M|Z) = 1$
- $\therefore \Pr(\neg M|V) = 1$

So a different formal apparatus is required. Of course under a system with one modal operator transitivity would also hold: $\Box(V \supset Z) \wedge \Box(Z \supset \neg M) \vdash \Box(V \supset \neg M)$. Naturally, this leads to an application of the multi-modal logic framework. We start with the following:

- $\Box_x(V \supset Z)$
- $\Box_y(Z \supset \neg M)$
- $\Box_z(V \supset \neg M)$

Clearly transitivity is no longer inevitable:

$$\Box_x(V \supset Z) \wedge \Box_y(Z \supset \neg M) \not\vdash \Box_z(V \supset \neg M).$$

How x , y and z are assigned depends on one's approach to setting relevant alternatives. We now turn to look at these approaches.

Dretske and Relevant Alternative Accounts that Deny Closure

For Dretske, $\Box_x(V \supset Z)$ is not going to range over irrelevant alternatives with the non-actual possibility that the zoo has a painted mule. $\Box_y(Z \supset \neg M)$ will presumably be able to range over all alternatives, since it is an analytic truth. $\Box_z(V \supset \neg M)$ will presumably once again range over another set of alternatives, ones where the possibility of a mule being in the zoo's enclosure are counted. Whilst 'mule' is not relevant to 'zebra', it is relevant to 'not-mule'. So according to this Dretskean picture, we have the following:

- $\Box_y A \supset \Box_x A$
- $\Box_y A \supset \Box_z A$
- $\Box_x(V \supset Z) \wedge \Box_y(Z \supset \neg M) \not\vdash \Box_z(V \supset \neg M)$

Since $x \neq z$, the information flow is not transitive and V does not carry the information that $\neg M$. $\Box_x(V \supset \neg M)$ is derivable, but since $\neg M$ is a channel condition relative to the set indexed by x , this does not represent an information relation.

In determining the relevant alternative indexes for an argument, it seems that both the proposition about which information is being carried and the signal that is carrying the information need to be factors.

If $\neg M$ were to be replaced by $\neg T$ (not-tiger), then we should be able to get something like:

- $\Box_x(V \supset Z) \wedge \Box_y(Z \supset \neg T) \vdash \Box_x(V \supset \neg T)$

Also, if the visual zebra signal were to be replaced by a DNA test that confirms that the creature is a zebra, then we should be able to get something like:

- $\Box_z(DNA \supset Z) \wedge \Box_y(Z \supset \neg M) \vdash \Box_z(DNA \supset \neg M)$

What is required with this account of scenarios such as the standard zebra-mule one is an explanation of why it is legitimate to shift the set of relevant alternatives. As we have seen, one reason for this is that it affords a way to counter the skeptic and maintain knowledge of ordinary propositions whilst accepting that we do not know skeptical heavyweight propositions.

As will be covered in the next section, contextualists such as Stine [172], who accuse Dretske of cross-context equivocation, prefer another way around the skeptic ³¹. They posit an interpretation of the relevant alternatives idea whereby each of the premises involved in the evaluation of a knowledge statement (or information in our case) shares the same set of relevant alternatives. In a low context this set is one in which not-mule is eliminated and in a high context this set is one in which not-mule is included. Depending on the context, either $V \sqsubset Z$ fails or $V \sqsubset \neg M$ holds. What determines which set of relevant alternatives is used is not “certain features of the putative subject of knowledge (his/her evidence, history, other beliefs, etc.) or his/her objective situation (what is true/false, which alternatives to what is believed are likely to obtain, etc.), but rather features of the knowledge *attributor(s)*’ psychology and/or conversational-practical situation” [160].

This dominant form of contextualism, ‘attributor contextualism’, means that in ordinary discourse the claim that one knows the animal is a zebra (based on $V \sqsubset Z$) and the claim that one knows it is not a mule are both true. In a skeptical discourse on the other hand, both claims are false. Without any change in the subject’s situation, change in the attributor’s conversational-practical situation means that the subjects go from knowing to not knowing.

This is a fairly crude outline of the contextualist approach, but it will serve as a reference point to further show why Dretske thinks that it is legitimate to shift the set of relevant

³¹Heller [98] offers a good explication of Dretske’s type of relevant alternatives account that answers opponent contextualists such as Stine.

alternatives. Dretske adopts what he terms ‘modest contextualism’, whereby the set of relevant alternatives for a proposition is determined by certain features of the putative knower, not the attributor. Unlike ‘radical contextualism’, the attributor’s conversational-practical situation should not ‘rob’ subjects of knowledge.

It is this feature of contextualism that Dretske finds disagreeable [55, 56]: if someone can be robbed of their knowledge simply by the raising of a question which augments the realm of relevant possibilities, which ‘increases the stakes’, then something is amiss. Can increasing the set of relevant possibilities just by contemplating skeptical possibilities really result in someone saying that they never knew some proposition which they claimed to once know?

So if one wants to maintain knowledge of ordinary propositions and accept not knowing skeptical propositions whilst *not* contextually relativising knowledge claims in such a way then shifting relevant alternative sets across different propositions/information relations is the way to go. By denying closure, if it is true that a subject knows an ordinary proposition based on the right information given their context, this knowledge claim stays true; it does not have to be given up.

As well as modifications to the technical apparatus for Dretske’s definition, another related part of Dretske’s account that would need to be revised is information nesting as mentioned in Section 4.1:

... if a signal carries the information that s is F , it also carries all the information nested in s ’s being F

The information that t is G is nested in s ’s being $F = s$ ’s being F carries the information that t is G . [51, p. 71]

He also distinguishes between information that is *analytically* nested and information that is *nomically* nested. As the zebra-mule example shows, caution must be exercised with nesting and its applicability regulated, for although ‘not-mule’ is analytically implied by ‘zebra’, the information that ‘not-mule’ is not necessarily nested in the information that ‘zebra’.

In closing this section it is important to re-emphasise that the Dretskean rejection of closure only applies to certain cases; it is certainly not the case that this position is anti-closure. We can generally still gain information/knowledge of some ordinary proposition B by inferring it from the information/knowledge of some A . For example, replace ‘mule’ with

‘tiger’ and you get information transitivity and knowledge closure in the zoo scenario. So in the majority of ordinary cases information/knowledge will be preserved under entailment. Indeed,

some of these reactions [to closure rejection] are, I think, a bit overdone. To deny closure is not to say that you can *never* know (find out, discover, learn) that Q is true by inferring it from a P that you know to be true. It is merely to deny that this can be done for any Q . [56, p. 17]

Contextualism and Relevant Alternative Accounts that Accept Closure

Contextualism provides another way to look at information and knowledge.³² Whilst there are different specific versions of contextualism, the general idea is that attributions of ‘knows that’ are made relative to contexts, so that they might be true in some relaxed contexts and false in some stronger contexts. As DeRose puts it:

the truth-conditions of knowledge ascribing and knowledge denying sentences (sentences of the form ‘S knows that P’ and ‘S doesn’t know that P’ and related variants of such sentences) vary in certain ways according to the contexts in which they are uttered. What so varies is the epistemic standards that S must meet (or, in the case of a denial of knowledge, fail to meet) in order for such a statement to be true. [45, p. 2]

So epistemological contextualism is a semantic thesis about how ‘knows that’ operates, not a theory of what knowledge is. It is embedded into a theory of knowledge, with some parameter(s) in that theory being made contextually variable. Relevant alternative theories of knowledge can be contextualised simply by making the set of relevant alternatives contextually variable. Against a certain set of relevant alternatives the subject might have ruled out all of the alternatives so that knowledge can be attributed to them. Against another higher-standard set they might not be able to rule out all relevant alternatives and so cannot be said to have knowledge.

A typical relevant alternatives contextualist about knowledge would treat closure as follows. Take the following formulation of the skeptical argument in terms of the zebra-mule scenario.

³²By contextualism here I generally mean contextualism of the attributor variety. See [160] for an overview.

1. John knows that there is a zebra before him
2. If John does not know that what is before him is not a cleverly disguised mule then he does not know that it is a zebra
3. John does not know that what is before him is a cleverly disguised mule

1,2 and 3 are mutually inconsistent. Dretske wants to hold onto 1 and 3 and reject 2, thereby rejecting closure.

The contextualist maintains that 1, 2 and 3 do not actually conflict. Instead, they retain closure and selectively reject one of 1 or 3. This inconsistency can be resolved by judging the truth/falsity of these statements relative to the one context. In ordinary standards contexts 3 is false whilst 1 and 2 are true. In skeptical contexts the higher epistemic standards are not met and 1 is false whilst 2 and 3 are true.

If we are never actually faced with a conflict between these three statements, why does it seem like there is an issue? A contextualist would respond by saying that:

since we most often find ourselves in low-standards contexts, we tend to evaluate knowledge attributions according to the epistemic standards that are in place in those contexts. Thus, we tend to reckon (1) true. However, since (3) makes explicit reference to BIVs, our evaluation of that claim tends to lead us to entertain the BIV skeptical scenario. Doing this can raise the epistemic standards it can push us into a context in which the epistemic standards are quite high - and so we tend to reckon (3) true. And so it seems that we are faced with a conflict between (1), (2), and (3). Yet it merely seems as if we are faced with such a conflict. For, as we have seen, when the epistemic standards are high, (1) is false while (3) is true. But when the standards are lower, (1) is true while (3) is false. [20]

There are two ways things can work here. One is to say that in ordinary contexts one can know zebra and know not-mule. The former is knowledge based on visual information and the possibility of not-mule does not preclude this knowledge as it is an irrelevant possibility. The latter is knowledge based on the fact that not-mule is not ordinarily relevant. When not-mule becomes relevant both pieces of knowledge go. I take this to be the standard approach to relevant alternatives contextualism.

Another way is to say that in ordinary contexts one can know zebra given their zoo experience and that not-mule is irrelevant. Based on this knowledge they can infer many other

ordinary pieces of knowledge such as not-tiger, not-giraffe, etc. But once a knowledge statement with not-mule is entertained, not-mule becomes relevant and hence the knowledge that zebra is lost. Closure is maintained with this approach because $K(z \supset \neg m)$ automatically implies a context in which not-mule is relevant and therefore $\neg Kz$ and $\neg K\neg m$.

In the spirit of epistemological contextualism it follows that we can entertain the possibility of treating information flow (and being informed) in a contextualist manner. Thus we shall continue this discussion in terms of information flow. Applying the modal logic framework to the zebra-mule scenario as a contextualist, let 1 and 2 index two different contexts or sets of relevant alternatives, such that the set indexed by 1 is a proper subset of that indexed by 2. We have:

- $\Box_1(V \supset Z) \supset \Box_1\neg M$
- $\Box_1(V \supset Z)$
- $\Box_2(Z \supset \neg M)$
- $\Box_x(V \supset \neg M)$

In the lower standards context indexed by 1, x is set to 1 and $\Box_1(V \supset \neg M)$ is derivable. There are a few ways to treat this derivation. As we have seen, this cannot represent an information relation for Dretske since $\neg M$ is a channel condition. In terms of a contextualist framework, it does not represent an information relation proper since $\Box_1\neg M$. But it could be considered a vacuous or default information relation relative to $\Box_1$ ³³. At any rate, $\neg M$ can be considered a piece of information within this context since it is true in every relevant alternative.

In the higher standards context, x is set to 2, so that the information $\neg M$ cannot be gained period:

- $\Box_1(V \supset Z) \wedge \Box_2(Z \supset \neg M) \not\vdash \Box_2(V \supset \neg M)$
- $\Diamond_2 Z \wedge \Diamond_2 M$
- $\Diamond_2(V \wedge \neg Z)$
- $\Box_2(V \supset Z \vee M)$

³³One idea is to introduce an independent $\Box$ that ranges over all alternatives. This could distinguish contextually set conditions from necessary truths: $\Box_1 \top \wedge \Box \top$ whereas $\neg \Box \neg m \wedge \Box_1 \neg m$

This contextualist idea can be used to devise a practical way of speaking about information flow in which statements of the form ‘*A* carries the information that *B*’ are associated with a set of background relevant alternatives. In order to get a better idea of how this strategy might work, we start with some objections to contextualism and consequent contextualist responses.

Yourgrau argues that the following dialogue exemplifies a contextualist scenario [180, p. 183]:

- Person A: Is that a zebra?
- Person B: Yes, it is a zebra.
- Person A: But can you rule out its merely being a cleverly painted mule?
- Person B: No, I can’t.
- Person A: So you admit you didn’t know it was a zebra.
- Person B: No, I *did* know then that it was a zebra. But after your question, I no longer know.

Person A’s utterance here has shifted the standards for knowledge from one context to another. But if nothing about Person B’s epistemic state changes then why should the simple mention of a skeptical possibility mean that they no longer have knowledge that they once had?

A standard contextualist response to this is that once the standards for knowledge are raised, one cannot say ‘I used to know that it was a zebra’: “Once the standards have been raised, B cannot both attribute knowledge to himself in the past and deny knowledge to himself in the present. He should now only deny himself knowledge; once the standards have been raised, neither B’s past self nor his present self knows that this is a zebra”. [20]

Whilst it addresses Yourgrau’s objection, for some, this type of conversational contextualism, in which there is only one active context as dictated by the possibilities that have been raised, fares little better. As mentioned earlier, Dretske in particular is uncomfortable with such an approach. To use one of his examples in arguing against (radical) contextualism, take Clyde, who can distinguish oranges from tangerines but cannot distinguish oranges from wax imitation oranges. If just oranges and tangerines are considered, then when Clyde sees an orange, this is enough to give him the knowledge that there is an orange before him. When wax imitation oranges are considered too, the visual signal is not enough.

Simply by raising this question, we have moved to a context, a skeptical context, in which Clyde not only doesn't any longer know they are oranges, but a context in which he can truly deny having known it before we asked the question. We've moved to a context in which (it is true to say that) Clyde never knew they were oranges. [55, p. 181]

This type of shift and its consequences can understandably not sit well. The adoption of a contextually flexible approach which can also preserve context-relative knowledge attributions is thus preferable.

I do not see the harm in simply associating a context with each knowledge statement. Rather than setting one active context and then judging each knowledge/information claim relative to it, for each knowledge/information claim there should be some context against which it is made. Unlike Dretske's account however, there is more than one acceptable candidate set of contextual relevant alternatives for each knowledge/information claim. Call this a *relativist approach* to information/knowledge.

As we have seen, with the zebra-mule example Dretske says that 'zebra' has a fixed set of relevant alternatives x and 'not-mule' has a fixed set of relevant alternatives y , with $x \neq y$. Given these sets, we have $\Box_x(V \supset Z)$ but not $\Box_y(V \supset \neg M)$. The contextualist will say that one of x or y is the active context. Against x we have $\Box_x(V \supset Z)$ and $\Box_x(V \supset \neg M)$ but against y we have not $\Box_y(V \supset Z)$ and not $\Box_y(V \supset \neg M)$.

Alternatively, by qualifying each information/knowledge statement with a set of relevant alternatives there is no problem in simultaneously asserting $\Box_x(V \supset Z)$ and not $\Box_y(V \supset Z)$; they are not mutually exclusive. In an ordinary context one should be able to go to the zoo and say that they know there is a zebra in the enclosure with the implicit qualification that they are ruling out the possibility that it is a disguised mule. If the possibility that it is a disguised mule is included then they do not know. But the sense in which they do know that it is zebra remains; it does not vanish simply because other skeptical possibilities are entertained. This is a feasible way to deal with information flow and knowledge.

In explicating their theory, contextualists have suggested that "know' either is or functions very much like an indexical, that is, an expression whose semantic content (or meaning) depends on the context of its use" [20]. For example, 'I' is an indexical, as its meaning depends on the context of its use and who is using it. But unlike uses of an indexical such as 'I', the term 'knowledge' seems more amenable to simultaneous cross-contextual application. If I am at the zoo enclosure marked 'zebras' and utter 'I am at the zebra enclosure' the expression is true. I cannot also truthfully utter 'I am not at the zebra enclosure', because

the context determines that this expression is false. I can though simultaneously utter ‘I know that there is a zebra’ relative to x and ‘I do not know that there is a zebra’ relative to y .³⁴

Whilst a Dretskean relevant alternatives account and a contextualist relevant alternatives account differ on how they use relevant alternatives, their procedure for determining relevant alternatives can share criteria. Take Cohen’s relevant alternatives contextualism. According to Cohen,

an alternative (to [some proposition] p) h is relevant (for [some person] S) = df
 S ’s epistemic position with respect to h precludes S from knowing p . (Cohen
 1988, p. 101)

He incorporates two criteria of relevance, one based on external conditions and one based on internal conditions:

External An alternative (to p) h is relevant if the probability of h conditional on reason r and certain features of the circumstances is sufficiently high [where the level of probability that is sufficient is determined by context]. (Cohen 1988, p. 102)

Internal An alternative (to p) h is relevant if S lacks sufficient evidence [where the amount of evidence that is sufficient is presumably determined by context] to deny h , i.e., to believe not- h (Cohen 1988, p. 103),

According to Cohen, both the internal criterion and the external criterion are sensitive to context. To this end, he takes the standard contextualist line that raising the stakes or highlighting skeptical possibilities allows certain alternatives to become relevant. Of course,

³⁴The adoption of a system in which each information/knowledge statement is relativised and paired with a variable, contextually determined set of relevant alternatives was suggested in early critiques of Dretske’s relevant alternatives approach. Foley suggests that that notion of a relevant alternative in Dretske’s account “is an ad hoc device to make sure that the account implies we have knowledge [information] in ordinary cases where we are inclined to think we have knowledge [information]. The notion does nothing to explain why the skeptic is wrong in suggesting that we may lack knowledge in such cases” [78, p. 181]. Rather than trying to deal with the skeptic by employing a fixed set relevant alternatives, it would just be better to make an admission and “relativize the notion of knowledge. It might be admitted that the skeptic is right to say that relative to all the possibilities the probability of there being a barn in front of S given that he sees in front of him an object which looks like a barn, is not 1. So, given Dretske’s account, S does not know relative to this set of possibilities that there is a barn in front of him. On the other hand, relative to a smaller set of possibilities (ones which we think to be most likely and hence ones which we *think* to be most relevant), the probability of there being a barn in front of him might be 1. And so, relative to this smaller set of possibilities S might know that there is a barn in front of him” [78, p. 181]. Whilst Foley’s dismissal of Dretske’s notion of determinate relevant alternatives is a little too cursory, it does highlight an understandable tendency towards eschewing this more involved notion in favour of a relativised usage.

Dretske would not agree that the standards of relevance can shift in this way. What he would endorse is the involvement of external and internal criteria. Regarding external conditions, if it was the case that the zoo sometimes did put disguised mules in the enclosure, then ‘mule’ would become a relevant alternative³⁵. Regarding the internal condition, if a person was not aware of the fact that zoos do not keep mules then ‘mule’ might be relevant. But under a Dretskean account these external and internal criteria would not be sensitive to contexts that are conversationally and practically determined; once the relevant alternatives for a proposition are set based on the internal and external conditions of an epistemic agent’s situation, they cannot simply change in such a way.

Cohen’s criteria could be applied to the relativist approach I have suggested, with each statement of information flow being relativised and paired with a variable, contextually determined set of relevant alternatives. For the external criterion, a set of probability thresholds would determine a set of relevant alternative sets. For the internal criterion, a set of evidence thresholds would determine a set of relevant alternative sets. These would be combined to determine the sets of relevant alternatives that could be paired with an information flow statement.

Apart from adopting criteria such as Cohen’s, I would like to suggest another condition to be incorporated into the relativist relevant alternatives framework for information flow. This condition, which would place a lower limit on sets of relevant alternatives, comes in response to certain uses of the notion of contextualism and sides more with Dretske’s account.

Lottery-style paradoxes³⁶ present an opportunity to initiate this matter. Dretske discusses the lottery paradox briefly in [51, p. 99]. According to his information-theoretic analysis one cannot know that they will lose a one million ticket lottery since they do not have the information that they will lose. Let s stand for the signal constituted by a description of the lottery situation. Since s is all the contestant has and $\Pr(\text{‘lose’}|s) < 1$, they do not have the information that they will lose.

Contextualists such as David Lewis have argued that there are contexts in which it is fine to say that someone knows that they will not win the lottery. It is perfectly fine to have contexts in which a subject knows they will not win the lottery because every alternative in which they do win the lottery is ruled out as irrelevant.

Support for this approach comes when the lottery proposition is connected to another proposition. As covered in [125], this can be presented as Lewis does:

³⁵A great anecdote here is that of a Palestinian zoo that a few years ago could not afford to replace its deceased zebras and so added some donkeys painted to look like zebras to its collection. To find out more search for ‘Palestinian zoo zebras’ on the internet.

³⁶Also discussed in Chapter 3, Section 3.5.3

Pity poor Bill! He squanders all his spare cash on the pokies, the races, and the lottery. He will be a wage slave all his days. We know he will never be rich. But [...] if we know that he will never be rich, we know that he will lose [the lottery]. But [...] we cannot know that he will lose. All the possibilities in which Bill loses and someone else wins saliently resemble the possibility in which Bill wins and the others lose; one of those possibilities is actual; so by the rules of Actuality and Resemblance, we may not properly ignore the possibility that Bill wins. [124, p. 134]

The apparent issue here consists of the following inconsistent triad:

1. We know that Bill will never be rich
2. If we know that Bill will never be rich, then we know that he will lose the lottery
3. We do not know that Bill will lose the lottery.

Lewis' intended solution to this issue is his brand of contextualism:

Salience, as well as ignoring, may vary between contexts. [...] When [...] explaining how the Rule of Resemblance applied to lotteries, [...] the resemblance between the many possibilities associated with the many tickets was sufficiently salient. But [...] when we were busy pitying poor Bill for his habits and not for his luck, the resemblance of the many possibilities was not so salient. At that point, the possibility of Bill's winning was properly ignored; so then it was true to say that we knew he would never be rich. Afterward I switched the context. I mentioned the possibility that Bill might win, wherefore that possibility was no longer properly ignored. [...] it was no longer true that we knew he would lose. At that point, it was also no longer true that we knew he would never be rich. [124, p. 134]

The general idea here seems to be that when one is in such a lottery case, alternatives in which they win the lottery given only a one in a million chance are ruled out as irrelevant in certain contexts. Given this, there would also be certain contexts for which alternatives in which they win the lottery given only 1/999,999 chance are ruled out as irrelevant too. But where does this contextualist threshold relaxation stop? There is the implication that there is a context in which they could know that they will not win the lottery given a 1/2 chance.³⁷

³⁷See [125] for a detailed discussion of this.

With the account of information flow I have in mind there is something wrong with permitting contexts in which the signal constituted by a description of the lottery situation carries the information that they will lose. Obviously, if it actually is the case that the contestant has the winning ticket then they cannot possibly receive the information that they do not have the winning ticket. But if they do not have the winning ticket, then what is the problem with excluding the alternative in which they have won?³⁸ As just mentioned, for this to be the case, in such a context, alternatives in which they win the lottery given only a one in a million chance would have to be ruled out as irrelevant. But given that there is another signal, namely the result of the lottery, which is readily available and would rule out the alternatives in which they win, I do not see a good reason to contextually dismiss the possibility of them winning as irrelevant.

More generally, I think that the criterion has to do with a consideration of the types of signals that are available at a certain level and whether the level is legitimate given the restrictions it involves. In the zebra-mule scenario, the standard level is that in which one goes to the zoo and knows that there is a zebra by seeing it. In this ordinary context, the zebra-like visual signal carries the information that ‘zebra’. At this level, since there is no other signal that one can readily or feasibly receive which would be able to distinguish between ‘zebra’ and other possibilities such as ‘disguised mule’, such possibilities can be legitimately ruled out as irrelevant. On a higher level, involving zoologists who have access to DNA testing, ‘disguised mule’ becomes a relevant alternative because they have access to a signal that can rule out such an alternative. Since these two different sets of relevant alternatives are acceptable given their level, both can be used and paired with a knowledge/information statement.

With regards to the zoo scenario, the following would by contrast be an illegitimate contextual exclusion of alternatives at any level. Suppose that the zoo-goer, who has no idea what animal is in the next enclosure they are walking to, is informed by a zookeeper that the animal in the next enclosure is an animal of the African Savannah. If a restriction were implemented in which all other Savannah animals (giraffes, elephants, ...) were excluded as irrelevant, then the zookeepers utterance would carry the information that the animal is a zebra. But given the fact that a visual signal which can rule out these possibilities is readily available to the zoo-goer, such contextualisation would be illegitimate.

Returning to the issue of the inconsistent triad of claims regarding Bill and the lottery, we saw that Lewis’s strategy was to implement a rather liberal contextualism, in which contexts can shift simply by entertaining possibilities. Another option would be to deny closure and say that we know that Bill will never get rich but that we do not know that

³⁸Given the equality of each ticket, if the possibility of one winning is excluded, then why not all of them? This would lead to an incoherent situation in which no relevant alternative has a winning ticket. This is a point particular to lottery-style situations though.

he will lose the lottery, even though the former implies the latter. But this is ill-suited and unnecessary. Firstly, Dretskean closure rejection has a different basis. Secondly, in examples of Dretskean closure the options are not exhaustive and what is known implies what is not known but the negation of what is known does not imply the negation of what is not known (e.g. $z \supset \neg m$ but not $\neg z \supset m$). On the other hand, with these lottery examples, not-rich implies lottery-loss and rich implies lottery win.

Given the relativist approach to information that I have outlined and the limits that are placed on what counts as an acceptable set of relevant alternatives, neither of the propositions ‘not rich’ and ‘not win’ are known. Unlike the zebra-mule scenario, there is a potential signal at the standard level that carries both the information that the contestant will not be rich and the information that the contestant will lose the lottery; namely a signal that provides the results of the lottery. So as Dretske would, we deny both the knowledge that the contestant will not be rich and the knowledge that they will not win the lottery.

4.4.4 Variable Relevant Alternatives and the Other Information Flow Properties

Since bringing variable contexts/relevant alternatives into the picture, the focus has mainly been on transitivity and to a lesser extent monotonicity. In this section I consider the effects that such variability has on conjunction and disjunction properties, with a particular focus on some of the repercussions of a Dretskean relevant alternatives account. The results of this investigation will be applied in the next chapter to the development of epistemic logics based on information-theoretic epistemology; if information leads to knowledge then the properties of information flow can have something to say about the properties of knowledge.

Simply due to their structure, the Reverse-Conjunction and Reverse-Disjunction principles have less room to be affected by shifting contexts:

- **Reverse-Conjunction:** $\neg \Box_x(A \supset (B \wedge C)) \vdash \Box_y(A \supset B) \wedge \Box_z(A \supset C)$
- **Reverse-Disjunction:** $\neg \Box_x((A \vee B) \supset C) \vdash \Box_y(A \supset C) \wedge \Box_z(B \supset C)$

Given that there is only one context x in the antecedent, it follows that $y = x$ and $z = x$ are natural assignments. If we break up an antecedent that ranges over a certain set of relevant alternatives then the information flow relations captured in the consequent conjuncts can simply range over that same set of relevant alternatives.

There is one point related to the Reverse-Conjunction principle that I would like to mention here since it will help to further elucidate things. Kripke levelled a purported counterexample against Nozick's tracking theory of knowledge which at first might also seem like a successful counterexample to Dretske's account.³⁹ In the 'red barn' example Bob is travelling through fake barn county, in which real barns and fake barns are distributed. Since for some reason it is not possible to fake a red barn all the red barns are real. If Bob sees a red barn what information does he receive?

Since in all relevant alternatives every time something looks like a red barn (LRB) it is a red barn, Bob receives the information that the object is a red barn ($Bo \wedge Ro$). Does it follow from this that he receives the information that the object is a barn: Bo ? Well, some like Kripke might say that since there are relevant alternatives in which something looks like a barn (LB) but is not a barn, then Bob is not receiving the information that the object is a barn. Given Dretske's information-based definition of knowledge, this would lead to the problematic conclusion that Bob can know that $Bo \wedge Ro$ but not know that Bo . The problem with this purported counter-example is that it switches from red barn to just barn when considering what visual signal Bob receives, when it should just be fixed on red barn. The information that Bo should be judged by the signal LRB , not the signal LB :

- $LRB \sqsupset (Ro \wedge Bo)$
- $(Ro \wedge Bo) \sqsupset Bo$
- $LRB \sqsupset Bo$

Due to their structure, the Conjunction and Disjunction principles have more room for investigation. With both of these principles as given below, it can certainly be the case that $x \neq y$; the whole point of this framework is that assertions of information flow relations do not need to be made against the same set of relevant alternatives. Given this, these principles are not logically valid, even if $z = x$ or $z = y$:

- **Conjunction:** - $\Box_x(A \supset B) \wedge \Box_y(A \supset C) \not\vdash \Box_z(A \supset (B \wedge C))$
- **Disjunction:** - $\Box_x(A \supset C) \wedge \Box_y(B \supset C) \not\vdash \Box_z((A \vee B) \supset C)$

Beyond this logical failure, is it possible to come up with actual and coherent examples of information flow failing the Conjunction or Disjunction principles?

³⁹Dretske [54, p. 117] briefly discusses this example and offers the case for why it does not affect his account, even if it affects Nozick's. Adams [4, p. 8] offers the case for why Kripke's example does not undermine Nozick's account either.

To begin with, let us first consider another principle:

Double Conjunction: $-(A \supset C) \wedge (B \supset D) \vdash (A \wedge B) \supset (C \wedge D)$

One thing to note about this principle is that even if the Conjunction principle were to hold, then the failure of this principle would mean that information-based knowledge, as will be given in the next chapter, is not closed under conjunction: $Kc \wedge Kd \not\vdash K(c \wedge d)$. For it could be the case that one comes to know C via A and comes to know D via B without receiving a signal that carries the information that $C \wedge D$.

Another thing is that consideration of this principle prompts consideration of the connective rules for relevant alternatives. Take the following scenario in which there are two perfectly functioning metal detectors, MD1 and MD2, which are being used by two amateur prospectors, P1 and P2 respectively. MD1 sounds (S1) and this carries the information that there is metal under P1's location (M1). A second later MD2 sounds (S2) and this carries the information that there is metal under P2's location (M2). So we have:

- $\Box_x(S1 \supset M1)$
- $\Box_y(S2 \supset M2)$

Now, x should range over all the standard alternatives in which MD1 is functioning correctly and y should range over all the standard alternatives in which MD2 is functioning correctly. If there are alternatives in which MD1 is functioning correctly and MD2 is defective and vice versa, $x \neq y$. Now, with the following:

- $\Box_x(S1 \supset M1) \wedge \Box_y(S2 \supset M2) \vdash \Box_z((S1 \wedge S2) \supset (M1 \wedge M2))$

what is the relevant set of alternatives indexed by z ? It cannot be the case that $z = x$ or $z = y$, for then it could be the case that $\Diamond_z((S1 \wedge S2) \wedge \neg(M1 \wedge M2))$. The only suitable option appears to involve making the relevant alternatives indexed by z a subset of both the relevant alternatives indexed by x and the relevant alternatives indexed by y : $(\Box_x \supset \Box_z) \wedge (\Box_y \supset \Box_z)$.

Leaving this aside, rather than focusing on Double Conjunction in particular, I will now cut straight to the Conjunction principle, showing in what ways it, and *a fortiori* Double Conjunction can fail.

Firstly, when $x \neq y$ it is no surprise that in some sense conjunction fails cross-contextually. Returning to the metal detector example, we saw that whilst a tone can signal metal relative to one context and signal functionality relative to another context, there is no context in which it can do both:

- T - metal detector sounds
- M - metal is present
- F - metal detector is functional
- $\Box_x(T \supset M) \supset \Box_x F \wedge \neg \Box_x M$
- $\Box_y(T \supset F) \supset \neg \Box_y F$
- Setting $x = y$ leads to contradiction, so there is no context in which $T \supset (M \wedge F)$

This metal detector example is a demonstration of what can technically be done with the variable relevant alternatives information flow framework. It is not a genuine challenge to the Conjunction principle though since it is not the case that both antecedent information relation statements can be made simultaneously; $T \supset M$ and $T \supset F$ are mutually incompatible and if one statement holds then the other fails to hold. This is because if some signal A carries the information that B given some channel condition C then A cannot also carry the information that C .

The conjunction principle is clearly going to hold under a contextualist account in which all given information relation statements are evaluated within the one context. On the other hand and more interestingly, as the following case illustrates it seems, perhaps somewhat surprisingly, that a Dretskean account of information flow means that the Conjunction principle will not always hold; Dretske cannot have an absolute Conjunction principle.

Returning to the zebra scenario, take the following two information relations:

- $V \supset Z$
- $V \supset Z \vee M$

In this example, $V \supset Z$ is as usual. $V \supset Z \vee M$ is made relative to a wider range of relevant alternatives, ones in which all animals, including mules are considered. With regards to the standard zoo scenario I do not see a problem in asserting both $V \supset Z$ and $V \supset Z \vee M$ under Dretske's account. All of this is represented in the following:

- $\Box_x(V \supset Z)$
- $\Box_y(V \supset Z \vee M)$
- the axiom $\Box_y p \supset \Box_x p$

From here it is legitimate to add the analytic relations:

- $\Box_y(Z \supset \neg M)$
- $\Box_y(M \supset \neg Z)$

from which can be derived

- $\Box_y(V \supset ((Z \wedge \neg M) \vee (M \wedge \neg Z)))$

So now we have:

- $V \sqsubset Z$
- $V \sqsubset ((Z \wedge \neg M) \vee (M \wedge \neg Z))$

Can these two information relations be joined as per the Conjunction principle to get $V \sqsubset (Z \wedge ((Z \wedge \neg M) \vee (M \wedge \neg Z)))$? Say there were a suitable set of relevant alternatives indexed by z against which this information relation were judged: $\Box_z(V \supset (Z \wedge ((Z \wedge \neg M) \vee (M \wedge \neg Z))))$. $\Box_z(V \supset (Z \wedge \neg M))$ is a consequence of this, from which it follows that $\Box_z(V \supset \neg M)$ is also a consequence. But this would then mean that $V \sqsubset \neg M$, something which Dretske does not want.

In general, for any account of $\sqsubset$ which has the following, such a consequence will follow:

1. $B \vdash C, C \vdash B, A \sqsubset B \Rightarrow A \sqsubset C$ (substitution of logically equivalent propositions)
2. $A \sqsubset (B \wedge C) \Rightarrow (A \sqsubset B) \wedge (A \sqsubset C)$ (reverse conjunction)

Rather than abandon either of these two fundamental properties, I think that it is the Conjunction principle which should go; the statement $V \sqsubset (Z \wedge ((Z \wedge \neg M) \vee (M \wedge \neg Z)))$ is not valid. The basic diagnosis is that the statements $V \sqsubset Z$ and $V \sqsubset (Z \vee M)$ are made

relative to different sets of relevant alternatives in such a way that they cannot be coherently combined. This line of thought will be continued in the next chapter where I suggest that one way to construct a Dretskean epistemic logic involves sacrificing the closure of knowledge under conjunction.

This is one example of how a principle can be affected under a relevant alternatives account that denies closure. Whichever of the above principles one decides to give up in order to ensure that $V \sqsupset \neg M$ or $K\neg m$ does not follow, it is clear that the Dretskean denial of information/knowledge closure has further repercussions for the logic of information/knowledge. It seems that this is something which Dretske did not fully appreciate.

This example feeds into a similar example that counters the Double Conjunction Principle proper. Let E stand for a statement asserting that the animal in question is an equid. Even if for some reason $V \sqsupset ((Z \wedge \neg M) \vee (M \wedge \neg Z))$ was denied, the statement $E \sqsupset ((Z \wedge \neg M) \vee (M \wedge \neg Z))$ is indisputable ⁴⁰.

Suppose that the zoo goer was told by a reliable zoo keeper that the animal in the enclosure is an equid. They look at the animal and see that it is a zebra. Whilst they have the information that Z and the information that $((Z \wedge \neg M) \vee (M \wedge \neg Z))$ from the signals V and E respectively, they do not have the information that is implied in joining them together:

- $V \sqsupset Z$ holds
- $E \sqsupset ((Z \wedge \neg M) \vee (M \wedge \neg Z))$ holds
- $V \wedge E \sqsupset (Z \wedge ((Z \wedge \neg M) \vee (M \wedge \neg Z)))$ fails

Showing how the Conjunction principle fails is the key point of this section. Before closing a quick look at the Disjunction principle is in order. The failure of the Conjunction principle, as demonstrated above with the zebra-mule example, logically carries over to the Disjunction principle. Continuing on from that example, we have:

- $\Box_x(V \supset Z)$
- $\Box_x(\neg Z \supset \neg V)$
- $\Box_y(V \supset Z \vee M)$
- $\Box_y(\neg(Z \vee M) \supset \neg V)$

⁴⁰In its full form this statement is an analytic truth. For our purposes this is a simplified version, since if something is an equid it could also be a horse or donkey. Adding these would not affect the argument though, since all the options are mutually exclusive.

Suppose that there were a z such that $\Box_z((\neg Z \vee \neg(Z \vee M)) \supset \neg V)$. Then the following results:

- $\Box_z((\neg Z \vee \neg(Z \vee M)) \supset \neg V)$
- $\Box_z(\neg(Z \wedge (Z \vee M)) \supset \neg V)$
- $\Box_z(V \supset (Z \wedge (Z \vee M)))$

from which $V \supset \neg M$ results as above.

I have here looked at three approaches to information flow based on the idea of relevant alternatives. The discussion started with Dretske's account before taking a look at contextualism and ending with a brief outline of a general relativist approach. The modal logic framework introduced in this chapter has been employed in our analysis of these three approaches. Whilst its employment has facilitated the analysis, the logic by itself is insufficient and extra-logical considerations need to be factored in when assessing arguments. Determining the principles of information flow endorsed by these accounts is not only a logical matter and depends on how indexes are assigned to modal operators and how certain elements are interpreted. Placing constraints on acceptable assignments can be used to determine a set of acceptable principles.

As we have seen, the contextualist approach works by fixing a context for evaluation and judging the argument (premises and conclusion) against that context. Therefore, the set of valid information flow relations for a contextualist approach can essentially be found in the set of corresponding valid arguments of the mono-modal logic, which in this case was **T**.

On the other hand, the relativist approach I briefly introduced towards the end gives little in the way of prescribing absolutely valid principles of information flow. Indexes can be liberally assigned with the imposition of some minimum extra-logical constraints and the modal logic is responsible only for assessing the argument's validity once the argument has been set up.

Dretske's account calls for the most 'customisation'. Going by the observations we have made throughout this chapter, the following restriction on modal operator indexing will serve to validate wanted properties and invalidate unwanted properties.

- For all arguments of the form $\Box_x A \vdash \Box_y B$ such that $A \vdash B$, $\Box_y C \supset \Box_x C$ is a theorem of the system.

Thus going by this condition, if we have an argument of the form $\Box_x(A \supset B) \vdash \Box_y(A \supset C)$ such that $B \vdash C$, then the relevant alternatives associated with x form a subset of the relevant alternatives associated with y : $X \subseteq Y$.

Another way to generate a suitable list of conditions in a mono-modal logic is to define $A \sqsupset B$ as $\Box(A \supset B)$ in the non-normal modal logic characterised by the following rules and axioms⁴¹:

- Propositional Calculus (PC)
- The axiom: $\Box A \supset A$ (T)
- The rule: $\frac{A}{\Box A}$ (N)
- The rule: $\frac{A \supset B}{\Box A \supset \Box B}$ (RM)

This gives us the following validities and invalidities:

- $A \sqsupset (B \wedge C) \vdash (A \sqsupset B) \wedge (A \sqsupset C)$
- $(A \vee B) \sqsupset C \vdash (A \sqsupset C) \wedge (B \sqsupset C)$
- $\vdash A \sqsupset A$
- $A \sqsupset B \vdash \neg B \sqsupset \neg A$
- $A \sqsupset B \vdash (A \wedge C) \sqsupset B$
- $A, A \sqsupset B \vdash B$
- $(A \sqsupset B) \wedge (A \sqsupset C) \not\vdash A \sqsupset (B \wedge C)$
- $(A \sqsupset C) \wedge (B \sqsupset C) \not\vdash (A \vee B) \sqsupset C$
- $(A \sqsupset B) \wedge (B \sqsupset C) \not\vdash A \sqsupset C$

Given this logic, it might seem that a probabilistic approach based on a high threshold of conditional probability could do the job. Take the following definition:

$$A \sqsupset B =_{df} \Pr(B|A) \geq 0.99$$

⁴¹Labelled EMNT by Chellas' naming convention [33, p. 207].

But as the following example shows, such an approach will not work. Going back to the zebra-mule counter-example to the conjunction principle, we make the following assignments:

1. $\Pr(Z|V) = 0.99$
2. $\Pr((Z \wedge \neg M) \vee (M \wedge \neg Z)|V) = 1$

Statement 1 expresses an approach that some might take; rather than employ the relevant alternatives strategy and assign a conditional probability of 1 against a certain set of relevant alternatives, instead make a probabilistic statement, according to which the probability that something is a zebra given that it looks like a zebra is very high but not 1, since there is a minute probability that some skeptical scenario could obtain. Statement 2 simply expresses the fact that the visual signal ‘zebra’ guarantees that the creature is at least a zebra or a disguised mule.

But it is a fact of probability that if $\Pr(B|A) = x$ and $\Pr(C|A) = 1$, then $\Pr(B \wedge C|A) = x$. Given this setup, the unwanted $V \sqsubset \neg M$ is a consequence:

- $\Pr(Z \wedge ((Z \wedge \neg M) \vee (M \wedge \neg Z))|V) = 0.99$
- $\therefore \Pr(\neg M|V) \geq 0.99$

As a result of context shifts and closure failure the logic behind Dretske’s information flow and epistemology will be anything but straightforward. Whilst this is not a problem in itself (and makes things more fun) an explication of the logic of information flow for Dretske’s account and what it entails is an important task. Beyond being of interest in shedding light on Dretske’s account, the results of the investigation in this chapter and the next could be of use in “[the] war [over closure that] has been raging in epistemology over the last few decades, [which] the good guys [(those who reject closure)] are losing” [98, p. 196].

4.5 Another Requirement on Semantic Information?

By now it is well-established that according to the definition of semantic information endorsed in this work, the following conditions must be met in order for σ to count as an instance of semantic information:

1. σ consists of one or more data

2. the data in σ are well-formed
3. the well-formed data in σ are meaningful
4. σ is truthful

Are these conditions sufficient as well as necessary? By connecting semantic information with environmental information another condition could be added to this definition. Whilst the resulting definition need not in general have exclusive claim to the term ‘semantic information’, as we will see it is this ‘stronger’ form of semantic information which will be of particular value.

Take the following example given by Floridi:

Epistemic luck does not affect informativeness negatively. To see why, one may use a classic Russellian example: if one checks a watch at time t and the watch is broken but stopped working exactly at $t - 12$ hours and therefore happens to indicate the right time $t - 12$ at t , one still holds the information that the time is t , although one can no longer be said to know the time. [71, p. 150]

This quote counters the idea of information flow covered in this chapter, in which there is a requirement of regularity for the presence of information flow. In fact elsewhere Floridi writes “is it possible that (1) S has the true belief that p and yet (2) S is not informed that p ? Barwise and Seligman seem to hold it is, I shall argue that it is not” [76, p. 41].

According to the account of information flow given in this chapter, the watch is not carrying information about the time, which implies that one cannot be informed that the time is such-and-such in this scenario. In fact with Dretske’s type of informational epistemology, the requirement of regularity for information flow is supposed to explain why knowledge (defined as information-caused belief) is not present in the above example.⁴²

This point has also been recognised by others:

holding a piece of well-formed meaningful data which incidentally happens to be true is not sufficient for being informed. This we see by considering Gettier-like cases in which true consequences can be derived from a set of (partially) false data. [9, p. 4]

⁴²Floridi does develop his own informational epistemology [76], but it is quite different to those such as Dretske’s.

Indeed, defining semantic information as true semantic content without such further qualification does lead to some questionable uses of ‘information’ and ‘informed’. For example, say someone guesses that there are 214 jelly beans in a jar and happens to be right. In terms of the first definition of information, this example is on par with the lucky watch example. Yet it seems questionable to say that they are informed and hold the information that there are 214 jelly beans in the jar.

Also, this definition leaves open the possibility of counting as information true semantic content that is generated from false semantic content. For example, valid reasoning with false or misinformative premises can still lead to true conclusions.⁴³

We thus have (1) a conception of semantic information and being informed that requires true semantic content *simpliciter* and (2) a conception of semantic information and being informed that requires environmental information. Denying (1) is not a necessary move. The conception of information broadly defined as true semantic content still accords with many intuitions about information, including its association with practical action and success. Also, with regards to instructional information, intuition suggests that true semantic content suffices. For example, if one correctly guessed the seven ingredients for a recipe for some dish, wrote them down on a piece of paper and handed that paper to a friend, that recipe would still count as instructional information on how to make the dish, irrespective of its aleatoric nature.

The contrariety between (1) and (2) will be addressed here by simply distinguishing these two types of information, with one being a subset of the other.⁴⁴

In order to make this distinction, let us use ‘weak’ semantic information (information_W) and ‘strong’ semantic information (information_S), with the latter representing a definition involving the new condition. So we can settle for this distinction without having to unnecessarily debate the two and information_W retains a place on the information map.

Nonetheless, the distinction is crucial. A failure to distinguish this ‘stronger’ conception of semantic information would detract from our toolkit of information concepts. As we have seen, given the manner in which it is generated, information_W in general cannot be considered an ‘epistemically legitimate’ form of information, given that such information should have a watertight connection to the truth. It is only information_W that can satisfy certain ideas about information and provide a certain bridge between information and knowledge.

⁴³Take the simple case where A is false, B is true and $A \supset B$ is true. One can still validly infer the true B from the presumed truth of A and $A \supset B$.

⁴⁴We want to say that when one acquires semantic information, they become informed, so it will not do to say that one can acquire semantic information but not be informed because the information was not generated in the right way.

So from this perspective semantic information would be something like true semantic content that is generated from or caused by environmental information. σ is an instance of semantic information (information_S) if and only if:

1. σ is information_W
2. σ is generated from environmental information

For the remainder of this work we will equate semantic information with information_S .⁴⁵ So if the watch in the above example were functioning correctly, then the person who checked the time would be informed of the time and the content of their true belief would be semantic information. In this context it is appropriate to refer to information_W as ‘mere’ true semantic content, thus paralleling a standard classification within epistemology:

	epistemology equivalent
semantic information	knowledge
true semantic content	true belief
misinformation	false belief

Following on from this classification, in the next chapter we will investigate the value of acquiring semantic information over true semantic content.

4.6 Conclusion

In this chapter an account of information flow was developed. The starting point of this investigation was Dretske’s account of information. Whilst core ideas from this account were retained in the development, several deficiencies and problems with Dretske’s account were addressed. Along the way a couple of alternative accounts of information flow were critiqued, an exercise which itself contributed to development of the account that was finally presented. It was shown that employing a system of multi-modal logic is a simple but effective way to capture information flow. Some of the main results of this account were:

- An analysis of the issue of information closure

⁴⁵Referring back to some of the discussion in Section 1.3.2 of Chapter 1, one thing to consider is whether information_S further suggests a conception of information that implies physical implementation.

- An explication and accommodation of approaches to information flow which are based on the notion of relevant alternatives
- An investigation into the properties of information flow

The last section of this chapter represents an essential theme of this thesis developed over the final two chapters, namely the significance of information_S. In even the most mundane aspects of life, we crucially rely on the reliability of our information sources. When one uses a beverage vending machine, the label for each button indicates what type of drink they will receive if they press that button. Imagine if this regularity was replaced by the random selection of drinks; if you want a bottle of water and press the button labelled water, the machine could return a bottle of water, but it could instead return a can of Sprite, or a bottle of orange juice, etc. In summary, it is not just that we acquire truths, but that we acquire them in an inherently consistent and reliable way. It is important that this fundamental aspect is accounted for in a formal framework for semantic information.

Chapter 5

Information and Knowledge

Information and knowledge are commonly associated with each other; colloquially, in dictionaries, the two terms are often treated as synonymous. Within philosophy however, information-theoretic or informational epistemology goes beyond this casual, colloquial association. It involves the development of specialised accounts of information and uses them to develop accounts of knowledge, to show how information causes or leads to knowledge. From this perspective, “information is a commodity that, given the right recipient, is capable of yielding knowledge” [51, p. 47]. With the investigation into information conducted in the previous chapter at hand, in this chapter we look at informational epistemology, using Dretske’s account as the basis for our investigation.

The goal of this chapter is to contribute towards the case for such an informational epistemology. The three main tasks of this endeavour are:

- An analysis of Dretske’s account and discussion of ways in which it could be modified/supplemented.
- A look into developing epistemic logics for Dretskean and contextualist informational epistemologies and applying the results to some issues.
- A proposal for answering the value problem for knowledge with informational epistemology.

5.1 Some Background

One of the main tasks of epistemology has been the analysis of propositional knowledge [170], a central task being to define what it is; what the necessary and sufficient conditions for propositional knowledge are. Traditionally knowledge has been defined as justified, true belief. According to this justified, true belief (JTB) analysis of knowledge, S knows that p if and only if:

1. p is true
2. S believes p
3. S is justified in believing p

Something like this definition can be found as far back as Plato¹. The truth and belief components are straightforward. Firstly, if a proposition is known then it is true; false propositions cannot be known. Secondly, if a proposition is known by a subject, then they must believe that proposition. Although an explication of the notion of justification is not straightforward, the basic idea in incorporating this requirement is that knowledge involves something beyond mere true belief and a condition is required that excludes true beliefs which are not formed by a methodical or responsible process and just so happen to be true. If Jane guesses that there are 102 books on the shelf, then although her belief happens to be true, we can say that it was lucky and is not strong enough to constitute knowledge. If, on the other hand, she formed her belief after counting the number of books on the shelf then her true belief is justified and is a case of knowledge.

The second half of the twentieth century saw a conclusive challenge to the JTB analysis which prompted a revision in epistemological theorising and spawned a wave of new work on defining propositional knowledge. In 1963 Edmund Gettier published a short landmark paper titled ‘Is Justified True Belief Knowledge?’ [82], in which he refuted the JTB account of knowledge by providing a couple of simple examples showing that there are clearly cases of justified true belief that are not knowledge. Here is one of those examples. It is supposed that a person Smith has strong evidence for the following proposition:

(A) Jones is the man who will get the job, and Jones has ten coins in his pocket.

Say, as Gettier does, that this evidence consists of the president of the company telling Smith that Jones would in the end get the job and that Smith had counted the coins in Jones’s pocket ten minutes ago.

¹see *Meno* and *Theaetetus*

Now, (A) entails the following

(B) The man who will get the job has ten coins in his pocket.

Based on this evidence and reasoning Smith comes to believe (B). As it turns out, (B) is true; the man who will get the job does have ten coins in his pocket. So Smith has the justified true belief that (B).

But now imagine the following is the case. Smith, who is unaware that he also has ten coins in his pocket, actually gets the job (the boss made a mistake in working out who was getting the job). Although Smith's true belief is justified, it does not appear to be a genuine case of knowledge. Smith's true belief has been 'Gettierised' or 'Gettierisation', which can be broadly defined as the phenomenon whereby a justified true belief is not knowledge because it is lucky or the truth of the belief is not connected with the justification/evidence in the right way, has occurred.

Interestingly, another Gettier-style example of justified, true belief that is not knowledge can be found pre-Gettier in the writings of Bertrand Russell. His example, which will be used later on, involves someone who consults a clock at time x . The clock displays time x and as a result the person comes to form the justified, true belief that the time is x (since looking at a clock is a justified way of telling the time and the time that they read is in fact the correct time). The twist is that unbeknownst to this person, the clock broke down at time x on the previous day. Given this, it seems right to say that the person does not know that the time is x ².

Whilst the extent to which the subjects in these examples are justified can be debated, it is fair to say that these examples suffice to show that truth, belief and justification are not sufficient conditions for knowledge. In Gettier's example the justified true belief comes about as the result of a deduction that uses a justified false belief, namely the false belief that 'Jones will get the job'. This led some early responses to Gettier to conclude that the definition of knowledge could be easily adjusted, so that knowledge was justified true belief that depends on no false premises. This 'no false premises' solution did not settle the matter however, as more general Gettier-style problems were then constructed in which the justified true belief does not result from reasoning that uses a justified false belief.

Much work has been done in epistemology since Gettier's paper in order to try to find an adequate definition of propositional knowledge that can, amongst other things, deal with

²Russell's passage can be found in his book *Human Knowledge: Its Scope and Limits*, in Section D of the chapter 'Fact, Belief, Truth and Knowledge' (Chapter 11). It is quite surprising that Russell's example was largely overlooked by the epistemology community.

Gettierisation. The common goal is to search for what needs to be added to true belief in order to get knowledge. There is a variety of proposals concerning just what this addition(s) might be and there can be more than one way to flesh out a proposal ³. Furthermore, there are classificatory dividing lines according to which accounts can be categorised depending on what position they take with regard to some conceptual debate.

According to one prominent account from Lehrer and Paxson [120] a defeasibility condition is added to the JTB analysis, so that knowledge is defined as *undefeated* justified true belief. A justified true belief is undefeated when there is no extra truth that the subject could come to acquire that would defeat their present justification for the belief. This explains why in Gettier's example Smith does not know that the man who will get the job has ten coins in his pocket, for if Smith had known the extra truth that Jones is not getting the job, this would have defeated the justification for his belief and he consequently would not have held it.

Regarding some classificatory features of this account, it is an approach that seeks to supplement and use as a foundation the JTB account rather than revise the foundation. It also "employs the original normative intuition that it is the quality of the reasons which distinguishes knowledge from mere true belief" [116]. In this way it is located within the internalist camp of the internalist/externalist divide [170, 147].

Roughly speaking, the internalist position maintains that the factors that distinguish mere true belief from knowledge are factors internal to the epistemic agent. These factors are within the set of the agent's consciously held beliefs and the agent has introspective access to them. In this sense justification as it features in JTB-grounded accounts is an internalist element. Conversely, the externalist position maintains that the extra condition(s) involves some factors in the world external to the epistemic agent. For a true belief to be knowledge it must stand in some appropriate relation to the true state of affairs believed; this is an external factor which the agent need not be directly aware of. The reliabilist account propounded by Goldman [84] is one prominent form of externalism. According to this reliabilism, for a true belief to count as knowledge it must be the result of some reliable process of belief-formation. The reliability of this process is a factor external to the agent. Whether or not a belief is generated by a reliable process is something separate from the agent's internal state and something that they may very well be unaware of. Dretske, as we shall soon see, and Nozick [137] have also provided externalist accounts of knowledge with "reliabilist contours" [85]. Nozick employs counterfactuals in his 'tracking theory' of knowledge in order to implement its requirement that to know that p the method employed in generating the true belief that p must reliably track the truth and be sensitive to varying

³See [19] and [171] for some literature on the matter.

relevant conditions.

Amongst the search for an adequate definition of propositional knowledge some philosophers have adopted information-based approaches. Dretske and Floridi are two philosophers who have thought that information could be the missing key and have used information to construct accounts of knowledge that, amongst other things, overcome the Gettier problem. Dretske's influential information-theoretic epistemology involves the definition of information that was covered in the previous chapter and defines knowledge as information-caused belief. Floridi [77] offers an account of knowledge that differs from standard approaches to defining knowledge, abandoning them in favour of a non-doxastic, informational and erotetic approach. In short, information is true semantic content and knowledge is information that has been correctly accounted for. Thus under both of these accounts information is a fundamental precursor to knowledge. Truth is embedded into knowledge because knowledge is based on information, which itself is also veridical. We shall now look at Dretske's account in more detail.

5.2 Dretske on Knowledge

In *Knowledge and the Flow of Information* Dretske constructed an original and influential information-theoretic account of knowledge (ITE). Whilst Dretske's informational approach to epistemology is a promising line to take, there is also room for further investigation into this approach. In this section my aims are to explicate certain aspects of this account, discuss some extant issues and suggest some modifications and additions which could be made in order to improve the case for such an informational epistemology.

Let us begin by cutting straight to Dretske's key definition of knowledge resulting from his analysis:

(ITK): K knows that s is F = K 's belief that s is F is caused (or causally sustained) by the information that s is F .

Clearly information plays a central role in this definition. In the last chapter a few suitable candidate accounts of information were covered. Whilst it is Dretske's account that will form the background for most of the following discussion, it is important to note that the different accounts of information can be 'plugged' into this definition. Consequently, the account of information that is used will determine the account of knowledge that results; the knowledge that p will in part depend on whether the account determines that the information that p is

present. Whilst all such accounts will share common features, such as the requirement of some sort of external regularity, there are going to be significant differences and particularities. To take a few examples, firstly, as we have seen, knowledge closure will fail for Dretske's account but not for the contextualist account. Secondly, the relativist approach will mean that each knowledge statement can be individually qualified and made relative to a set of relevant alternatives. With the exception of some key points particular to Dretske's account, the following discussion applies to such other instantiations of this definition.

The account given is restricted to perceptual knowledge of contingent states of affairs (so states of affairs having an informational measure of something greater than zero given Dretske's definition of information) and deals with only *de re* knowledge, not *de dicto* knowledge (if one sees the flag of Switzerland without knowing that it is the flag of Switzerland, then they can come to know *de re* that the flag of Switzerland has a cross without knowing *de dicto* that the flag of Switzerland has a cross).

With this information-theoretic account of knowledge Dretske eschews (or at least demotes) the 'philosopher's usual bag of tricks' such as justification, reasons and evidence in order to provide a more viable account of perceptual knowledge. One shouldn't need reasons, evidence or rational justification for their belief that the temperature in the room is 26°, as long as the thermometer they are consulting is functioning properly and providing the right information. Since Gettier cases have exposed the shortcomings of the justified, true belief account of knowledge an epistemology which is not based on this tripartite foundation is sought. As will be covered in more detail shortly, this information-theoretic account of knowledge is immune to Gettier-like difficulties because though one can get into an appropriate justificatory relationship to something false, one cannot get into an appropriate informational relationship with something false.

Also, eschewing these traditional epistemological elements allows for a more general, less anthropomorphic account of knowledge that could at least accommodate the possibility that animals could know things, without having to suppose that they are capable of sophisticated, human operations like reasoning and justifying. It also paves the way for a possible epistemological framework for artificially intelligent agents.

Dretske's account of knowledge clearly falls under the banner of externalism. As he puts it:

Externalism is the name for an epistemological view that maintains that some of the conditions required to know that *P* may be, and often are, completely beyond the ken of the knower. ... The idea is that the information required to

know can be obtained from a signal without having to know that the signal from which you obtain this information actually carries it. [59, p. 39]

So such an informational epistemology defines knowledge without necessarily informing us whether we have it. A consequence of this is that the KK principle is rejected. According to the KK principle, if one knows that p , then one knows that one knows that p [99]. With ITE the conditions required for information flow and knowledge do not themselves need to be known. When this is the case, it follows that one can acquire knowledge without knowing this. In order to know that they know that p , one would need the information that the signal carrying the information that p is carrying the information that p . As we saw in the previous chapter, signals generally do not carry this type of meta-information. Thus one can know that there is metal below them if their functioning metal detector sounds but not know that they know that there is metal below them. This is because the metal detector's sound does not rule out the possibility that it is due to a malfunction.

Such failure of the KK principle makes sense given this externalist informational account of knowledge. Beyond this reasoning that accompanies Dretske's picture, I shall now try to give a more rigorous analysis of meta-knowledge in Dretskean terms. Let K_ap stand for a knows that p . In judging $K_ap \Rightarrow K_aK_ap$, let us treat K_ap as a proposition which falls within the scope of the outer K_a operator. Given this, K_aK_ap would require the satisfaction of the following three conditions:

1. the truth of K_ap
2. a believing that K_ap
3. a must have the information that K_ap . In terms of the Dretskean analysis we are working with, this would mean that they would need to rule out all relevant alternatives to K_ap .

(1) and (2) are straightforward. The satisfaction of (1) follows from the satisfaction of the above three conditions for p . When judging the KK principle this is given. Given that agents will believe what they believe to be true, if they know that p then they will believe that they know that p .

(3) is the interesting condition. The failure of KK implies that the set of relevant alternatives against which K_aK_ap is judged is different to and greater than the set of relevant alternatives against which K_ap is judged. Let r stand for the appropriate belief-causing signal. If K_ap is true, then $r \sqsupset p$. But since KK fails, it follows from this that $r \sqsupset K_ap$ is not the case.

Going back to the metal detector example from the previous chapter, relative to a certain set of relevant alternatives indexed by x the metal detector's sounding (s) carries the information that there is metal (m): $\Box_x(s \supset m)$. Whilst relative to this set the detector's sounding is correlated with knowing that there is metal [$\Box_x(s \supset Km)$], Km is to be judged against a different set instead. Amongst this set are alternatives in which the sounding of a malfunctioning detector causes the belief that there is metal even though this is not the case: $[\Diamond_y(s \wedge \neg m \wedge \neg Km)]$.

At an extreme level, y could index the set of all alternatives; every alternative is relevant. Given the presence of skeptical brain-in-a-vat situations amongst these alternatives one could never be in a position to possess this kind of meta-knowledge. More realistically, under some contextualist framework where the restricted y set would still be a superset of the x set, it would be possible for the knower to rule out all the relevant alternatives. For example, with regards to the metal detector scenario, given that the x set requires the functionality of metal detector as a channel condition, knowing that there is metal via the detector's tone is not enough to know that you know. But given a bigger set y which includes scenarios where the metal detector is broken, knowing that there is metal by digging it up and seeing is enough to know that you know.

Another related consequence of Dretske's analysis, as covered for information in the previous chapter, is that knowledge closure also fails (Using the notation of epistemic logic we will represent this as $Kp \wedge K(p \supset q) \not\vdash Kq$).

This raises the important distinction between knowing that p and knowing that p is true. Unlike the former, the latter would at least require meta-knowledge. Dretske though seems to confuse the two, or has at least been careless at times with his terminology:

Closure tells us that when S knows that P is true and also knows that P implies Q , then not only must Q be true (*modus ponens* gets you this much), S must know it is true. [56, p. 13]

Contrary to what is suggested by the above quote, knowing that p is different to knowing that p is true. It is a hallmark of such an externalist informational epistemology that one can know that p without being in a position to know that p is true; you don't need to know that the proposition is true in order to know the proposition.

To put it another way, knowing that p is true means knowing that you know p and since in Dretske's account the KK principle fails knowing that p does not imply that p is known to be true. If one did know that p is true, then they would be in a position to know everything

that p entails. Since closure fails for knowledge, this is another way to show that knowing that p is different to knowing that p is true.

It does though suggest the possibility of some form of closure that involves meta-knowledge. Perhaps one idea that could be implemented under some contextualist framework is a weaker closure principle: $KKp \wedge K(p \supset q) \vdash Kq$. If one comes to know that there is a zebra by seeing it at a zoo (Kz) this does not imply KKz . Alternatively, if they come to know that there is a zebra by running a DNA test on it, then both Kz and KKz are the case. From KKz and $K(z \supset \neg m)$ would follow $K\neg m$. We will touch upon this idea again in Section 5.3.

Given Dretske's definitions of information and knowledge, it might seem that his analysis is viciously circular. We recall that in his definition of information k stands for the background knowledge the epistemic agent already has about the source. If knowledge is analysed in terms of information, and information is in part analysed in terms of knowledge, a seeming problem is that the reference to k prevents the definition of knowledge from getting off the ground; one would already have to have a definition for knowledge in order to apply it.

Although the definition is circular, it is not viciously so. The idea is that recursive application of the definition will end in termination. To make this clearer, consider the following example, a simplification of Dretske's Shell Game example. There are three shells and a peanut is located under one of them. You investigate shell 1 and find that the peanut is not under it. Given what you already know, 2 possibilities remain. You subsequently investigate shell 2 to discover that it also does not contain the peanut. Consequently, for you this signal carries the information that the peanut is located under shell 3. The final discovery carries this information because of what you already know about the situation. Now, let us go through this example backwards. You know that the peanut is under shell 3 because you receive the information that it is not under shell 2 via visual signal and you already know that it is not under shell 1. How did you come to know that it is not under shell 1? Well, you came to know that it is not under shell 1 purely via visual signal, without any background knowledge. It is at this stage that the analysis terminates and no appeal to background knowledge is required.

With this out of the way, the next question to ask is what are beliefs and how can they be caused by information? In fact, how could information cause anything, let alone beliefs? Dretske illustrates this with a simple example. Suppose that you are waiting for your friend at home and that the two of you have pre-established that the action signalling his arrival will be three quick knocks on the door, followed by a pause, followed by another quick three knocks. It is that particular signal, that particular rhythmic pattern that constitutes the information that your friend has arrived. Things like the knock having a certain pitch or

amplitude are irrelevant. When it is this particular rhythmic pattern of knocks that causes you to believe your friend has arrived, then we can say that the information that your friend has arrived causes you to believe he has arrived. The knock might have other incidental consequences; a loud knock might frighten away a nearby mouse or cause the window to vibrate, but what causes these things is not the information, because these things would have occurred with any loud rhythmic pattern.

Finally, the causally sustained qualification is an important one. Imagine your football team played a game last night of which you do not yet know the result. You take a guess and come to form the mere true belief that they won the match. You then watch a replay of the game which gives you the information that they won and your mere true belief that the team won then becomes knowledge. Although this piece of knowledge was not initially caused by the information, it was causally sustained.

5.2.1 Testing and Supplementing Dretske's Informational Epistemology

With an outline of Dretske's informational epistemology given, it is now time to take a look at how this account of knowledge deals with certain examples. Firstly, I will show how this account successfully deals with some standard Gettier cases. Secondly, I will go through some examples that challenge this account. I will argue that ultimately they do not refute the basis of such an informational epistemology and at worst mean that Dretske's definition of knowledge needs to be refined or that such an account needs to be furnished with some supplementations.

As exemplified by Gettier's examples, the involvement of luck in cases of justified true belief formation precludes knowledge. Despite attempts to salvage the JTB account, Floridi [72] argues that an account of propositional knowledge based on justified true belief is doomed. Such an approach

can become adequate only if it can solve the Gettier problem. However, the latter can be solved only if the problem of a successful coordination of truth and justification can be achieved. But such coordination problem is unsolvable because it is equivalent to the 'coordinated attack' problem, which is demonstrably unsolvable in epistemic logic. It follows that the tripartite account is not merely inadequate as it stands, as proved by Gettier-type counterexamples, but demonstrably irreparable in principle, so that efforts to improve it can never succeed. [72, p. 61]

Floridi uses this verdict to pave the way for his own informational analysis of knowledge [77]. I am not sure exactly what to make of the results given by Floridi in [72] and how wide their reach is. Floridi considers the following lemma to be the one that correctly represents an approach that would deal with the Gettier issue by revising the tripartite account to avoid the counterexamples whilst not introducing other problems:

Lemma 1: (i) if the tripartite account can become adequate, at least in principle, then Gettier-type counterexamples are avoidable, at least in principle; but (ii) Gettier-type counterexamples are not avoidable, even in principle, therefore (iii) the tripartite account is irreparably inadequate in principle. [72, p. 62]

The point of those hopeful that the tripartite account can be salvaged is to show that (ii) is not the case. The three ways an attempt to revise the tripartite account might be made are [72, p. 62]:

1. by strengthening/modifying the only flexible feature of the account, namely the justification condition
2. by adding at least one more condition that would prevent the Gettierization of the required justified true beliefs or, alternatively, allow their de-Gettierization; or
3. by combining (a) and (b)

Whilst I am content to conclude that justification alone will not suffice and strategy (1) will not do, I fail to be convinced though that (2) or (3) are not workable. Particularly if at least one extra primary condition was to be added and justification was to be given a secondary, supplementary role, then it seems that such an approach could be effective.

In a recent thesis [114, p. 145], Christoph Kelp puts it this way:

[Floridi argues] not only that the relevant type of analysis will never solve the Gettier problem but even that the Gettier problem is logically unsolvable by analyses of this type. The relevant type of analysis is of the following form:

(TB+X) S knows that p if and only if

1. p is true
2. S believes that p

3. S's belief that p has property X (where having property X does not entail that the relevant belief that p is true)

So the Gettier problem is unsolvable by any attempt to define the knowledge that p that adds a property (or properties) X to true belief that does not entail that the belief that p is true. Of course, with Dretske's account X is the property of being information caused, which does entail the truth of the belief. Hence such an analysis would be able to successfully deal with this issue. Also, Dretske's account at the very least demotes justification and so it is without the boundaries of Floridi's purview.

Thus Dretske's account can easily deal with the above example from Gettier. Smith has not received the information that the man who will get the job has ten coins in his pocket. Clearly the boss' word is not a source of information and one of the premises used in the deduction is false. There are two ways of looking at this. The first is to say that if some piece of non-information is used in deducing some truth then that truth is not information. The second way is to say that the signal consisting of the boss' word *and* counting Jones' ten coins does not carry the information that the person who will get the job has ten coins in his pocket.

As can be gathered from the last chapter, this account also easily explains why knowledge is not present when one forms a belief based on looking at a broken clock that happens to be stuck on the right time; the clock's signal does not carry the information that the time is such-and-such.

Other examples of epistemic luck further demonstrate the workings of this information-theoretic epistemology. Consider someone who, upon looking into a sheep's field sees something that really looks like a sheep and forms the true belief that there is a sheep in the field. Unfortunately for this person however, what they are looking at is in fact not a sheep but a sheep dog that looks like a sheep. Nevertheless, their belief is true since there is a sheep in the field, hidden from view behind the dog. This example is different to the ones provided by Gettier, but nonetheless involves an element of luck that precludes knowledge.

In terms of ITE, we can say that the sheep-like visual signal does not carry the information that there is a sheep in the field. There are relevant alternatives in which the dog's presence results in a sheep-like visual signal even though there is no sheep present. Since this signal causes the belief in question, the belief that there is a sheep is not caused by the information that there is a sheep.

Next is the barn-facades case.⁴ Henry often drives through an area in which there are many fake barn facades (he is not aware they are fakes) and only one real barn. Naturally when driving through this area he will on numerous occasions form a false belief that there is a barn because he is looking at a barn facade. Since Henry has no reason to suspect that he is the victim of such a setup, his beliefs are justified. Now suppose further that on one of those occasions when he believes there is a barn in front of him he happens to be looking at the one and only real barn in the area. This time his belief is justified and true. Since he could very well have been duped like he often is, Henry can be considered lucky that he was in front of the one real barn at the time. Thus the consensus is that his belief is not an instance of knowledge.⁵

For this case ITE would say that Henry's true belief that there is a barn in front of him is not based on the information that there is a barn in front of him, since in this context the belief-causing visual signal of a barn does not carry the information that there is a barn. Given the context, there are relevant alternatives where the visual signal of a barn results from a barn facade being in front of Henry, so it is possible amongst the relevant alternatives that 'barn-like signal' and 'not-barn'.

These last two examples in particular demonstrate a benefit of such an externalist epistemology. If the conditions external to the knowing subject are not right, the absence of the right conditions can easily be used to account for the absence of knowledge.

There is a range of purported counterexamples and objections to Dretske's account of knowledge. The usual strategy is to provide an example where it is correct to say that the subject does not know p and claim that Dretske's account is committed to saying that the subject does know p . Some objections are better than others, as the following sample shows. The objections generally fail because the cases they give would not actually be classed as knowledge given a genuine and correct application of Dretske's account. Some of the examples though compel us to rethink Dretske's definition.

Doyle [48] discusses a few examples in the literature used to critique Dretske's account. One example due to Lehrer and Cohen is taken from a collection of commentaries on *Knowledge and the Flow of Information* [52]. Sopor visits a 'marvellous planetarium', so marvellous that their night sky simulation is indistinguishable from the real night sky. There are portions of the planetarium's ceiling where the simulation is replaced with windows through which the actual sky is visible. Sopor, true to his name, falls asleep in the planetarium and after a while wakes up without recollection of his location. Looking up, he sees a real star

⁴See [171] [<http://plato.stanford.edu/entries/epistemology/#GET>]. The barn-facades case is originally due to Goldman. We saw Kripke's modified version of it in Section 4.4.3 of Chapter 4.

⁵Note that some epistemologists do not regard the fake barns case as being a genuine Gettier case. There is a touch of vagueness in the concept of a Gettier case [102].

through one of the windowed portions of the ceiling and forms the true belief that he is looking at a star. Intuition says that this is not a case of knowledge. What is the verdict under Dretske's account?

This case is similar to the fake barn scenario.⁶ In this case, the straightforward response is that the visual signal Sopor receives does not carry the information that there is a star, simply because there are relevant alternatives in which he is receiving a visual signal 'star' which is caused by a virtual star in the planetarium. So contrary to Lehrer and Cohen, the verdict under Dretske's account is that Sopor does not have knowledge.

Foley [78] dedicated a sizeable paper to critiquing Dretske's account. In one of his main purported counterexamples, he modifies Dretske's example of three quick knocks at the door carrying the information and causing the belief in a spy that the courier has arrived:

Suppose that the courier has arrived and that three quick knocks on the door cause the spy to believe that the courier has arrived, and suppose also the conditional probability that the courier has arrived, given three quick knocks and given what else the spy knows, is 1, whereas the probability that the courier has arrived given what else the spy knows (minus the knocks) is not 1. According to Dretske's account then, the spy knows that the courier has arrived. But, suppose that the way in which the three knocks cause the spy to believe that the courier has arrived is 'wayward'. Suppose, for instance, that the spy, unbeknownst to him, just prior to the knocks has become deaf and that the three quick knocks cause him to believe that the courier has arrived only because they cause the spy's partner to trip, which in turn causes a box to fall on the spy's head which in turn somehow jars his brain in a way that he suddenly comes to believe that the courier has arrived. [78, p. 169]

The verdict that the spy does not have knowledge in the second scenario is correct. But apart from being a far-fetched example, the claim that Dretske's account does not rule out this being a case of knowledge is incorrect. To explain why, let us start by identifying the three stages involved in this deviant causal chain:

- Three knocks cause partner to trip
- The partner tripping causes a box to fall on the spy's head
- The box's falling on the spy's head causes the spy to believe that the courier has arrived

⁶Given the general take on such examples by externalist, relevant alternatives accounts such as Dretske's historically, I would say that it is surprising that Lehrer and Cohen offer this purported counter-example.

Now, as Foley footnotes,

we need to further suppose here that only three quick knocks (and not three slows knocks or two quick ones) would have caused the spy's partner to trip, thus unleashing the chain of causes which result in the spy's believing the courier has arrived. For, otherwise the knocking would not have caused the spy to believe the courier has arrived in virtue of its having those properties which allow it to carry this information. [78, p. 169]

This supposition already somewhat weakens the example; it generally would be the case that any startling knock would cause the partner to trip. But let us grant the possibility that the information-carrying property could cause the partner to trip; say, if the partner also knew about the arrangement and as a result of being nervous about the impending arrival of the courier got startled because they came to know via the knocks that the courier had arrived.

This though leads to the point that for the spy's belief that there is a courier to be caused by the information that there is a courier, this information needs to be transmitted through each stage in this causal chain. In this case, the information needs to flow from the knock all the way to the spy's brain.

As Adams notes that:

... it is quite clear that an information-theoretic account needs the *proximate* cause of the belief (or sustaining cause) to carry the information that *p*. In Foley's example, that is not the case. Even if the three knock's contained the information that *p* can for some strange reason the spy's partner would not have tripped unless the knocks did contain that information, the rest of the story doesn't preserve information. The communication channel has been broken by the time the box falls on the spy's head and his brain is jarred. The conditional probability that the courier has arrived given the knocks is 1, but the conditional probability that the courier has arrived given that the events in the jarred brain has occurred, is not 1. This is because the spy's brain, since jarred sufficiently hard by the box, might be in that state (seeming to hear three knocks) even if the courier had not arrived and there had not been three knocks at the door. [3, p. 233]

To this I would add that it could even be the case that the information is lost earlier,

in the second stage; even if the partner tripped due to the information that the courier has arrived, it need not be the information that caused the box to fall. If there was a small earthquake for example, the box could have fallen off anyway and still have caused the knock that resulted in the spy's belief.

We can take Foley's example and this rejoinder to it as an opportunity to reemphasise the point made in Section 4.1 of Chapter 4 that the flow of causation is insufficient for the flow of information. Although a causal chain might be started by something that carries some information, this information can get lost along the way.

The next example, taken from a commentary given by Ginet [52], exposes an important point. We have a machine that is equipped with a red indicator light which comes on if and only if the machine becomes too hot. K walks by the machine and observes that the red light is on. Not knowing what the red light means, K asks H, having good reason to think that H would know. As it turns out, H does not know either what the red light means. This does not stop H from taking a lucky guess and telling K with confidence that the red light means that the machine is too hot. As a consequence, K comes to have the true belief that the machine is too hot.⁷

It is clear that K does not know that the machine is hot. However Ginet's claim that this is a Gettier counterexample to Dretske's account might demonstrate a limited or ungenerous application of it. As Doyle explains in countering Ginet's case, "K's belief, while partially caused by a signal which carries the information that the machine is too hot, is also partially caused by H, and his conjecture fails to carry the information that the machine is too hot" [48, p. 37]. So going by Doyle's argument, contrary to Ginet's claim, Dretske's account can determine that this is not a case of knowledge. Yes, the red light signal plays a part in causing K's belief. But it is not *the* cause of the belief; alone the red light was insufficient to cause the belief.

The general point to be made here is that K's belief is not knowledge because H's non-information played a part in the process that resulted in K's belief. But I am not sure how far Doyle's strategy can go. If it took both the red light and H's utterance to cause K's belief, then we could say that the signal consisting of both the red light and H's utterance caused K's belief. But the conditional probability of the machine being too hot given this signal is 1⁸, so technically it still carries the information that the machine is too hot and thus according to this definition the belief is still caused by a signal carrying the information.

⁷Another similar example due to Maloney can be found in [129].

⁸Let r stand for red light, u stand for H's utterance and h stand for the machine is hot. $\Pr(r \wedge u) > 0$. So if $\Pr(h|r) = 1$ then $\Pr(h|r \wedge u) = 1$.

The general problem as I see it is a certain ‘sloppiness’⁹ in the formulation of this externalist ITE which neglects certain aspects of the way information-carrying signals are used and the way in which background beliefs are used. If a receiver has to rely on background beliefs in order to extract some piece of information carried by a signal then it is imperative that these beliefs are not only true, but also qualify as information_S/knowledge. In this example, the problem is that because K used H’s non-information to extract the information from the signal, we can say something like the belief was not caused by the information in the right way. To expand upon all of this, we could even replace H’s utterance with K’s lucky guess that red light means too hot. In this example, it is not the case that K receives a signal that fails to carry the information that the machine is too hot. Rather, his belief that red light means too hot is not information-based and therefore there is something wrong with the way he uses the information-carrying signal in forming his belief that the machine is hot.

All this brings up Barwise’s concern mentioned towards the start of Section 4.1 in the previous chapter. It is important that the receiver knows about the nomic relations involved. In this case, K does *not* know about the nomic relation between red light and too hot. To requote Barwise in part, “while information *is* out there, it *informs* only those attuned to the relations that allows its flow” [52].

If K has no belief about what the red light means, then he is obviously not attuned to the relations and cannot extract the information carried by the signal. If he is misinformed that the red light means that the machine is too cold then he also cannot extract the information carried by the signal. In the example that is given K has a true belief about the nomic relation ($p \supset q$). He also is informed and knows that the red light is on (p). He uses these two to infer that the machine is too hot (q). But since he is not informed and does not have knowledge of this nomic relation, he is not informed (i.e. does not hold the information_S) that the machine is too hot.

Continuing on, another example against Dretske’s account that Doyle considers more challenging is due to Harman [52]. In this modification of Dretske’s eight-employee selection example, again one employee must be selected amongst eight. The employees have opted to inform their boss of their selection not via a message with the employee’s name, but via the colour of the envelope they send. Each employee was assigned a colour before the selection process, with Herman being assigned pink. Herman is selected and a pink envelope is thus passed on to the boss. Upon receiving the envelope the boss forms the belief that Herman was selected. He does so though not based on the convention adopted by the employees; he does not know about the convention. Rather, it is because the pink colour of the envelope

⁹This term is taken from [22, p. 134]

reminds the boss of Herman's hideous pink ties and so he associates the pink envelope with Herman.

The envelope from the employees carries the information that Herman was selected in virtue of its pinkness. Since this property causes the belief that Herman was selected, it seems that ITE is committed to attributing knowledge to the boss. Unlike the red light example above, it is not clearly the case that the belief is the result of some piece of information (i.e. the red light) being coupled with a piece of non-information (i.e. H's word); the pink envelope carries the information that Herman was selected and the boss's association of pink with Herman is reliable.

Doyle offers a way to explain the absence of knowledge here in terms of ITE:

in order to be assured that the information and not merely the envelope's pinkness causes the belief we must first be assured of the following: in a situation where the pink envelope did not carry the information it would not have caused the belief in question. In other words if the belief that Herman has been selected would have been brought about in the employer by a pink envelope, where this did not carry the requisite information, then we can say that its pinkness alone - and not the information it embodies - is the cause of the belief. [48, p. 45]

So one option is to say that ultimately it is the pinkness of the envelope, not the information that Herman was selected, which causes the belief that Herman was selected. In this case the two coincided, but even if they did not, say because pink was assigned to another employee, the belief that Herman was selected would still have resulted.

Thus there is something more to information belief causation than is suggested by some of the explanatory discussion of ITE. Dretske explains that when "a signal carries the information that s is F in virtue of having property F' , when it is the signal's *being* F' that carries the information, then (and only then) will we say that the information that s is F causes whatever the signal's being F' causes". [51, p. 87] Recall, not just any knock tells the spy that the courier has arrived; the signal is three quick knocks followed by a pause followed by another three quick knocks. Other properties of the knock might have a causal effect, but such causation would not involve the information.

In this sense there is little to distinguish the information-bearing knock on the door causing the true belief that so-and-so is at the door and the information-bearing envelope's pinkness causing the true belief that Herman was selected. The knock carries the information in virtue of its pattern and it is this pattern that causes the respective belief. The envelope

carries the information that Herman was selected in virtue of its pinkness and it is this pinkness that causes the respective belief. Thus going by this picture it would seem that the information that Herman was selected causes the belief that Herman was selected.

With the establishment of this point a crucial distinction becomes clearer. The distinction is that between (1) a signal that carries the information that p and its information-carrying property causing the belief that p and (2) the information that p causing the belief that p . The latter involves the former but also requires something extra. Not only must the signal cause the belief that p as a result of that property it possess with which it carries the information that p . The receiving agent must also be informed of this connection.

These considerations raise the prospect of tightening the original formulation of knowledge. Starting with the basis of knowledge as information-caused belief, some sort of qualification which required that the belief was caused by the information in the right way would do the job. Imposing such conditions on the nature of the causation in the case of belief seems reasonable; after all, information causing a belief seems to be a more sophisticated phenomenon than, say, a white ball hitting an eight ball and causing it to move.

Given all of this, let us start with the following revised alternative definition of knowledge:

(ITK₂): K knows that s is F = K 's belief that s is F is based on K being informed that s is F

What is involved here in an agent being informed that s is F ? Firstly, a signal r carrying the information that s is F must be received by the agent. Secondly, that piece of information must be extracted by the agent in the right way, so that the agent becomes informed that s is F . In order to extract the information that s is F from the signal, the agent must be attuned to the right informational relation; they must have the information/know that the signal indicates that s is F . Whilst this notion of being attuned to the right informational relation is something that I am sure the reader can appreciate, this idea is one to flesh out. To what extent, if any, this requirement detracts from the externalist/naturalistic purity of Dretske's original formulation is something to be considered. Firstly, care must be taken with the terminology employed. Being attuned to the informational relation does not mean being informed that r carries the information that s is F ; as we have seen, such meta-information is not generally available according to this externalist account. In terms available given this informational account, we can define this idea of being attuned as follows. An agent being attuned to the informational relation between r and s being F means that the agent has received a reliable signal that carries the information and has made them aware that there is a reliable correlation between r and s being F (this correlation will hold over a given set

of relevant alternatives).

One way the boss would have become attuned is by the employees informing him of their convention. Associating Herman with pink based on his hideous pink ties is not a way to be attuned to the fact that the pink signal indicates that Herman was selected, because there are relevant alternatives where this association is made but pink does not carry the information that Herman was selected. The same applies to Ginet's red light example. H's utterance does not carry the information that the red light indicates that the machine is too hot because given the unreliability of H's utterance there are relevant alternatives in which H utters this even though the red light does not stand for too hot.

There is another interesting way that one could go about tightening the original formulation of knowledge to deal with such problematic cases. The idea is prompted by the pink envelope example above and the following quote from Adams:

As we've seen, ITKs admit that a belief that p can be knowledge, even if the belief that p does not contain the information that p , as long as it was caused by the information that p . [3, p. 233]

Eliminating gaps that allow for beliefs that are caused but do not carry the relevant information might do the trick. So let us look more into the idea of adding a requirement whereby for a belief that p to count as knowledge, it must carry as well as be caused by the information that p . We have the following definition:

(ITK₃): K knows that s is F = K 's belief that s is F is caused (or causally sustained) by and carries the information that s is F .

A first thing to note is that this approach offers an alternative account to Doyle's of the problematic pink envelope example. Rather than saying that it was not the information that Herman was selected that caused the boss's belief that Herman was selected, with ITK₃ another response available is to say that the boss' belief did not carry the information that Herman was selected. We have established that the boss could very well have believed that Herman was selected even if Herman was not. This is because the boss' reason (Herman's hideous pink ties) is not based on the colour assignment convention adopted by the employees. So although the boss believes that Herman was selected, it is not necessarily the case that had Herman not been selected, the boss would not have believed that Herman was selected (the convention could have easily been different and if so, Herman still would have believed). So the belief does not carry the information that Herman was selected. Adopting

the formal notation from the previous chapter, let us use $Bp \sqsupset p$ to stand for ‘the boss’ belief that Herman was selected carries the information that Herman was selected’. In terms of our modal logical system, this can be converted to $\Box_x(Bp \supset p)$. Now, in the set of relevant alternatives ranged over by $\Box_x$, it is fair to include ones in which the boss holds his hideous pink tie association but the employee assignment convention has changed. In such alternatives it is the case that $Bp \wedge \neg p$, so $\Diamond_x(Bp \wedge \neg p)$ and $Bp \sqsupset p$ fails to hold.¹⁰

We can frame all of this in terms of certain relations. As was covered in the previous chapter causation is insufficient for information flow. For example, if either of B or C causes A then neither $A \sqsupset B$ nor $A \sqsupset C$ holds. In the pink envelope case, although a signal carrying the information that Herman was selected causes the belief, since signals that did not carry that information would also have caused it then information does not flow into the belief. In cases where B is the only thing that causes A , then $A \sqsupset B$ does hold. In such cases, there is a one-one relationship between A and B . So not only would it be necessary for the belief that p to be caused by a signal that carries the information that p , but given a certain set of alternatives such signals must be the only cause of the belief that p . This clearly fails for Herman’s boss; one such relevant alternative for example would have a colour code system whereby say yellow stood for Herman and pink stood for another employee.

Having both causation by the information that p and carriage of the information that p in this modified definition of informational knowledge consolidates a one-one link between the information and the belief. It is this one-one relationship that would need to obtain; within a relevant set of alternatives the belief that p if and only if the information that p .

Recall Doyle’s claim that “in order to be assured that the information ... causes the belief we must first be assured of the following: in a situation where the pink envelope did not carry the information it would not have caused the belief in question” [48, p. 45]. A possible logical problem with this claim is its suggestion that for the information to cause the belief it must be the case that the absence of the information means the absence of the belief. But this contrapositive relationship is not a necessary condition of causal relations. Rather than concentrating on whether it is the information that causes the belief, ITK₃ can deal with the pink envelope case by appealing to the qualification that not only must the information that p cause the belief that p , it must be the only possible cause of p within some set of suitable alternatives. The boss’ belief fails this test.

This qualification that the belief must also carry the information can also deal with other cases that are similar but also importantly different to the pink envelope case. In determining that $Bp \sqsupset p$ fails for that case, the main factor was that there are relevant

¹⁰In a way this example is similar to Foley’s modified spy scenario, in that the information has not flowed all the way through the causal chain from start (the employee selection process) to finish (the boss’ belief).

alternatives in which the convention changes yet the boss still would believe that Herman was selected given a pink envelope: “a pink envelope carrying the information that, say, Shirley was selected, combined with the stipulation that no other colour but pink could have begotten the employer’s conviction that Herman was the nominee, could nevertheless have caused the belief that Herman was selected” [48, p. 46].

Alternatively, even if for some reason the convention was the same over a set of relevant alternatives, the epistemic state of the agent in relation to the convention could preclude knowledge another way, as demonstrated in the following example

Eight employees have to choose who is to perform a task. They will write down the name of who is selected on a piece of paper and pass it on to their boss in order to inform him of the selection. Two of these employees are named William; William Jones and William Smith. If William Jones is selected, they will write ‘Will’ on the piece of paper, as that is what they generally refer to him as. If William Smith is selected, they will write ‘Bill’ on the piece of paper, as that is what they generally refer to him as. This system results in a perfect and reliable information carrying source. In MTC terms, each selection generates 3 bits of information and for any message; the probability that the employer was selected given their name being on the piece of paper is 1.

But now suppose that the boss is not aware of this naming system, and thinks that a message with ‘Will’ or a message with ‘Bill’ can refer to either of the two men. For the boss, $\Pr(\text{Jones selected} | \text{‘Will’}) = \frac{1}{2}$ and $\Pr(\text{Smith selected} | \text{‘Will’}) = \Pr(\text{Jones selected} | \text{‘Smith’}) = \frac{1}{2}$ and $\Pr(\text{Smith selected} | \text{‘Smith’}) = \frac{1}{2}$. After going through their selection procedure, the employees select William Jones for the task. The boss receives ‘Will’, guesses that it refers to William Jones and forms the true belief that William Jones was selected.

Now, this does not feel like a case of knowledge. Going by the first modified definition ITK_2 , we can say that the boss is not properly informed that ‘Will’ corresponds to William Jones and so is not informed that William Jones was selected. Although the information caused the belief, it did not cause the belief in the right way. Going by the second modified definition ITK_3 , the boss’ belief that William Jones was selected does not carry the information that William Jones was selected. This is the verdict that would be reached given that there are relevant alternatives in which the boss receives the message ‘Bill’ (Smith was selected) and associates it with William Jones, since he has no guaranteed way to correctly associate each of the messages with their respective outcomes.

In a variation of this example, the employees and the boss start off by agreeing on a message system convention. The only William amongst the employees is William Jones, so it is established that either of ‘Will’ or ‘Bill’ will refer to him. Just before the employees

are to make a selection, Sally is replaced by William Smith. As a result of there now being two Williams amongst the group, Jones is once again assigned 'Will' and Smith is assigned 'Bill'. The boss however is not notified of this new setup. William Jones is selected and the message 'Will' is received, causing the boss to form the true belief that William Jones was selected. Certain intuitions suggest that the boss' belief here is not knowledge, since it relies on an element of characteristically knowledge-robbing luck; if Smith was selected then the boss would still have believed (falsely) that Jones was selected.

ITK₃ offers a response to this example not available to Doyle's approach to the pink envelope example or ITK₂. With regards to Doyle's strategy, it is clear that the signal 'Will' carries the information that William Jones was selected and had this not been the case the boss would not have come to this belief. So it is fair to say that the information did cause the belief. With regards to ITK₂, the boss receives the signal 'Will' that carries the information that William Jones was selected and is also informed of the connection between William Jones and 'Will'. Therefore the boss is also informed that Jones was selected. Perhaps one option is to expand ITK₂ to ensure that the boss is properly informed of all the signal/source correlations in the system (in this case there would be failure for the Smith/'Bill' correlation). At any rate, with regards to ITK₃ it is clear that the boss' belief does not carry the information that Jones was selected, since if William Smith was selected, surely a relevant alternative, the boss would receive 'Bill' and still believe that William Jones was selected.

As we will see in Section 5.2.4, points of view can differ on whether or not this last example is a case of knowledge. More generally, it is an open matter to explore whether ITK₃ might be too strong a definition because it excludes certain cases of true belief that it would be fair to count as knowledge.

The examples discussed thus far have stretched and tested Dretske's original definition of knowledge given at the start of this section. As result I have suggested two alternative definitions of knowledge in keeping with the original. To what extent these definitions align is an open question. If neither subsumes the other then perhaps they could be combined. Before ending this section, I will pose a few more potential issues.

Some epistemologists, particularly those with internalist inclinations, have concocted examples which target the nature of externalism. In Bonjour's clairvoyant example, Sally is a clairvoyant who has a special faculty or power that tracks the truth about *p*.

Sally clairvoyantly believes that the president is in New York. The president is in New York. Sally is aware of a massive media blitz orchestrated by the

secret service saying that the president is at that time in D.C. The cover up is designed to thwart an assassination attempt on the life of the president. Sally has no reason to think that she has clairvoyant powers. Indeed, she has good reason to think that there are no clairvoyants. Sally also has no good reason to believe that the media blitz is a cover up or hoax. Bonjour claims that Sally would be epistemically irresponsible to continue to believe that the president is in New York in the face of the overwhelming evidence to the contrary. Let us suppose that, because of her clairvoyance, Sally is unable to share her belief. Bonjour claims that despite the fact that Sally is tracking the truth about the whereabouts of the president, she does not know that he is in New York because she is being epistemically irresponsible in her belief. [4, p. 28]

This example provokes a conflict between people's intuitions. Whilst Bonjour claims that this is not a case of knowledge, Adams is happy to disagree and maintain that Sally knows the president is in New York:

Her belief is true. It is not an accident that her belief is true. There is no close world in which were she to apply her clairvoyance, it would lead her astray. Her belief is produced by a nomically reliable primary cognitive process. What more could one want from a cognitive process but that it be nomically guaranteed to produce truth? [4, p. 28]

Whilst Adams' position is reasonable, there is a case to be made using an information-theoretic account for the absence of knowledge in Sally's case. If being informed that p requires correctly processing a signal which carries the information that p , then clearly Sally is not informed and does not know as per the first modified definition I suggested.

But without resorting to such an explanation I think that the main element in Dretske's original definition could be applied here. Part of the problem with this example is that the notion of clairvoyance is not a well-defined one. If it means, as I think it does, that there is some perfect correlation between the truth and Sally's beliefs, then we can say that such a belief that p carries the information that p . The problem is that it is not also the case that the belief is *caused* by the information that p ; correlation does not imply causation and something can carry the information that p without being caused by something which carries the information that p . If this is the case, then the absence of causation could explain the absence of knowledge.

There are similar but different examples though where the belief that p both carries the information that p and is caused by the information that p . Take Lehrer's Truetemp case.

Suppose, that quite without being aware, Mr. Truetemp has a high-tech, belief-producing thermometer chip implanted in his brain. The chip measures the surrounding temperature and then directly enters a belief that it is the temperature into Truetemp's belief box, if you will. ... there is no signal that Truetemp perceives and which carries the information. The information is mainlined into Truetemp's belief box, as it were. Suppose Truetemp is cognitively unable to withhold belief despite having no evidence that his beliefs are correct.

In this case, a signal that carries the information that p causes the belief that p and the belief that p carries the information that p . I think that neither the non-knowledge nor the knowledge verdict is unequivocally the correct one here. Whilst denying that this is a case of knowledge is not unreasonable, insisting "that this is a case of knowledge, albeit a very unusual kind of case" [4, p. 33] is also plausible. Denying knowledge based on a conflicting intuition is a start but not enough, especially if an informational account that determines knowledge in this case has greater overall coherence and explanatory power.

This example does come close to the bone of pure externalism. If one did want to deny knowledge under an informational account, one could appeal to the notion of being informed as per the first modified definition. If being informed that p requires the conscious processing of a signal which carries the information that p , then since in this case such a process is bypassed, Truetemp is not informed. To be informed, the information that p must be extracted from the signal using only other information and a conscious sensation carrying the information must cause the belief.

It is worth mentioning that those such as Dretske and Adams who endorse an externalist information-theoretic epistemology think that one of its selling points is that it can account for knowledge in pre-introspective infants or animals. As Adams says,

Raven [the dog] knows that a squirrel ran up the tree. Her belief is based solely on her untutored cognitive processes. She has no secondary checks in place on whether her cognitive processes are reliable or whether her relying on them is justified. That is, she has no mechanism for confirming whether her cognitive mechanisms are reliable. Yet nearly everyone agrees that Rave can know the squirrel ran up the tree. [4, p. 28]

I think that this is a fair point and one of the benefits of an information-theoretic epistemology that can be used to help tip the scales in its favour (particularly against internalist requirements). I do also however think that whilst humans (including infants) and animals

can have the same informational basis for knowledge, there is nothing untoward in stipulating that knowledge for each group is different, that there are different types of knowledge. Human knowledge shares the same roots as canine knowledge for example but the former is more sophisticated and hence is associated with extra requirements, just like the human brain and the canine brain share the same basis though the former has extra developments. Examples suggesting that human knowledge involves extra elements can arise because of the fact that humans can interact with the world in a richer or more sophisticated way.

The minimum lesson to be learned from this is that there could very well be a problem in sticking with an account that is too ruthlessly externalist. For human knowledge, it is reasonable that some minimum level of processing an information-carrying signal is required. Going further, one obvious option would be to introduce some notion of justification or epistemic responsibility into the picture, coupling it with information. How exactly this would be required or work I leave open. Such a move is one easy way to satisfy certain intuitions that might creep in as a result of some examples.

Even given the extended and expansive applicability of the modified definitions that I have suggested, examples that further test are still conceivable. The following and final example suggests that even though the notion of relevant alternatives can be employed to show why knowledge is not present due to the use of non-information in forming a belief, there are limits to stretching this strategy. It raises the possibility that despite the presence of total information, a belief that is formed is not knowledge if the agent fails to be properly aware of how the information system they are using works.

Consider a perfectly functioning, information providing standard 12-hr analogue clock that is currently showing the time 6:30. So it has both hands located at 6 on the clock. Harry consults the clock and forms the true belief that it is 6:30. But Harry is mistaken about something; he thinks that the small hand indicates the minutes and the big hand indicates the hours. Given this incorrect interpretation of the clock's signal, it seems fair to say that Harry does not have knowledge (though I would say that this verdict is not unequivocal). It is though fair to say that the following hold:

1. the clock's signal carries the information that it is 6:30
2. this information causes the belief
3. this belief carries the information

Point (1) is straightforward. Points (2) and (3) are open to discussion but it is not clear that one could be denied in order to salvage the informational definitions of knowledge

suggested. One strategy is to go from this one instance of the information system (time is 6:30) to a universal generalisation (any time). The test then becomes to see whether these three conditions hold for all times (events in the system). Condition (3) would not be affected given this shift, for there would still be a consistent correlation between beliefs and times: suppose that another person Sally knew about Harry's misunderstanding and he told her that the time is 6:25. From this she could extract the information that the time is 5:30. There would be repercussions for (2) though, which would fail to hold. For example, given the time of 5:30 Harry would believe that the time was 6:25 although this belief would not be caused by the information that the time is 6:25. As some related discussion in Section 5.2.2 will make clearer such an approach might not be satisfactory though; this strategy of judging one successful information caused/carrying belief in terms of other instances in the information system can be problematic.

Another option would be to impose a requirement such as that introduced for the ITK_2 definition, according to which Harry must be informed of/attuned to the connection between the clock's signal and the time of 6:30. Now, there is indeed something improper in the way Harry uses the signal of both clock hands located on the 6 mark as an indication of the time 6:30. But the requirement associated with definition ITK_2 is not enough to reach this verdict. According to that definition, being informed or attuned to the fact that signal r indicates that s is F means that the agent has come to hold this fact via some reliable means. In other words, had there not actually been such a correlation the association between the signal and the source event would not have been made.

Going by this definition, it seems that Harry is informed of the informational connection between both clock hands located on the 6 mark and the time 6:30. This is because given his inverted method of telling the time, it would still always be the case that a small hand on the 6 and a big hand on the 6 would be associated with 6:30. Contrast this with the red light machine indicator example earlier, where it was not the case that K was informed that the red light indicates that the machine is overheated simply because their method of coming to believe this (H guessing and telling K) does not ensure that this connection was made; H could have still mistakenly suggested this correlation even if red light did not indicate too hot. Rather than focusing on the information connection between a specific signal and specific event in the system, it seems that a more general requirement that the agent correctly understands how the information system works would be better. In particular, they need to understand how the signals in the system work and what they stand for.

The last few examples, particularly the last, have raised the possibility that knowledge needs information plus something else. What that something else is remains open. Whilst one option is to incorporate some notion of justification, more generally we could say that

the belief-forming agent must exercise some epistemic responsibility. But without letting internalist notions dig their heels in too much, the something else could simply be that the agent must properly understand how the information system through which they access the information works. Whatever that something else is, all this suggests, to use Dretske's terminology, that perhaps a slightly impure informational epistemology is the right way to go.¹¹

5.2.2 How reliable does an information source have to be?

To what extent must an information source be regular in order for its signals to carry the information required for knowledge? This question is an interesting one, which also pertains to issues in Section 5.4. It is something that Dretske also considers in writing on the matter and a good place to start is, as he does, by framing the issue in terms of MTC [59].

Recall the employee selection example in Section 4.1 of Chapter 4, where a boss asks eight of his employees to select one from amongst themselves to perform some task. The information source that is the message generating selection process gives a reduction of 8 possibilities to 1 and it generates 3 bits of information; its entropy (H), the average amount of information associated with the source S is 3 bits, calculated as:

$$H = - \sum_{i=1}^n \Pr(i) \times \log_2 \Pr(i) = 3 \text{ bits}$$

for each of the n possible outcomes/symbols i .

As well as being the entropy of the information source, 3 bits of information is also the amount of information generated by each event; for any employee x , their selection generates 3 bits of information:

$$I(s_x) = -\log_2 \Pr(s_x) = 3 \text{ bits}$$

11

An impure theory of knowledge, for example, would contain a mixture of conditions on knowledge, some involving information (to know that p one must stand in the right informational relation to the fact that p), others, perhaps, involving subjective conditions (to know that p , one must also be subjectively justified in thinking that the appropriate informational relations exist). A pure theory, on the other hand, forgoes all subjective requirements (except belief) and identifies knowledge of the facts with beliefs that stand in the appropriate informational relations to the facts. [53]

Given an equivocation-free channel, the amount of information received equals that of the information generated; $I_s(r_x) = I(s_x)$. So if Herman is selected the amount of information transmitted and received will be the same as the amount of information generated and when the boss receives the note, it will carry the information required for knowledge.

In a modified example also given in Section 4.1 of Chapter 4, everything is the same except that the employees decide that should Barbara be selected, they will write Herman's name down on the note instead. This equivocation affects the information of the source, as the following calculations will demonstrate.

As we have seen, for each source event s_i and for some message r_a , the equivocation associated with r_a is calculated as:

$$E(r_a) = - \sum_{i=1}^n \Pr(s_i|r_a) \times \log_2 \Pr(s_i|r_a)$$

The average equivocation associated with a source is the probability-weighted sum of the equivocation associated with each message:

$$E = \sum_{i=1}^m \Pr(r_i) \times E(r_i) \quad (5.1)$$

for each of the m messages r_i .

When the channel is perfect, equivocation is 0; either $\Pr(s_i|r_j)$ takes the value of 0 or $\log_2 \Pr(s_i|r_j)$ takes the value of 0.

If the channel were completely random in this employee setup, then $\Pr(s_i|r_j) = \frac{1}{8}$ for each message and equivocation is maximum (i.e. the equivocation equals the amount generated from the source), so no information is transmitted. In this scenario, the calculation would be:

$$\begin{aligned} E &= (8 \times \frac{1}{8} \times -(8 \times \frac{1}{8} \times \log_2(\frac{1}{8}))) \\ E &= (1 \times -(8 \times \frac{1}{8} \times -3)) \\ E &= 3 \end{aligned}$$

In our current example, equivocation is somewhere between the maximum 3 and minimum 0. Since the message 'Herman' is responsible for the non-zero equivocation (all other messages have zero equivocation), we only need to calculate $E()$ for when r is 'Herman' (r_H). This will involve the outcomes 'Herman' (H) and 'Barbara' (B)

$$\begin{aligned}
E(r_H) &= -[\Pr(B|r_H)\log_2\Pr(B|r_H) + \Pr(H|r_H)\log_2\Pr(H|r_H)] \\
E(r_H) &= -(\tfrac{1}{2} \times -1) + (\tfrac{1}{2} \times -1) \\
E(r_H) &= 1
\end{aligned}$$

Since r_H will appear if either H or B, $\Pr(r_H) = \frac{2}{8}$

Plugging this into Equation 5.1, we get

$$E = \frac{1}{4} \times 1 = 0.25$$

Thus the average equivocation on the channel rises from 0 to 0.25 and the average amount of transmitted information is now 2.75.

Due to this equivocation, not as much information is transmitted on average as is generated by the selection of an employee. The equivocation associated with the message ‘Herman’ explains why the boss cannot come to know that Herman was selected based on reading the message, even if Herman was actually selected, for it could have been Barbara instead.

The same verdict is reached if we speak in terms of specific signals rather than source averages; the amount of information carried by the signal ‘Herman’, as given in the following calculation is going to be less than the amount of information generated by Herman being picked, which is 3:

- $I(s_H) = 3$ bits
- $E(r_H) = (\Pr(H|r_H) \times \log_2\Pr(H|r_H)) + (\Pr(B|r_H) \times \log_2\Pr(B|r_H)) = 1$ bit
- $I_H(r_H) = I(s_H) - E(r_H) = 3 - 1 = 2$ bits

This is what we would expect. There is not enough information given by the message ‘Herman’ to generate the knowledge that Herman was selected. In terms of Dretske’s definition of information, the probability of Herman being selected given the message ‘Herman’ is not 1.

This unproblematic result and the accord between results for average and specific signals disappear in cases where another employee is selected. Suppose that Nancy is another employee candidate. Whilst ‘Herman’ is equivocal, the message ‘Nancy’ is not and reliably indicates that Nancy was selected; the probability of Nancy being selected given the message ‘Nancy’ is 1. Calculations for the specific signal ‘Nancy’ (r_N) will easily show that there is no equivocation, since $\Pr(s_i|r_N)$ takes the value of 0 or $\log_2\Pr(s_i|r_N)$ takes the value of 0.

The key question now concerns what calculations judgements of sufficient information should be made against; averages or those involving the relevant specific signals. Is it a problem that the average information transmitted in the system is less than the information carried by the particular message ‘Nancy’ (or for that matter any other message besides ‘Herman’)? These messages reliably indicate who was selected, but since there is some equivocation associated with the system of which they are a part it is not clear-cut that their information carriage is sufficient to generate knowledge. This is where points of view can diverge.

On one side, there are those who will say that the way MTC deals with averages “disqualifies it for rendering a useful analysis of when a signal carries information in the ordinary sense of information” [59, p. 17]. This is why Dretske was motivated to develop an account dealing with the amount of information associated with particular signals. It is not the equivocation of the channel in general, but rather the equivocation associated with a particular signal(s) that is of concern. Yes, the average amount of information transmitted is less than the average amount generated. But as long as the particular signal is unaffected and reliably correlates with Nancy being selected then that is enough for it to provide the information required for knowledge.

The opposing thought is that an information source that is sometimes unreliable is not good enough as a source of knowledge-generating information, even when it does provide a signal which is right.

A channel of the sort described here, a channel that (unknown to the receiver) sometimes transmits misleading messages, is a channel that should never be trusted. If it is trusted, the resulting belief, even [if] it happens to be true, does not possess the “certainty” characteristic of knowledge. If messages are trusted, if the receiver actually believes that Nancy was selected on the basis of a message bearing the name “Nancy”, the resulting belief does not, therefore, add up to knowledge. To think otherwise is like supposing that one could come to know by taking the word of a chronic liar just because he happened, on this particular occasion, and quite unintentionally, to be speaking the truth. [59, p. 17]

In terms of Dretske’s probabilistic requirement, we can put all of this as follows. Let S_x stand for ‘ x was selected’ and let M_x stand for a message bearing the name x . Which of the following should we go by when considering the case where Nancy is selected?

1. $\Pr(S_N|M_N) = 1$

$$2. \neg(\forall x)(\Pr(S_x|M_x) = 1)$$

I would now like to offer some considerations in support of the view that it is specific signals rather than averages which should be used; so the boss can know that Nancy was selected based on the message ‘Nancy’ in the second example (although for some (obvious) reasons, including some demonstrated in Section 5.4, a channel completely free of equivocation is optimal).

I begin by offering a comment on the liar analogy Dretske used in the above quote, which I think is off the mark. If a chronic liar, happened on one particular occasion to tell the truth that p , this does not mean that their signal carried the information that p . Presumably, their utterance of the truth is random and they do not always tell that truth. For example, suppose that the liar, when asked how old they are, generally lies about their age. But on one occasion they tell the truth. Formulating this in terms of a probabilistic relation we have:

$$\Pr(\text{liar is age } x | \text{given their utterance of age } x) < 1$$

So even those who subscribe to a specific-signal scheme are not in a position to say that this is an information-carrying relation. On the other hand, if this chronic liar were to generally lie but always the truth when it came to their age, then:

$$\Pr(\text{liar is age } x | \text{given their utterance of age } x) = 1$$

It is only in this case that a scheme based on specific signals (rather than averages) determines that there is information carriage.

One very important thing to note is that it is not the presence of average equivocation *per se* that causes the possible disqualification of the source as a knowledge providing source. Rather, it is that the presence of average equivocation leads to the possibility that the receiver uses a signal that does not carry sufficient information, or even worse, misinforms. But given certain constraints on the usage of a source which shield the receiver from its bad signals things are fine.

For example, the issue of having some signals which are not right, of having a communication channel with at least some equivocation, would be rendered null if the agent using this information source was aware of which signals were not right. Suppose that a partially

functional thermometer has a scale ranging from -30 degrees Celsius to 100 degrees Celsius. Due to some issue, it provides an accurate reading from -30 to 50. Once the temperature is above 50, the thermometer stays stuck on 50. This means that only for all temperatures between -30 and 49 is the thermometer able to provide information about the temperature. If a person using the thermometer was aware of this limit then they could restrict their usage of it to the working range, within which the thermometer provides information sufficient for knowledge generation.

Extending this idea, I would say that such meta-knowledge of the working conditions of a system are not necessary for this type of shielding. Similar to the thermometer example, take a car with a speedometer that functions correctly between speed ranges 0 km/h and 100 km/h but becomes unreliable beyond that. The car is located in a country that has a maximum speed limit of 100 km/h and the car's driver (who is unaware of the potential malfunction) strictly observes road rules and regulations so will never drive above 100 km/h. Given this, it is fair to say that the driver has knowledge of their car's speed based on their reading of the speedometer.

So often it will be the case that one can have knowledge from a reliable signal even though the signal emanates from a system with overall average equivocation. What is essentially at issue is whether or not the receiver can receive bad signals from the system given the possible circumstances. If the driver is never going to go over 100 km/h, then the set of circumstances is limited to those in which the driver does not exceed 100 km/h. Given this, the faulty range of the machine will never be tested. For want of a better term, let us call such circumstances *applicable alternatives*.

One problem with not accepting a specific-signal scheme and insisting instead that the average information of received signals is no less than the average information generated is that it becomes too easy to rob signals of their knowledge generating information carriage. Suppose that the driver lives in a state where the maximum speed limit is 100 km/h. Each day they drive around, never exceeding the speed limit. The twist is that they spend one day each year driving in another state, where the maximum speed limit is 110 km/h and so when in this state they exceed 100 km/h. What are the consequences of this? Does the speedometer still provide the information required for knowledge in the original state or does this extremely occasional expansion of applicable alternatives generally nullify its status as an instrument capable of generating information about the car's speed? If the former holds then what about when the car is in the other state? Saying that the speedometer ceases to generate the right information between 0 km/h and 100 km/h only when in the other state seems awkward and unsatisfactory.

So the average-signal scheme is caught between such problematic alternatives and comes with the attendant headache of demarcating applicable alternatives. The signal-specific scheme on the other hand does not suffer from such issues and is less complicated. Whilst the average-signal scheme is affected by issues of indeterminacy, the signal-specific scheme can be determined precisely. Thus without having argued directly for the signal-specific scheme, by showing the costs of the average-signal scheme we have one reason to choose the former over the latter.

Apart from the concerns with applicable alternatives, there are other slippery-slope limit setting issues. Admittedly, there is at some level some intuitive resonance with the idea that an unreliable source, which cannot generally be trusted because some signals do not reliably indicate a corresponding event, will never give a signal that is enough to generate knowledge. If a system had 1000 signals and only 1 of them was reliable, then given this extremely high level of unreliability, it seems acceptable to say that many bad apples do spoil the bunch, and the one signal which is reliable would be rendered ineffective as a knowledge-generating information source.

On the other hand, if there were 1000 signals and 999 of them were good then surely the knowledge-generating information potential of the 999 signals should not be nullified. One bad apple does not spoil the bunch; surely a trivia book with 1 incorrect answer due to a misprint and 999 correct ones is still a source of information-generating knowledge. But since trying to determine a boundary would lead us down a slippery slope, it is preferable to accept the determinate specific-signal scheme. So a signal which reliably indicates some event carries the requisite information to generate knowledge about that event, irrespective of other signals and the general equivocation of the system.

One final issue within this theme of indeterminacy concerns where the boundaries would be set for the system being assessed.

Take the following computer function:

```
function increment(i) {
  return i + 1
}
```

With this function, for all numbers x , the output x carries the information that the input was $x - 1$. Next take the following function:

```
function increment2(i) {
```

```

    if(i == 5) {
        return 5
    }
    else {
        return i + 1
    }
}

```

In this case the output ‘5’ does not carry the information that the input was ‘4’ (also, ‘6’ will never be output). So there is some average equivocation with the system consisting of this function. Whilst my contention is that this should not affect the knowledge-generating information carriage of all the outputs other than ‘5’, let us entertain the average-signal scheme. Applied to the system consisting of this function, it is clear that the system has average equivocation.

But next consider the following function. It randomly assigns a Boolean value to one variable and then uses this to determine which of two sub-routines is to be used to operate on the input:

```

function increment_decrement(i) {

    random = rand()

    if(random == true) {
        return "increment: i + 1"
    }
    else {
        if(i == 5) {
            return "decrement: 5"
        }
        else {
            return "decrement: i - 1"
        }
    }
}

```

The system consisting of this function has some average equivocation; the output “decrement: 5” does not carry the information that the input was 6. But I think it is definitely fair to say that this should not affect an output such as “increment: 10” from carrying the information that the input was 9. This will not be the case if an unreliable part is allowed

to ‘infect’ the whole system. Once again, unlike the unaffected specific-signal scheme, the average-signal scheme faces issues of indeterminacy. By using averages instead of specific correlations, the good parts in a system are unnecessarily affected. One option is to demarcate sub-systems and apply average calculations to them, but the problem then becomes how to determine such boundaries; it would not be too difficult to come up with examples more complicated than that above which would make this strategy more troubled. These unnecessary issues can be circumvented by sticking with the specific-signal scheme.

In closing, let us introduce some terminology to capture all of this:

- S stands for the set of signals in a system
- E stands for the set of events in a system
- Drawing from meta-logical terminology, define the properties of *soundness* and *completeness* for an information system as follows:
 - Soundness - for every signal s there is some event e such that s carries the information that e :

$$(\forall x)(x \in S \supset (\exists y)(y \in E \wedge (x \sqsubset y)))$$

- Completeness - for every event e there is some signal s such that s carries the information that e :

$$(\forall y)(y \in E \supset (\exists x)(x \in S \wedge (x \sqsubset y)))$$

According to the specific-signal scheme, a signal s can still carry the knowledge-sufficient information that some event e even if it is part of an unsound system.

To briefly illustrate these definitions, take a system consisting of the time and a clock. Each time is an event and each clock position is a signal. If the clock is functioning perfectly then the system is sound and complete. If it breaks down and stops at 6:00pm, it is no longer sound or complete. The signal ‘6:00pm’, its only signal, does not carry the information that x for any x . Furthermore, for no event e (i.e. time) is it the case that a signal s carries the information that e . So this system is totally incomplete and provides no information. The partially functioning thermometer on the other hand provides some information.

A system can be unsound and complete. For example, take a light warning system which uses two lights to indicate when some event (e_1) happens. The first light (s_1) is the main light and the second light (s_2) is a backup. In this case, the first light is working properly but the second is not and often flashes even when the event does not happen. So the system

is complete, because s_1 carries the information that e_1 (the only event), but is unsound, because there is no event e such that s_2 carries the information that e .

If a system is incomplete then it follows that it is unsound. For if it is incomplete, it follows that there is some event e such that when it occurs there is no signal to indicate it. This means that there is at least one signal which is correlated with at least two events. To see this, start with a complete and sound system, such that each event is paired with one and only one signal and vice versa. Represent this collection with the following set of tuples: $\{(e_1, s_1), (e_2, s_2), \dots (e_n, s_n)\}$. So each s_x is generated if and only if e_x occurs.

Now, there are two ways to make this system incomplete. One consists of removing one tuple and adding another such that a signal is assigned to another event. For example, take out (e_2, s_2) and add (e_2, s_1) . Now e_2 does not have a corresponding signal but s_1 causes the system to be unsound because it is not uniquely paired with a single event. It might be thought that one can instead simply remove the signal s_2 by taking out the tuple (e_2, s_2) . Although this means that an event (e_2) is no longer correlated with an active signal, there is still something that carries the information that e_2 , namely the absence of an active signal; and an absence of primary data is still itself data.¹² Remove another signal s_3 and take out the tuple (e_3, s_3) and this means that the signal consisting of no active signal (silence) does not carry the information that e for any particular e (it carries the information that $e_2 \vee e_3$).

If a system is incomplete, then the number of signals (including the absence of a primary signal) is less than the number of events. From this it follows that the system is also unsound. Thus it is important that the quantification over signals in the above definitions of soundness and completeness include the ‘silent signal’.

5.2.3 Dealing with Knowledge of Necessary Truths

According to Dretske’s and similar accounts, logical, mathematical, and analytic truths generate no information. Therefore a signal cannot carry the information that x , where x is such a necessary truth. According to Dretske, this would suggest that no information is needed to know that $\sqrt[3]{27} = 3$.

Or, to put the same point differently, informationally speaking anything whatsoever is good enough to know a necessary truth. Bubba’s assurances are good enough to know that 3 is the cube root of 27 because his assurances carry all the information generated by that fact. Mathematical knowledge appears to be

¹²See Section 1.1.1 of Chapter 1

cheap indeed. [59]

This is no doubt an issue, one which Dretske suggests some ways to deal with in a discussion of epistemology and information [59]. Some of them revolve around adding a subjective element. In *Knowledge and the Flow of Information* the account of knowledge is restricted to contingent empirical knowledge.

Since a contingent fact is a fact for which there are possible alternatives, a fact that might not have been a fact, a fact that (because it has a probability less than one) generates information, one will always have a channel of communication between knower and known that is possibly equivocal, a channel that might mislead. If a theory of knowledge is a theory about this limited domain of facts, a theory (merely) of empirical knowledge, then communication theory is prepared to say something about an essential ingredient in such knowledge. It tells you what the channel between source and receiver must be like for someone at the receiver to learn, come to know, empirical facts about the source. [59, p. 26]

So we restrict the applicability of this account to such knowledge. An account of the information associated with necessary truths and the knowledge of such truths is another project. Amongst the tasks of such a project is an explication of the sense in which logical truths do yield information and an account of the difference between having ‘implicit’ access to all necessary-truth information and ‘explicitly’ accessing that information by deduction.¹³

For our purposes it is enough to simply accommodate knowledge of necessary truths. This is done straightforwardly. The fact that no signal can carry the information that x , where x is a necessary truth, does not mean that x is not information, or that it is information which an agent cannot use for knowledge. Rather, x is a piece of information that the agent already holds by default, that they don’t require a signal to provide them with. This is related to the presence of the necessitation rule in a logic of being informed [67]. Since an agent already holds the information that x , its empirical information measure is zero. So relative to an account of empirical information, think of being informed of necessary truths as holding ‘base information’. All we need to say is that the information that x , where x is a necessary truth which an agent holds to begin with, can lead to the belief and resultant knowledge that x .

Continuing on, we can say that:

¹³See [61, 164, 23] for some recent work.

1. Perception is a source of knowledge.
2. Memory is a common way of storing and retrieving knowledge
3. Reasoning and inference are effective ways for extending/gaining knowledge

Point 3 merits some further scrutiny. Whilst reasoning and inference are definitely effective ways for extending and gaining knowledge, we must consider that there are certain conditions which must be met in order for some piece of reasoning or an inference to qualify. Firstly, as has already been covered extensively, there are grounds for constructing an account in which information is not universally closed under deductive inference. But even aside from more complex debates such as this, there are certain conditions that must hold. Some of these are nothing new. As Russell pointed out

it is clear that a true belief is not knowledge when it is deduced from a false belief. In like manner, a true belief cannot be called knowledge when it is deduced by a fallacious process of reasoning, even if the premisses from which it is deduced are true. If I know that all Greeks are men and that Socrates was a man, and I infer that Socrates was a Greek, I cannot be said to know that Socrates was a Greek, because, although my premisses and my conclusion are true, the conclusion does not follow from the premisses. [159]

So in using reasoning and inference to extend and gain knowledge it is essential that

1. the premisses of the deduction are true
2. the process of deduction is valid

If either of these two fail to hold, whilst the result might be information_W , it is not information_S .

Within the framework of this informational epistemology a third condition arises. As is by now clear, information implies truth but truth does not imply information in the stronger sense. Given this, another condition to add is that the premisses used must all be information_S . For example, let A and B be true statements. Now, the following is certainly a sound deduction:

$$\frac{A \quad B}{A \wedge B}$$

But now suppose that A is a piece of information _{S} and B is a piece of information _{W} . It then follows that $A \wedge B$ is a piece of information _{W} . So if someone holds:

- A = the true semantic content that it is 6pm (acquired from a functioning clock)
- B = the true semantic content that the room's temperature is 20° (generated by a guess)

then they are in a position to know that A based on the information _{S} that A but are not in a position to know that B or to know that $A \wedge B$, both of which are information _{W} .

5.2.4 Knowledge, Information and Testimony

Testimony is an important way in which we come to know things about the world. Under the standard view testimony is the *transferring of knowledge* from a speaker to a hearer. This can be broken down into a sufficiency claim and a necessity claim [6]:

- (KN) If H knows that p by accepting S 's testimony that p , then S knows that p
- (KS) If S knows that p and H accepts S 's testimony that p , then H knows that p

Under the framework of information-theoretic epistemology, an alternative account to this has the following two counterparts:

- (IN) If H knows that p by accepting p on the basis of S 's testimony, then S 's assertion that p to H carries the information that p
- (IS) If S 's assertion that p to H carries the information that p and H accepts p on the basis of S 's testimony, then H knows that p

This alternative informational version involves treating knowledge from testimony in a similar way to knowledge from perception. It has been notably endorsed by Graham [88]. As evidenced by the following counter-examples he uses to undermine KN/KS and support IN/IS, an analysis of testimony based on information transmission could supplant one based on knowledge transmission. The following two examples are given to undermine the sufficiency of knowledge and support the necessity of information for testimony to result in knowledge.

In the Newspaper case, the military of a country plans on staging a coup. It

pays off or threatens the reporters of the country's newspapers to run a story that the President has been assassinated. All the reporters give in but Andy. He will report what *really* happens. The assassination attempt is successful and Andy is the only eyewitness; the other reporters do not know or even inquire into what actually happened. Andy writes in his by-lined column that the President was assassinated. [88, p. 370]

Jenny, who usually does not consult the newspaper that Andy is associated with, randomly happens to come across Andy's report and forms the true belief that the president has been killed. Since Jenny was lucky in consulting the one media source that was informed and knew that the president had been assassinated, it is fair to say that she does not know that the president was assassinated.

An information-theoretic analysis could accommodate this verdict by saying that Jenny's belief was not caused by the information that the president was killed. The signal consisting of a media report about the president's status does not carry the information that he has been killed given that there are alternatives in which the signal occurs but the president is not killed and these alternatives are relevant (e.g. a mishap results in the assassination attempt being botched and Jenny reading her usual newspaper, which falsely reports that the president has been assassinated). If this analysis holds, then we have a case where Andy knows that the president is assassinated, Jenny accepts that the president was assassinated on the basis of Andy's testimony, but Jenny does not know that the president was assassinated.¹⁴

In the second case, originally due to Dretske [60], George, a connoisseur of fine wines, knows a Medoc wine when he tastes one and also knows a Chianti wine when he tastes one. Hence he knows that there is a distinction between the two on the basis of their tastes. George knows that Medocs are Bordeaux, since Medoc is a region of Bordeaux. However he falsely believes that Chianti wine is also Bordeaux, because he mistakenly thinks that Tuscany is also a region in Bordeaux. George consumes some Medoc wine, correctly identifies it and comes to know that he drank wine from Bordeaux. A while later he is asked by his friend Michael what type of wine he drank, to which George replies it was from Bordeaux. It is claimed by Graham that although George knows that the wine was from Bordeaux, Michael does not know, since George's utterance does not carry the information that the wine he drank was from Bordeaux.

¹⁴This is related to the discussion in Section 5.2.1 concerning casual deviance and the loss of information in a causation chain.

This is an interesting example. As Graham admits, “many do not find the example convincing at first pass” [88, p. 389] based on the reason that George does not know in the first place because of his false geographical belief. Although Graham discusses and defends this example in more depth elsewhere [89], my analysis leads me to at least doubt that George does know that the wine he drank was from Bordeaux. Granted in this particular scenario, he does not use any false beliefs or information_W to deduce his belief. Furthermore,

the fact that Tuscany is not a region of Bordeaux is not a defeater for George’s belief that the wine he drank was a Bordeaux, for if he were to come to know that Tuscany is not in Bordeaux he would still know that Medoc is in Bordeaux and would still believe that the wine, which he correctly identified as a Medoc, was a Bordeaux. [89, p. 137]

Coady [35] offers an objection to this case, arguing that George does not know:

It is not clear how we can treat as true the claim, ‘it is possible that Michael should have had this very evidence but his belief should have been false’, without making it true also that ‘it is possible that George should have had this very evidence but his belief have been false’. We could treat Michael’s evidence as merely some utterance of George’s to the effect that some wine he had at a dinner party was a Bordeaux, but then why not treat George’s evidence equally abstractly as some taste sensation which he identifies as belonging to a type of wine which he believes to be a Bordeaux? To do so is to allow that in both cases the evidence might remain the same and the belief be false. [35, p. 228]

So according to Coady, once we describe the example at the right level then George’s evidence is no better than Michael’s. Thus George’s evidence is not that the wine tastes like a Medoc and is therefore from Bordeaux, but that the wine tastes like a Bordeaux. Since based on this evidence George would believe that it tastes like a Bordeaux even when a Tuscan Chianti, he does not know. But as Graham (whose response I favour) argues, since George can distinguish between the wines, his evidence is the more accurate ‘tastes like a Medoc’, so Coady’s objection does not hold.

According to Graham,

what the example shows is that when the speaker can rule out more relevant alternatives than the hearer, a speaker can sometimes know that P and sincerely

state that *P* but not enable his hearer to learn something on the basis of his testimony. This is because the speaker's belief, though it is caused or causally sustained by a signal that carries the information that *P*, does not itself carry the information that *P*; and it does not carry the information that *P* because, in this case, the speaker has a false general belief that sometimes causes him to believe that *P* when not *p*. Because his belief does not carry the information that *P*, when he states that *P* his statement will not carry the requisite information either. Hence he will know something even though he cannot bring his audience to know it by communicating the belief to them. [89, p. 138]

It is right to say that the signal consisting of the Medoc taste carries the information that the wine is a Bordeaux for George. Furthermore, since George bases his belief on this signal his belief is caused by the information that the wine is a Bordeaux. Michael on the other hand bases his belief on George's word, which does not carry the information, since there are relevant alternatives in which George says that the wine is a Bordeaux when it is not.

But focusing on George, we can see that an argument against him having knowledge can be based on similar reasons for denying that the boss has knowledge in the two William Smith/Jones examples I discussed in Section 5.2.1. The second of these examples in particular is structurally similar to this wine example. In my discussion of those examples I suggested the requirement of belief information carriage and that since the boss' beliefs did not carry the information that Jones was selected the boss did not have knowledge. So there seems to be a tension here between the intuition I worked with in discussing that example and Graham's view in this example. If definition ITK_3 and the requirement of belief information carriage is too strong then that leaves ITK_2 as a suitable qualified definition.

Despite the problematic nature of this second example from Dretske/Graham, I maintain that the first newspaper example suffices to undermine KS and support IN.

The third example given by Graham undermines the necessity of knowledge and supports the sufficiency of information. It relies on cases where one person who has a 'Gettierised' belief that *p* tells another person that *p*. Although the source's belief is Gettierised, their statement carries the information that *p* and the receiver is in a position to know that *p*. Judy and Trudy are identical twins who work at the library. Bill also works in the library, knows both of the twins and can tell them apart. Susan on the other hand also works in the library but only knows Judy; she is not aware that Judy has an identical twin. Susan works in the same section of the library as Judy; although Trudy does occasionally work in that section (there has never been an occasion where Susan saw both of them). One evening,

Judy and Trudy are at work in adjacent sections of the library re-shelving books. Susan can see only Judy and Bill can see only Trudy. Judy knocks over a shelf (P), something which both creates a big crashing sound and Susan sees. Bill, who also hears this crash, calls Susan and asks if she knows what happened, to which she replies that Judy knocked over a shelf.

Now, given that Trudy knocking over the shelf (Q) is a relevant alternative that Susan cannot rule out, she does not know that Judy knocked over the shelf. Bill on the other hand does come to know this, for Susan's testimony carries for Bill the information that Judy knocked over the shelf. For someone, unlike Bill, who did not know that it could not have been Trudy, Susan's statement carries the information that $P \vee Q$. Since Bill knows that $\neg Q$, he gets the information that P .

This case thus undermines the necessity of knowledge and supports the sufficiency of information. There is an objection to Graham's usage of this case that involves the distinction between *accepting* a report and *making use of* a report, but he does fend it off [88, p. 372].

At any rate, it does not take much to think of examples where H comes to know that p via S's testimony even though S does not even believe that p . Take a regularly moving clock which is fifteen minutes fast. So, for example, the signal '6:15pm' is correlated with the event 6pm. John, who does not know that the clock is fast, looks at it and forms the false belief that it is 6:15pm. Bob, who knows that both the clock is fifteen minutes fast and that John does not know this, asks John what time he sees, to which John responds 6:15. Now, it is clear that this utterance carries the information that it is 6:00 and that Bob can access this information. Accordingly, the true belief that Bob forms is knowledge. One man's misinformation is another man's information.

This discussion shows that an information-theoretic account of knowledge can contribute to and (re)shape our accounts of testimony. Here are some other points to consider.

Over this chapter and the last we have established that under a Dretskean account of information and knowledge, one can know that p without knowing that they know that p . This failure of the KK-principle is a result of the fact that a signal which carries some piece of information on which the knowledge is based will not be self-authenticating, will not necessarily carry the information that it is carrying the information. With regards to this section, this means that one does not necessarily know that they know if the knowledge in question is based on testimony. We also have seen that for reasons related to the failure of the KK-principle, knowledge is not closed. With regards to testimony, this means that one can come to know that p on the basis of another's testimony without knowing that the source of the testimony is not misinforming or disinforming them.

Another effect results from the fact that if one does not have the information themselves then they cannot be in a position to make a transmission that itself will result in knowledge. This results in conceivable cases where expert testimony that is truthful does not lead to knowledge because information was lost somewhere along the chain.

Take the following example, which plays on the classic Russellian example of a consulted broken clock which happens to display the actual time. A famous historical event occurred at time x . Someone present recorded the time at which the event occurred and it is this record which the history books use. As it turns out, the clock that the recorder used was broken and happened to be stuck on the actual time. Now, although the recorder recorded a fact (a true proposition), according to the information-theoretic analysis they did not know nor did they have the information that the event occurred at time x . If the recorder did not have the knowledge or information that the event occurred at time x , is what they recorded a piece of information that can be consulted to gain knowledge? If their record is used in an encyclopaedia entry about the event, can someone who reads this entry fifty years later come to know that the event occurred at time x . A faithful reading of the information-theoretic account would suggest that since there was no information in the first place, there was no information to be stored or transmitted. So what the encyclopaedia contains is just true semantic content, not a piece of information that can lead to the knowledge that the event occurred at time x .

The picture painted here can be a stringent one. Isolated blemishes in a chain of communication can result in a cessation of the flow of information hence the absence of knowledge after the blemish. For example, suppose that a piece of information p (information_S, true semantic content that was generated from a signal carrying that information) is passed down via informed scribes though the centuries. Each scribe is informed of this piece of information and reliably rewrites it, thereby passing on the information. At one stage the duty to pass on this information falls to an irresponsible scribe, who is generally drunk. He reads the information whilst drunk before accidentally burning the piece of paper that contains the information. Being in a drunken state, he cannot remember if it was the case that p or $\neg p$. He takes a guess and writes down p , thereby luckily passing on the true content. Under our paradigm of information and knowledge the information that p is not passed on and the next scribe who comes into the possession of this message does not know that p .

Whilst it might seem that requirements of this account can all too easily and harshly rob recipients of knowledge, it should be noted that there is room for flexibility. This account can serve as a paradigm to guide usage of the hazy term ‘knowledge’ as applied in the imperfect ‘real world’. There is no perfect circle in the world, yet we have a precise conception of this geometrical entity and deploy this conception in determining whether or not something

is a circle in the real world. Similarly, this Platonic notion could apply to knowledge; an intelligible conception of the form of knowledge is required in order to judge real-world approximations.

So beyond the possibility that we have covered of employing contextualism in order to set relevant alternatives, other forms of gradation can feature in practical applications of the term ‘knowledge’. ‘Pure knowledge’ might be lost if the requirements do not strictly obtain, yet granted a sufficient percentage of requirement adherence a sufficient ‘impure knowledge’ might be generated. Whilst a community that has relied on the scribes in order to gain the knowledge that p might not have the ideal knowledge that p , the fact that their reception of p has come via a highly reliable lineage means that they have high grade ‘impure’ knowledge. Just like a triangle whose sides add up to 179.99 degrees is a higher grade triangle than one whose sides add up to 179 degrees, knowledge of p resulting from an information chain with one drunk scribe is of a higher grade than knowledge resulting from an information chain with ten drunk scribes.

5.3 Epistemic Logics for Informational Epistemologies

The development of epistemic logics that capture the type of informational epistemology we have covered is an interesting open task. In this section I would like to:

1. Give a simple contextualist epistemic logic for contextualist informational epistemology in general
2. Explore the prospects for developing an epistemic logic which captures Dretske’s notion of knowledge as a semi-penetrating operator.

Our starting point is the standard treatment of epistemic logic as a type of modal logic which can be represented using possible world semantics [100]. The universal modal operator K_a is such that K_ap stands for ‘Agent a knows that p ’. As usual, K_ap holds at a world/state w if and only if p holds at all states accessed by w . Each accessed state w' can be seen as an epistemic alternative to world w for agent a .

Commonly, the type of modal logic used is a normal one. Starting off with the base normal modal logic **K**, axioms are added in accordance with the desired properties of knowledge and the epistemic agents. One essential axiom is that corresponding to reflexive frames, which simply expresses the veridicality of knowledge; if a knows that p , then p is true: $K_ap \supset p$. Two other axioms and their corresponding epistemic interpretations are:

- $K_ap \supset K_aK_ap$ (4) - Introspection: a knows what they know.
- $\neg K_ap \supset K_a\neg K_ap$ (5) - Negative Introspection: a knows what they don't know.

Now, there is a veritable range of limitations and issues with this basic modal approach to formally capturing knowledge and epistemic logic has thus developed into several other approaches and branches involving a variety of sophisticated logical tools and methods. Setting these issues aside, in this section we work with basic accounts of epistemic logic and focus on getting the right validities and invalidities pertaining to an informational epistemology such as Dretske's.

We begin with some preliminary observations on what would such a logic look like. The first and most prominent issue is closure, associated in epistemic logic with the K axiom:

$$K_a(p \supset q) \supset (K_ap \supset K_aq)$$

which leads to the following validity:

$$K_ap \wedge K_a(p \supset q) \vdash K_aq$$

Closure will not be universally valid in the type of epistemic logic we are interested in.

Secondly, since the KK principle is rejected for externalist epistemologies such as Dretske's, introspection [the (4) axiom] will not be universally valid. We do not consider negative introspection here, which is excluded from many standard treatments of epistemic logic for good reason. At any rate, similar reasons used here for rejecting introspection could apply to negative introspection. One uncontroversial axiom though in such a logic is the T axiom $K_ap \supset p$. This all leaves us with a rather stripped down epistemic logic.

5.3.1 Alternative Sets of Relevant Alternatives and Multi-Modal Logic

In the previous chapter, multi-modal logic was used to analyse information flow. The upshot was that different boxes can correspond to different sets of relevant alternatives and that the validity of properties such as transitivity would be relative to the boxes associated with each of the information statements constituting the premises. It takes little leap of the imagination to realise that this type of thing could also be done with epistemic operators.

With this approach, epistemic operators are of the form K_x^y , for each agent x ($a, b, c, \dots$) and each context/set of relevant alternatives y (1, 2, 3, ...). The modal logic for each operator is at least **T**, which is **K** + T .

Since each operator K has $Kp \supset p$ as an axiom, truth is still preserved under closure:

$$K_a^x p \wedge K_a^y (p \supset q) \vdash q$$

Under this approach, the (Dretskean) failure of closure could be represented by invalidities of the form:

$$K_a^x p \wedge K_a^y (p \supset q) \not\vdash K_a^z q$$

where $\neg(x = y = z)$. So across contexts, closure fails.

For the contextualist, epistemic closure holds within an epistemic context:

$$K_a^x p \wedge K_a^x (p \supset q) \vdash K_a^x q$$

But once the context is shifted and the stakes are raised, then closure does not apply, since the initial proposition is no longer known: $\neg K_a^y p \wedge \neg K_a^y \neg p$.

Once again, an axiom such as $K_a^2 p \supset K_a^1 p$ would represent the fact that the set of relevant alternatives associated with 1 is a subset of the relevant alternatives associated with 2 and would give the following validities:

- $K_a^1 p \wedge K_a^2 (p \supset q) \vdash K_a^1 q$
- $\vdash K_a^2 K_a^1 p \supset K_a^1 K_a^1 p$

For contextualists, the KK principle could hold within the one context: $K_a^x p \supset K_a^x K_a^x p$. For example, relative to standard possibilities where only normal zoo animals are kept in the zoo's enclosures, a comes to know that there is a zebra in the enclosure. Given the restricted set of relevant alternatives, by default the visual signal of a zebra carries the information that it is carrying this information and that it is true that there is a zebra. This enables the subject to know that they know relative to the context. With one-context introspection and given $K_a^2 p \supset K_a^1 p$, the following is valid: $K_a^2 p \equiv K_a^2 K_a^1 p$.

Given this, rather than using $K_a^1 K_a^1 p$ to represent meta-knowledge that p , a fitting representation of this would be $K_a^2 K_a^1 p$, which would lead to the following:

$$K_a^2 K_a^1 p \wedge K_a^2 (p \supset q) \vdash K_a^2 q$$

How could this be used in the zebra-mule example? One goes to the zoo, sees what they believe to be a zebra in front of them and acquires the knowledge that there is a zebra in the enclosure. This true belief is knowledge relative to the elimination of a certain set of relevant alternatives (that associated with K_a^1). Constituting this set of relevant alternatives are standard possibilities where only normal zoo animals are kept in the zoo's enclosures. Now, suppose that K_a^2 ranges over a superset of alternatives, not only those involving normal zoo animals, but those involving all types of animals. Clearly the following type of closure is invalid:

$$K_a^1 p \wedge K_a^2 (p \supset q) \not\vdash K_a^2 q$$

But now the person in question goes one step further and performs a DNA test on the animal they see. This test confirms that the animal is indeed a zebra and rules out the possibility that it is any other type of animal. This person can be said to have knowledge about their lower-level knowledge; by confirming that a certain set of irrelevant alternatives do not obtain, they have a form of meta-knowledge. Closure relative to K_a^2 will then hold.

Continuing on, we can also use this logical apparatus to easily provide an acceptable representation of DeRose's [45] *abominable conjunctions*. An abominable conjunction is a conjunction that results from denying closure. In general, they are of the form $Kp \wedge \neg K\neg q$, where p is an ordinary proposition and q is a heavyweight proposition implied by p . The following is an example:

I know that I have hands (p), but I do not know that I am not a handless brain in a vat (q)

If one utters this, then each conjunct is made relative to a different context. Whilst inconsistent within one context, the following representation of this is not inconsistent:

- $K_a^1 (p \supset q) \wedge K_a^2 (p \supset q) \wedge K_a^1 p \wedge \neg K_a^2 \neg q$

There has been some, but to my knowledge little, work done on contextual epistemic logic.¹⁵ Whilst the multi-modal system approach I have suggested is relatively unsophisticated compared to such other proposals, it is flexible and I think that it sufficiently captures the basic idea of a logic for contextualist epistemology. That being said, it is a contextual logic perhaps best suited to contextualism proper or radical contextualism, as opposed to Dretske's modest contextualism [55]. For Dretske, sets of relevant alternatives are not variables that can be set in accordance with any context set by discussion and knowledge attributors. Rather, the set of relevant alternatives for a proposition is determined by the proposition and the subject's situation. Despite the variability between relevant alternative sets for different propositions, ultimately there is one knowledge statement Kp to judge for each proposition p , not a range of candidates $\{K_a^1 p, K_a^2 p, \dots\}$. With this in mind, it is time to explore the possibility of developing an epistemic logic with only one K operator per agent that still satisfies all the conditions of a Dretskean epistemology.

5.3.2 'Knows that' as a semi-penetrating operator

As early as his 1970 paper *Epistemic Operators* [49], Dretske treated knowledge as a *semi-penetrating* operator. So, whilst closure fails for knowledge, the following two hold:

- $K(p \wedge q) \vdash Kp \wedge Kq$ (Conjunction Distribution)
- $Kp \vdash K(p \vee q)$ (Disjunction Introduction)

Although it is not mentioned explicitly, since Dretske endorses a conjunction principle for information, we will also experiment with the following conjunction principle for knowledge:

- $Kp \wedge Kq \vdash K(p \wedge q)$ (Conjunction Closure)

With these three principles serving as yardsticks, we turn to the task at hand, starting with a look at one attempt in the literature (the only serious one that I am aware of) to formalise a Dretskean epistemic logic.

5.3.3 One Approach to Epistemic Logic and Relevant Alternatives

Holliday [109] offers an interesting proposal to formalise a Dretskean epistemic logic. Using a variation of modal preference logic [18], he formalises two versions of the relevant alternatives

¹⁵[154] is one paper that I am familiar with.

theory, the Dretskean version and the Lewisian contextualist version.

Where W is a set of possible worlds and a proposition P is understood as a subset of W (so that the complement of P in W , $W \setminus P$, is the proposition not- P), his starting point is the following characterisation of the two positions:

- Dretskean - $\mathbf{RA}_{\forall\exists}$ - for every proposition P , there is a relevancy set $R_C(P) \subseteq (W \setminus P)$ such that in order to know P one must rule out $R_C(P)$
- Lewisian - $\mathbf{RA}_{\exists\forall}$ - there is a set of relevant worlds R_C such that for every proposition P , in order to know P one must rule out $R_C \cap (W \setminus P)$

Here are the key details of the language used to formalise these positions.

- $\Box \sim \phi$ means that ϕ is true in every possibility consistent with the agent's information
- $\Box \prec \phi$ means that ϕ is true in every strictly-more-relevant possibility
- $R_{\sim}$, the accessibility relation for $\Box \sim$ is an equivalence relation (S5)
- $R_{\prec}$, the accessibility relation for $\Box \prec$ is a total preorder (S4.3)
- $wR_{\sim}v$ means that possibility v is consistent with the agent's information at w
- $wR_{\prec}v$ means that alternative v is at least as relevant as alternative w
- $\max_{\preceq}(S) = \{v \in S \mid u \preceq v \text{ for all } u \in S\}$ is the set of most relevant alternatives in a set $S \subseteq W$
- Given some model M , $[\phi]_M$ is the truth set of ϕ : $[\phi]_M = \{v \in W \mid M, v \models \phi\}$

The truth definitions for Dretske-knowledge (K_d) and Lewis-knowledge (K_l) are:

- (Dretske) $M, w \models K_d \phi$ iff $\max_{\preceq}(W \setminus [\phi]_M) \cap \{v \mid wR_{\sim}v\} = \emptyset$
- (Lewis) $M, w \models K_l \phi$ iff $\max_{\preceq}(W) \cap (W \setminus [\phi]_M) \cap \{v \mid wR_{\sim}v\} = \emptyset$

With the definition of Dretske-knowledge, $\max_{\preceq}(W \setminus [\phi]_M)$ stands for the set of maximally relevant $\neg\phi$ worlds. If the intersection between these relevant $\neg\phi$ worlds and the possibilities consistent with the agent's information $\{v \mid wR_{\sim}v\}$ is null, then the agent has ruled out the relevant alternatives as required in order to know that ϕ .

With Lewis-knowledge, there is a general set of maximally relevant alternatives $\max_{\preceq}(W)$ to begin with. For any ϕ , in order to know that ϕ an agent must rule out the $\neg\phi$ worlds in that set.

K_d and K_l are definable using the following formulas:

- $K_d\phi =_{df} \Box_{\sim}(\neg\phi \supset \Diamond^{\prec}\neg\phi)$
- $K_{dl}\phi =_{df} \Box_{\sim}(\neg\phi \supset \Diamond^{\prec}\top)$

They can be read as follows. If $K_d\phi$ is the case, then if $\neg\phi$ is true in an accessed possibility, there must be some other possibility that is more relevant with regards to $\neg\phi$ (accessed possibilities in which $\neg\phi$ is true cannot be maximally relevant with regards to $\neg\phi$).

If $K_l\phi$ is the case, then if $\neg\phi$ is true in an accessed possibility, there must be some other possibility that is more relevant in general. This is guaranteed using $\Diamond^{\prec}\top$.

Some of the main results of this system are as follows:

- $K_l\phi \wedge K_l(\phi \supset \psi) \not\vdash K_l\psi$
- $K_d\phi \wedge K_d(\phi \supset \psi) \not\vdash K_d\psi$
- $K_d(\phi \wedge \psi) \not\vdash K_d\phi \wedge K_d\psi$
- $K_d\phi \not\vdash K_d(\phi \vee \psi)$

These results suggest that the Dretskean knowledge operator is not even semi-penetrating. Of course, it could just be that this limited proposal does not adequately capture Dretske's account. One offhand observation is that it does nothing to consider implementing a connection between the relevant alternatives of propositions. For example, if $A \vdash B$ then would the relevant alternatives for A form a subset of the relevant alternatives for B ?

Holliday writes:

It remains a challenge for defenders of Dretske's views to formulate a plausible version of the RA theory according to which closure of knowledge under known implication fails while closure under conjunction elimination and disjunction introduction holds. [109, p. 11]

For the sake of attempting to give Dretske's views a full and fair go, we shall continue with the interesting challenge of finding an epistemic logic which captures Dretske's notion of knowledge, one in which closure of knowledge under known implication fails whilst conjunction distribution and disjunction introduction hold.

5.3.4 Going Non-normal

The K axiom $[K(p \supset q) \supset (Kp \supset Kq)]$ is a mainstay of normal modal logic and its removal brings us into the territory of non-normal modal logic. Since the absence of K means the absence of closure, employing a non-normal modal might seem like one way to go.

Let us start this investigation with the minimal non-normal modal logic **E**, which is based on neighbourhood semantics [141, 33]. It consists of:

- Propositional calculus (PC)
- The rule (RE): $\frac{p \equiv q}{Kp \equiv Kq}$, where $p \equiv q$ is a theorem

To this straightforwardly add the T axiom: $Kp \supset p$.

Next, let us add the axiom for Conjunction Distribution:

- $K(p \wedge q) \supset Kp \wedge Kq$ (M)¹⁶

The presence of (M) gets us the rule called *Right Monotonicity*: $\frac{p \supset q}{Kp \supset Kq}$, where $p \supset q$ is a theorem.

Proof.

- | | | |
|----|--------------------------------------|---------------------------------|
| 1. | $p \supset q$ | Given theorem |
| 2. | $p \equiv p \wedge q$ | Follows from 1 using PC |
| 3. | $Kp \equiv K(p \wedge q)$ | Using RE on 2 |
| 4. | $K(p \wedge q) \supset Kp \wedge Kq$ | M rule |
| 5. | $Kp \supset Kp \wedge Kq$ | PC transitivity using 3 and 4 |
| 6. | $Kp \supset Kq$ | |

□

¹⁶This axiom corresponds to neighbourhood frames being closed under supersets

So the addition of M also gets us Disjunction Introduction $Kp \vdash K(p \vee q)$, since $p \supset (p \vee q)$.

Next, let us add the axiom for Conjunction Closure:

- $Kp \wedge Kq \supset K(p \wedge q)$ (C)¹⁷

We find that Conjunction Distribution and Conjunction Closure cannot both be added without getting K and closure:

Proof.

- | | | |
|----|--|------------------------------------|
| 1. | $Kp \wedge K(p \supset q)$ | Given assumption |
| 2. | $K(p \wedge (p \supset q))$ | C on 1 |
| 3. | $(p \wedge (p \supset q)) \equiv p \wedge q$ | PC reasoning |
| 4. | $K(p \wedge q)$ | RE and PC reasoning on 2 and 3 |
| 5. | $Kp \wedge Kq$ | M on 4 |
| 6. | Kq | |

□

So at least one of (M) or (C) has to go. The choice to keep (M) over (C) is a clear one and there is a variety of reasons for such a choice. To begin with, as Williamson observes:

... it is far more plausible that knowing a conjunction requires knowing the conjuncts than that knowing the conjuncts requires knowing the conjunction; one may fail to put two pieces of knowledge together. [176, p. 29]

Of course, motives such as the last are not ones based on concerns particularly relevant to Dretskean informational epistemology. Nonetheless, there is a case to be made for rejecting Conjunction Closure based on such relevant concerns. One point to begin with is that as was shown above the addition of (M) for Conjunction Distribution also gets us Disjunction Introduction, another desired property of semi-penetrating knowledge. So what we end up with is pretty much the same non-normal modal logic that was introduced at the end of Section 4.4.4, depending on whether or not N is included; call this system **EMT** (add N if included).

¹⁷This axiom corresponds to frames being closed under finite intersections

More importantly for the case against (C), in Section 4.4.4 of the previous chapter it was shown how the conjunction principle for information flow can fail to hold. Using the zebra-mule scenario, the following example was given:

$$V \sqsupset Z, V \sqsupset (Z \wedge \neg M) \vee (M \wedge \neg Z) \not\models V \sqsupset Z \wedge ((Z \wedge \neg M) \vee (M \wedge \neg Z))$$

If the conjunction principle for information flow can fail to hold then it follows that knowledge can also fail to close under conjunction. One could know that z and know that $(z \wedge \neg m) \vee (m \wedge \neg z)$ without being in a position to know their conjunction because they do not have the information that is their conjunction.

Also, denying the validity of (C) means the following invalidity: $Kz \wedge K(z \supset \neg m) \not\models K(z \wedge (z \supset \neg m))$. So without (C) in a logic such as **EMT** closure fails: $Kz \wedge K(z \supset \neg m) \not\models K\neg m$. We can interpret this by saying that if one knows that there is a zebra based on the visual information and that not-mule is a channel condition, then this piece of knowledge should not be ‘mixed’ with a knowledge statement that zebra entails not-mule. One is empirical knowledge that depends on a certain channel condition and the other is analytic knowledge of a conditional that involves the channel condition as a consequent.

Furthermore, we can also see that a principle of Conjunction Closure for knowledge does not sit well with certain combinations involving iterated knowledge. With regards to the standard zebra scenario, it is fair to think that Dretske would assert the following; the first two are explicit in his account and the third presumably follows from the second:

1. Kz
2. $\neg K\neg m$
3. $K\neg K\neg m$

If Conjunction Closure were to hold here then the resulting $K(z \wedge \neg K\neg m)$ is awkward and problematic; relative to what set could this statement be made? The case for keeping (M) over (C) is thus clear and the failure of conjunction closure is one consequence of Dretske’s account.

We can apply this analysis to one of the debates in the literature concerning closure. Hawthorne [97] argues that rejecting closure entails rejecting distribution of conjunction. Start with the following:

The Equivalence Principle: If one knows *a priori* (with certainty) that P is equivalent to Q and knows P, and competently deduces Q from P (retaining one's knowledge that P), one knows Q

Hawthorne claims that Dretske's reasons for denying closure do not affect this principle and suggests that Dretske will accept it:

[Dretske's] argument against closure relies on the following idea. Following recent usage, let us say that R [a conclusive reason] is "sensitive" to P just in case were P not the case, R would not be the case. Suppose one believes P on the basis of R, and that P entails Q. R may be sensitive to P and still not to Q. But notice that where P and Q are equivalent, there can be no such basis for claiming that while R can underwrite knowledge that P, it cannot underwrite knowledge that Q. [97, p. 31]

The second principle is distribution:

Distribution: If one knows the conjunction of P and Q, then as long as one is able to deduce P, one is in a position to know that P (and as long as one is able to deduce Q, one is in a position to know that Q).

Hawthorne reasons that although distribution is extremely plausible, Dretske is committed to denying it. As can be gathered, his reasoning actually seems to conform to certain aspects of a normal modal logical framework for epistemic logic. But as I have made the case it is reasoning that conforms to a certain non-normal modal logic system that is required.

His example is the following:

Suppose one knows that some glass g is full of wine on the basis of perception (coupled, perhaps, with various background beliefs). The proposition that g is full of wine is *a priori* equivalent to the proposition

g is full of wine and $\sim$ g is full of non-wine that is colored like wine

So by equivalence one knows that conjunction. Supposing distribution, one is in a position to know that

$\sim g$ is full of non-wine that is colored like wine

Despite Hawthorne's argument, Dretske does not ultimately have to give up conjunction distribution. To begin with, Hawthorne's wine example has the same structure as Dretske's original zebra example that we have been working with, where wine corresponds to zebra and non-wine corresponds to mule. So we will return to the zebra-mule example.

In [5] the authors argue, in concordance with my own analysis, that Dretske's account actually entails giving up the equivalence principle instead of conjunction distribution. According to them, in cases like the zebra-mule scenario Dretske's reasons for denying closure do have force against the equivalence principle. Let:

- $p =$ a certain animal x is a zebra
- $q = x$ is not a painted mule
- $R = x$ appears to be a zebra (R is a conclusive reason)

As they write:

Suppose (i) R is sensitive to p , that is, if it were not the case that x is a zebra, x would not appear to be a zebra; (ii) S knows p on the basis of R ; (iii) S knows *a priori* that p if and only if $p \& q$; (iv) S competently deduces $p \& q$ from p ; and (v) if x were a painted mule, x would appear to be a zebra. Although the Equivalence Principle implies that S knows $p \& q$, the following considerations show that Dretske's analysis of knowledge implies that S doesn't know $p \& q$ because R is insensitive to $p \& q$. In virtue of (i) and (v), it follows that if it were not the case that x is a zebra, then x would not be a painted mule (for otherwise x would appear to be a zebra). But if it were not the case that x is both a zebra and not a painted mule, i.e., if it were the case that x is either a non-zebra or a painted mule, then x might be a painted mule, in which case x would appear to be a zebra. Consequently, R is insensitive to $p \& q$. That is, if it were not that case that x is a zebra and not a painted mule, x might appear to be a zebra. Thus, even though the Equivalence Principle implies that S knows that $p \& q$, contrary to Hawthorne's contention, Dretske's reasons for denying closure *do have force* against the Equivalence Principle. [5]

So once we get our heads around the unorthodox logic that accompanies Dretske's denial of closure we can see that his reasons for denying closure do have force against the Equivalence

Principle. In informational terms, the point being made here is that whilst $V \sqsupset Z$ and $Z \equiv Z \wedge \neg M$, it is not the case that $V \sqsupset (Z \wedge \neg M)$, which goes together with it not being the case that $(\neg Z \vee M) \sqsupset \neg V$. Thus one can know Z without knowing $Z \wedge \neg M$, despite their analytic equivalence.

The non-normal modal logic we have arrived at offers a way to get the right validities/invalidities and technically deal with the issue Hawthorne raises. It both validates The Distribution Principle and invalidates The Equivalence Principle. The Distribution Principle is validated given that it is an axiom. The Equivalence Principle however is not valid: $Kp \wedge K(p \equiv q) \not\vdash Kq$. In terms of our example, it is not the case that z is logically equivalent to $z \wedge \neg m$. Rather, it is $z \wedge (z \supset \neg m)$ which is logically equivalent to $z \wedge \neg m$. So by logic we have $K((z \wedge (z \supset \neg m)) \equiv (z \wedge \neg m))$ to begin with but not $K(z \equiv (z \wedge \neg m))$ and it doesn't boil down to whether or not z is known but rather whether or not $z \wedge (z \supset \neg m)$ is known. But as we have seen, given the rejection of conjunction closure, although we have Kz and $K(z \supset \neg m)$ it does not follow that $K(z \wedge (z \supset \neg m))$. Also, it is worth noting that $K(z \equiv (z \wedge \neg m))$ breaks down into $K(z \supset (z \wedge \neg m))$ and $K((z \wedge \neg m) \supset z)$ anyway, so it seems that Hawthorne has implicitly performed a form of closure in arguing for closure.

As was suggested above, one way to understand the failure of conjunction closure here is in terms of certain conditions precluding the joining of certain types of knowledge statements. The knowledge statement $K(z \supset \neg m)$ represents knowledge of an analytic truth. Knowing this also means knowing that if z is true then $\neg m$ is true. But the knowledge statement Kz here simply means that the agent has the externalist empirical knowledge that z given the channel condition $\neg m$, not that they know that z is true. If on the other hand Kz were replaced with the higher-order knowledge statement KKz and this implies knowing that z is true, then the inference $KKz, K(z \supset \neg m) \vdash K\neg m$ could be made. This is just a suggestion and the technical details would need to be fleshed out.

If $z \supset \neg m$ (as well as all such analytic truths) was made a theorem of the language, representing a system in which such analytic truths are true in all models, then closure for the zebra-mule case would go through due to Right Monotonicity. In this case one option would be to sacrifice conjunction distribution, though this would be a very significant cost. Looking back at the proof of Right Monotonicity, we see that proof theoretically Conjunction Distribution could be retained in this case if the *(RE)* rule $[p \equiv q \Rightarrow Kp \equiv Kq]$ was removed. Obviously removing this axiom could not be accomplished within the semantics of this modal logic framework. It would also mean the loss of Disjunction Introduction.

For a proposal such as Dretske's, there are limits to a logic where $p \equiv q \Rightarrow Kp \equiv Kq$. Even if z and $z \wedge \neg m$ are (truth) equivalent propositions, if they are judged relative to

different sets of relevant alternatives or channel conditions then they are not equivalent with regards to their informational status. Although z qualifies this does not necessarily mean that $z \wedge \neg m$ qualifies. In the next section we will take an initiating look into how one might go about developing a system that models this phenomenon more accurately.

5.3.5 A Logic for Dretskean Epistemology

Experimentation with a non-normal epistemic modal logic in the last section facilitated our investigation into the logical nature of a Dretskean epistemology. Beyond this, there is room to look into developing formal epistemic logics that mimic the structure of Dretske's epistemology more closely. Such attempts could incorporate higher-order or extra-logical resources. It goes without saying that Dretske's account of knowledge poses a challenge to formal epistemic logic. It is perhaps a good example of the separation between 'mainstream' and 'formal' approaches to epistemology and the potential benefits of bringing the two into closer contact ¹⁸. In this section I would like to offer some final observations and suggestions regarding the matter.

One important consideration is that under Dretske's account knowledge statements are made relative to certain channel conditions/relevant alternatives. These conditions depend on how the knowledge was acquired and need to be factored into the assessment of a deduction. Based on the analysis that has been carried out here, another important point is that the propositions that represent channel conditions cannot be a logical consequence of propositions that are known.

The zebra-mule argument $Kz, K(z \supset \neg m) \vdash K\neg m$ cannot be judged on logical grounds alone. What needs to be ascertained are the conditions, if any, that accompany the premises. For example, the conditions accompanying Kz depend on the method used to gain the knowledge. If the statement Kz is made on the basis of visual information as per the standard zebra-mule scenario, then it follows that $\neg m$ is a channel condition and therefore $K\neg m$ cannot be deduced. If, on the other hand, the statement Kz is made on the basis of a DNA test, then it follows that $\neg m$ is not a channel condition and therefore $K\neg m$ can be deduced.

So there are two things to incorporate here. Firstly, knowledge statements could be accompanied by certain channel conditions. Secondly, channel conditions cannot be a logical consequence of propositions that are known. One way to address these requirements involves simply extending the standard epistemic logic definition of the K operator and adding extra

¹⁸See [101]

conditions to reflect these requirements. What results is a system whereby the knowledge operator is broken down into constituent parts. Let us implement this approach and take the constituents of a knowledge statement to be the following:

1. The modal component of standard epistemic logic
2. The condition that if one knows that p it cannot be the case that p is a channel condition.
3. The condition that if one knows that p then it cannot be the case that there is some proposition x such that $p \vdash x$ and x is a channel condition.

So we have:

$$Kp =_{df} \Box_K p \wedge \neg[p] \wedge \neg(\exists x)((p \vdash x) \wedge [x])$$

$\Box_K$ is that of some standard normal epistemic modal logic. $[p]$ stands for ‘ p is a channel condition’. The existential quantification occurs over all propositions in the domain.

Formalising this idea

Here is one simple way to formalise this idea. We start off by introducing the two modal operators that are to be used in the definition of knowledge.

1. $\Box_K$ corresponds to the K operator in standard epistemic logic and here is that of the modal logic **T**.
2. $\Box$ will be used to check logical consequence relations and is that of the modal logic **K**.

Next, determine a (possibly empty) set of channel conditions $X = \{c_1, c_2, \dots, c_n\}$ for the argument to be evaluated. Corresponding to this set are two formulas $C_1 = c_1 \wedge c_2 \wedge \dots \wedge c_n$ and $C_2 = \neg c_1 \wedge \neg c_2 \wedge \dots \wedge \neg c_n$. If this set is empty then $C_1 = C_2 = \top$ (*verum*).

The definition of knowledge now incorporates and is relative to C_2 :

$$Kp =_{df} \Box_K p \wedge \Diamond(p \wedge C_2)$$

The $\Box_K$ component simply deals with knowledge before externalist preclusions are brought in. The $\Diamond(p \wedge C_2)$ component ensures that p does not logically imply one of the channel conditions.

There is one final requirement. Let Γ stand for a collection of premises. If C_1 represents channel conditions that are relevant to an argument $\Gamma \vdash_{C_2} B$ that is to be evaluated, then $C_1 \in \Gamma$. The turnstile is subscripted with C_2 to signify the incorporation of C_2 into the definition of K .

Following is a demonstration of this system to the zebra-mule scenario. For the sake of this demonstration, we will include two channel conditions. The first is the standard condition that the creature is not a disguised mule ($\neg m$) and the second is the condition that it is not a robot zebra that perfectly imitates a real zebra ($\neg r$). So let $X = \{\neg m, \neg r\}$, $C_1 = \neg m \wedge \neg r$ and $C_2 = m \wedge r$.

Here are some results:

Result	Comment
$C_1, Kz, K(z \supset \neg m) \not\vdash_{C_2} K(z \wedge (z \supset \neg m))$	Failure of closure under conjunction
$C_1, Kz, K(z \supset \neg m) \not\vdash_{C_2} K\neg m$	Failure of closure under implication
$C_1, Kz, K(z \equiv (z \wedge \neg m)) \not\vdash_{C_2} K(z \wedge \neg m)$	Failure of the equivalence principle
$C_1, K(z \wedge t) \vdash_{C_2} Kz \wedge Kt$	Conjunction distribution holds
$C_1, K(z \wedge \neg m) \vdash_{C_2} \perp$	It is not possible to know a proposition that logically entails a channel condition
$C_1, Kz \vdash_{C_2} K(z \vee \neg m)$	If one disjunct is a channel condition, this does not preclude disjunction introduction:
$C_1 \vdash_{C_2} \neg K(\neg m \vee \neg r)$	If both disjuncts are channel conditions, then the disjunction cannot be known
$C_1 \vdash_{C_2} \neg Km$	If p is a channel condition, then $\neg p$ cannot be known

The basic idea developed in this section offers a way to directly capture what is going on in Dretske's account. Before closing, it is worth mentioning the possibility of using the definitions discussed in Section 5.2 and defining the knowledge that p as the information

caused/carrying belief that p . One rough way to express this using doxastic logic and our modal logical definition of information carriage is:

$$Kp =_{df} Bp \wedge \Box(Bp \supset p)$$

$Kp \supset p$ follows from this definition and it also satisfies the criterion that knowledge is more than true belief.

5.4 The Value of Information and Knowledge

The value of knowledge has always been a topic of interest within epistemology. One of the big questions is the value problem that concerns explaining what it is about knowledge (if anything) that makes it more valuable than mere true belief [151]¹⁹. Given our informational account of knowledge, this translates more specifically to explaining why information caused/carrying belief is more valuable than mere true belief. Using the distinction introduced in Section 4.5 of the previous chapter, this could also be construed in terms of the value information_S has over information_W . Before closing this chapter I would like to offer some suggestions as to why (ITE) knowledge (information_S) is better than mere true belief (information_W), focusing on the instrumental and practical value that the former has over the latter. This will serve to contribute to the ITE cause, since an adequate response to this value problem has become a criterion in modern epistemology for assessing accounts of knowledge.

Whilst the value of true belief has also been scrutinised, I will here take it as a given that truth (information) is good. To begin with, it has instrumental value, as true belief is a primary factor in successful action. We generally prefer to form and act upon true beliefs over false ones since we are more likely to get what we want. As William James remarked, truth “is the good in the way of belief” [113, p. 42]. On top of this, it is reasonable to think that truth is more than instrumentally valuable. Most of us at one time or another have wanted to know the truth of a matter for its own sake, indeed as Michael P. Lynch writes, “curiosity is not always motivated by practical concerns” [128, p. 502]. Even if there were no practical consequences or differences between a belief being true or false, we would surely prefer it to be true and we would also prefer to not believe in trivial falsehoods [128, p. 502]. Furthermore, it can be said that we have a second-order desire for having true beliefs, in that we not only desire the truth, but desire to desire the truth [128, p. 504]. Given all of

¹⁹A.K.A. The Meno Problem

this, it is fair to say that true belief is an epistemic good and truth an appropriate object of value.

Since knowledge encapsulates truth, it follows that if truth is an appropriate object of value, then so is knowledge. This much is clear. However it is widely thought that knowledge is of distinctive value, that the value of knowledge is distinct not only from the value of its core constituent true belief, but also from whatever falls short of knowledge. Explaining just why knowledge is distinctly valuable, *the value problem*, is to what I shall now turn.²⁰

It is firstly important to note that when arguing for the case that knowledge is more valuable than true belief, we are not committed to saying that knowledge is *always* of greater value than corresponding true belief, for cases are conceivable where the true belief that *p* is of no lesser value than the knowledge that *p*. Rather, we are committed to the more modest task of arguing that knowledge is typically or sometimes of greater value than corresponding true belief. An answer to the value problem will be sought here in terms of instrumental value. Taken in isolation, the mere true belief that *p* will be of just as much instrumental value as the knowledge that *p*. After all, as Socrates notes in Plato's *Meno*, a true belief about the correct way to the city of Larissa is of just as much instrumental value as knowledge of the way to Larissa; both, if acted upon, will get one to Larissa. Despite this, I contend that overall (i.e. not just with regards to the goal of a particular true belief) the knowledge that *p* will generally encompass a greater quantity of instrumental value (in terms of true belief) than the mere true belief that *p*. In order to discuss this idea in more detail, consideration of the 'swamping problem' serves as a good starting point. A standard explanation of this problem is given with the following example by analogy due to Linda Zagzebski:

Imagine two great cups of coffee identical in every relevant respect - they look the same, taste the same, smell the same, are of the same quantity, and so on. Clearly, we value great cups of coffee. Moreover, given that we value great cups of coffee, it follows that we also value reliable coffee-making machines - i.e., machines which regularly produce good coffee. Notice, however, that once we've got the great coffee, we don't then care whether it was produced by a reliable coffee-making machine. That is, that the great coffee was produced by a reliable coffee-making machine doesn't contribute any additional value to it. In order to see this, note that if one were told that only one of the great identical cups of coffee before one had been produced by a reliable coffee-making machine, this would have no

²⁰The term *value problem* is taken from Duncan Pritchard, who uses it in his paper 'The Value of Knowledge' [152]. In this paper, he defines three levels of the value problem. The *primary value problem* is to explain why knowledge is more valuable than mere true belief. The *secondary value problem* is to explain why knowledge is more valuable than any proper sub-set of its parts. The *tertiary value problem* is to explain why knowledge has not just a greater amount but also a different kind of value than whatever falls short of knowledge. Here we are only concerned with the value problem as the primary value problem.

bearing on which cup one preferred; one would still be indifferent on this score. In short, whatever value is conferred on a cup of coffee through being produced by a reliable coffee-making machine, this value is ‘swamped’ by the value conferred on that coffee in virtue of it being a great cup of coffee. [153, p. 3]

Although this example specifically relates to reliabilist theories of knowledge, a more general point can be extracted and applied to other epistemological accounts. The gist is that if a property [like that of being reliably formed for beliefs] is only instrumentally valuable relative to some further good (e.g., true belief or great coffee), then in cases where the further good in question is acquired, the presence of the instrumentally valuable property confers no further value. So as another example, if justification is only instrumentally valuable relative to true belief, then in cases where true belief is acquired, the justification involved in its formation confers no further value.

Whilst plausible *prima facie*, it would be rash to conclude that the swamping problem in effect shows that there can be nothing (including knowledge) more epistemically valuable than mere true belief. Before offering some suggestions as to why knowledge is more valuable than mere true belief, I shall go over the case for why we should at least seek knowledge. This can also serve as a revisionary response.²¹

5.4.1 Knowledge and True Belief Generation

Irrespective of whether or not knowledge is distinctively valuable, there is justification for its key place in epistemological theorising, given that truth is a basic epistemic good which has value. One need only appeal to the teleological relationship between value and the deontic, where the deontic is antecedent to and explained in terms of the evaluative. Since true beliefs have value, it is good to have true beliefs. Since it is good to have true beliefs, we ought to pursue knowledge, since it is truth conducive and maximises the chances that our beliefs attain their goal of truth and avoid falsity. All of this is summed up in the following simple claim:

If you want to have true beliefs then you ought to have knowledge

²¹In addressing the value problem, one can seek a *validatory* answer by offering an explanation of why knowledge is distinctively valuable or a *revisionary* answer by offering an explanation of why although knowledge is *not* distinctively valuable, we are inclined to think of it as so [150, p. 1]. So it might not be necessary for knowledge to be pitted in a victorious ‘value battle’ against mere true belief in order to justify our particular focus on knowledge in epistemological theorising.

This then can explain the intuition that knowledge seems better than mere true belief. Marian David suggests something along similar lines when he writes that these intuitions²² “arise due to a confusion of sorts. They do not reflect any bonus of intrinsic value accruing to knowledge over and above ... [mere] true belief ...; rather they reflect our desire to have our desires satisfied” [42, p. 310]. Thus there is a *normative* difference between true belief (the good) and knowledge (the ought).²³ In light of this, it is not hard to see why knowledge, an epistemological ought (because it is truth conducive), is analysed and sought.

Estimating Truth Conduciveness

The truth conduciveness of knowledge, or more generally a reliable belief-forming process, is a straightforward matter. If we want to form true beliefs and avoid false ones, then obviously a method of information-caused belief is the right way to go. Whilst it is not impossible for unreliable methods to be just as truth conducive in some situation, they are less likely to be. Over the long run, those who form their beliefs based on information-carrying sources are probably better off.

Of course truth conduciveness corresponds to our notion of informativeness. Using some of the tools and methods developed in earlier chapters, we can give these ideas a simple formal expression. In Section 2.6 of Chapter 2, we looked at combining truthlikeness with decision theory in order to calculate the estimated informativeness of a statement when the actual state, against which information calculations are made, is not completely given. Apart from this, there is another way to use this estimation method in order to measure the truth/information conduciveness of a process.

In order to demonstrate this idea, we will use the Basic Features approach to truthlikeness covered in Section 2.4.1 of Chapter 2, where the similarity of statement A to constituent C_i is defined as:

$$s_\tau(A, C_i) = \tau \text{cont}_t(A, C_i) - (1 - \tau) \text{cont}_f(A, C_i)$$

Truthlikeness is defined as similarity against the true constituent:

²²David lists the following four intuitions, of which only the first is directly relevant to the current discussion

1. Knowledge seems better than mere true belief
2. Justified true belief seems better than unjustified true belief
3. Unjustified false belief seems worse than justified false belief
4. Unjustified true belief versus justified false belief? Intuition hesitates.

²³This idea is explored in [145]

$$\text{Tr}_\tau(A) = s_\tau(A, C_*)$$

For our purposes the parameter τ is such that $0 < \tau \leq 1$. Following is an example to demonstrate this method.

Example 5.1

John, Bob and Harry are taking a short quiz in which there are three true/false questions. Let the propositions representing questions 1, 2 and 3 be p_1 , p_2 and p_3 respectively. The answer to each question is true, so $S_1 = p_1 \wedge p_2 \wedge p_3$ is the true constituent/state description for this erotetic set, whose truth table is given in Table 5.1.

State	p_1	p_2	p_3
w ₁	T	T	T
w ₂	T	T	F
w ₃	T	F	T
w ₄	T	F	F
w ₅	F	T	T
w ₆	F	T	F
w ₇	F	F	T
w ₈	F	F	F

Table 5.1:

John studied comprehensively and knows the answer to every question. Bob on the other hand did not study at all, so when the time comes to answer each question all Bob can do is take a guess. Harry studied, but not comprehensively; he studied thoroughly the sections corresponding to the first two questions but not the section corresponding to the last question. Thus for each question, the probability that John will answer it correctly is 1 and the probability that Bob will answer it correctly is 0.5. Harry will answer the first two questions correctly, but has only a 50% chance of answering the last question correctly. As a result of these figures, what will be the estimated truthlikeness measures of their response sheets?

Let X stand for some belief or semantic content generating method. Then the estimated truthlikeness of the content generated by employing X is given by:

$$\text{Tr}_{est}(X) = \sum_{i=1}^n \text{Tr}(S_i) \times \text{Pr}(X(S_i))$$

where S_i stands for a state description, $\text{Pr}(X(S_i))$ stands for the probability that state

description i is generated using X and n stands for the number of possible states.

Setting $\tau = 0.5$ for calculations and starting with John's case, the only instance of S_i such that $\Pr(X(S_i)) \neq 0$ is $\Pr(X(S_1)) = 1$. Thus, using this formula, the estimated truthlikeness of his response sheet will be the maximal:

$$\text{Tr}(S_1) \times \Pr(X(S_1)) = 0.5$$

In Bob's case, since for each of the eight possible state descriptions there is a $(\frac{1}{2})^3$ chance it will result from his generation method, the estimated truthlikeness of Bob's response sheet will be:

$$(0.5 \times \frac{1}{8}) + 3(\frac{1}{6} \times \frac{1}{8}) + 3(-\frac{1}{6} \times \frac{1}{8}) + (-0.5 \times \frac{1}{8}) = \frac{1}{16} + \frac{1}{16} - \frac{1}{16} - \frac{1}{16} = 0$$

Given Harry's distribution as depicted in Table 5.2, the estimated truthlikeness of his response sheet will be:

$$(\text{Tr}(S_1) \times \Pr(X(S_1))) + (\text{Tr}(S_2) \times \Pr(X(S_2))) = (0.5 \times 1 \times \frac{1}{2}) + (0.5 \times \frac{1}{3} \times \frac{1}{2}) = \frac{1}{4} + \frac{1}{12} = \frac{1}{3}$$

State	p_1	p_2	p_3	$\Pr()$
w ₁	T	T	T	$\frac{1}{2}$
w ₂	T	T	F	$\frac{1}{2}$

Table 5.2:

Evidently, more reliable information-caused beliefs lead to a higher estimated truthlikeness measure.

□

With Harry's case in the last example, the random selection of an answer is effectively equivalent to a suspension of judgement on the final question. If Harry were to not respond, then the calculation would also amount to:

$$\text{Tr}(p_1 \wedge p_2) \times \Pr(X(p_1 \wedge p_2)) = 0.5 \times \frac{2}{3} \times 1 = \frac{1}{3}$$

So in this case there is no difference between Harry's choices of guessing the last answer or not responding. If parameters were adjusted to reflect a bias towards truth or against falsity, then Harry's choice becomes more interesting. Should he gamble on guessing the last answer in the hope that it is correct or should he exercise risk aversion? Here is a simple decision theoretic formulation of these scenarios, starting off with the relevant calculations:

- $\tau = 0.5$
- $s_\tau(p_1 \wedge p_2 \wedge p_3, w_1) = s_\tau(p_1 \wedge p_2 \wedge \neg p_3, w_2) = \frac{1}{2}$
- $s_\tau(p_1 \wedge p_2 \wedge \neg p_3, w_1) = s_\tau(p_1 \wedge p_2 \wedge p_3, w_2) = (0.5 \times \frac{2}{3}) - (0.5 \times \frac{1}{3}) = \frac{1}{6}$
- $s_\tau(p_1 \wedge p_2, w_1) = s_\tau(p_1 \wedge p_2, w_2) = 0.5 \times \frac{2}{3} = \frac{1}{3}$

The decision problem for the first scenario is tabulated in Figure 5.3, with the utilities being $s_\tau()$ for each statement/state pair.

	w ₁	w ₂
$p_1 \wedge p_2 \wedge p_3$	$\frac{1}{2}$	$\frac{1}{6}$
$p_1 \wedge p_2$	$\frac{1}{3}$	$\frac{1}{3}$
$p_1 \wedge p_2 \wedge \neg p_3$	$\frac{1}{6}$	$\frac{1}{2}$
Pr()	0.5	0.5

Table 5.3:

As calculations show, all three of the possible outcomes resulting from Harry's two choices have the same expected utility: $EU(p_1 \wedge p_2 \wedge p_3) = EU(p_1 \wedge p_2 \wedge \neg p_3) = EU(p_1 \wedge p_2) = \frac{1}{3}$

If τ were adjusted such that $\tau = 0.4$ to reflect a greater penalty for falsity, the decision problem becomes that tabulated in Table 5.4.

	w ₁	w ₂
$p_1 \wedge p_2 \wedge p_3$	0.4	$\frac{1}{15}$
$p_1 \wedge p_2$	$\frac{4}{15}$	$\frac{4}{15}$
$p_1 \wedge p_2 \wedge \neg p_3$	$\frac{1}{15}$	0.4
Pr()	0.5	0.5

Table 5.4:

The estimated utility calculations become:

- $\text{EU}(p_1 \wedge p_2 \wedge p_3) = \text{EU}(p_1 \wedge p_2 \wedge \neg p_3) = \frac{7}{30} = 0.23$
- $\text{EU}(p_1 \wedge p_2) = \frac{4}{15} = 0.267$

Clearly Harry would be better off not answering.

If τ were adjusted such that $\tau = 0.6$ to reflect a greater reward for truth, the decision problem becomes that tabulated in Figure 5.5.

	w ₁	w ₂
$p_1 \wedge p_2 \wedge p_3$	0.6	$\frac{4}{15}$
$p_1 \wedge p_2$	0.4	0.4
$p_1 \wedge p_2 \wedge \neg p_3$	$\frac{4}{15}$	0.6
Pr()	0.5	0.5

Table 5.5:

The estimated utility calculations become:

- $\text{EU}(p_1 \wedge p_2 \wedge p_3) = \text{EU}(p_1 \wedge p_2 \wedge \neg p_3) = \frac{13}{30} = 0.43$
- $\text{EU}(p_1 \wedge p_2) = 0.4$

Clearly Harry would be better off taking a guess.

Measuring Reliability Requirements

Extending this system, we can develop a way to measure the reliability adequacy requirement of some source or content generation method. Suppose that we put a minimum threshold of acceptability on the estimated truthlikeness of the content generated by a source or process. The highest value of estimated truthlikeness is τ , so this threshold can never be greater than τ , which would be the measure for a completely reliable source.

Take a source/process from which n bits of data are to be collected, each datum corresponding to a proposition that is true or false. The estimated truthlikeness of the content generated by the source/process depends on its success rate and τ . Let x stand for the probability that a datum given by the source/process is correct; that is, its success rate. Then the estimated truthlikeness (*et*) of the content generated by the source/process is measured with the following:

$$et = x + \tau - 1^{24}$$

Example 5.2

A machine on a production line assesses each instance of some product for a certain defect. If the product does not have the defect, then it passes the test and the machine shows a green light. If the product does have the defect, then it fails the test and the machine shows a red light. If the machine is 100% reliable, then $et = 1 + \tau - 1 = \tau$. If the machine is only 80% reliable, then $et = 0.8 + \tau - 1 = \tau - 0.2$ $\square$

Now, suppose that standards in this factory require that the et measure meets a certain minimum threshold of acceptability. Originally, this minimum is set with $\tau = 0.5$ and $et = 0.3$. The required success rate is therefore 80%.

There are two ways to modify these parameters so that a change to reliability must be made in order to meet the minimum threshold of acceptability:

1. The threshold of minimum acceptance is raised.
2. The value of τ is changed.
 - As an example of 1, if the threshold were raised to 0.4 and $\tau = 0.5$, then the required success rate would increase to 90%.
 - As an example of 2, if $et = 0.3$ and τ were changed to 0.4 in order to reflect a greater penalty associated with false test results, then once again the required success rate would have to increase to 90%.

5.4.2 The Value of Knowledge

The upshot of the previous section is hardly a revelation; we ought to pursue informations and knowledge because we want to have truth. But as the swamping argument suggests, it would seem that once the true content/belief has been acquired/generated, the fact that it was caused by a reliable information source confers no further value onto the content/belief.

$$^{24}x + \tau - 1 = \sum_{i=0}^n \left[\binom{n}{n-i} \times \left(\tau \frac{n-i}{n} - (1-\tau) \frac{i}{n} \right) \times (x^{n-i}(1-x)^i) \right]$$

In other words, true semantic content seems to be no more valuable because it was generated from an information carrying source.

In the case of the factory quality assurance machine, this means that even if its signal was not providing information due to unreliability, if it happened on one particular day to give nothing but correct results, this set of results is just as valuable as a set that was obtained from a completely reliable machine.

So we ask, is there something about an information-caused belief that p which can make it more valuable than its mere true belief that p counterpart? A comprehensive investigation into the matter is beyond the means of this work. The main aim is to show how the informational epistemology given here could address the question.

Reliabilist accounts of knowledge were targets of early contemporary work on the value problem. It has been claimed, particularly by epistemologists with alternative accounts of knowledge, that reliabilism is unable to offer an answer to the value problem. In particular, virtue epistemology has been touted as an alternative account of knowledge that is better suited to dealing with the value problem.²⁵ I am not sure to what extent an account of knowledge should be judged based on how it deals with the value problem. The significance of this problem might very well have been overplayed by those with accounts of knowledge that are tailored towards answering it. Nonetheless, it definitely is a question worth looking into. Since ITE is a form of reliabilism, the points made here will serve to help the reliabilist cause in this debate.

Whilst I think that some arguments can be made in terms of instrumental epistemic value, it goes without saying that if knowledge does not ultimately have any more instrumental value than mere true belief, then an appeal to some other form of value would be required. It should be pointed out that whilst intrinsic value, with which instrumental value has been traditionally contrasted with, is the obvious alternative, other categories of value which have been indentified might be applicable. For example, *final value*, “the value that something has as an end or for its own sake” [182] has been considered in the literature [151].²⁶

Before moving onto the specific arguments I have in mind, it is worth remarking that without considering any particular account of knowledge, perhaps a more general reason, using some form of value, for why we might value knowledge over mere true belief could be countenanced. For example, here is one offhand suggestion that appeals to a crucial difference between cases like the swamping problem example, in which the coffee is a product of a coffee

²⁵See [151] for an overview.

²⁶Final value is a form of non-instrumental, extrinsic value. The Mona Lisa is more finally valuable than an exact replica because it was created by Leonardo da Vinci, an extrinsic property of the object.

machine and cases in which belief is a product. This difference concerns the nature of the products' origins/causes. In the case of the coffee, its cause is of no value significance, for as far as one is concerned the machine is purely and totally an instrument for the production of the product (i.e. the coffee). With beliefs however, the causes, the originators of beliefs are humans, and although humans are instrumental in belief formation, they are also agents of value judgements and are bound to assess and judge their products (e.g. beliefs) not only intrinsically, but also with regard to extrinsic factors like the viability and quality of their production methods. A small modification to the coffee example will elucidate this point. Say, we replace the coffee making machines with two people. One person, Bob, has never made a coffee in his life. Another person, John, is one of the world's finest baristas. John, as usual, produces an optimal cup of coffee. Bob, as luck has it, throws a bunch of coffee ingredients together and produces a fluke coffee that is just as good as John's. Despite the equivalence of the instrumental values of these coffees, I get the feeling that one would opt to consume a coffee put together by the world's best barista rather than a coffee produced by someone who is not associated with the principles of coffee making excellence. The barista-made coffee would have a special final value because of who made it. The only unique value that Bob's 'luck coffee' would have is *shock value*. To further hone in on this point, consider if Bob were to make the effort to go to a barista school. Upon graduation, he goes to the effort to gather the finest coffee ingredients and puts together a coffee which is of the same instrumental and intrinsic value as the coffee he randomly produced in the previous example. Surely Bob would nonetheless value this coffee over the 'luck coffee', because of his efforts in its causal history. Likewise, the principles of knowledge are the principles of true belief formation excellence. Hence if someone forms a true belief by following these principles, they are going to value their product more than a corresponding product which does not have these principles and efforts associated with it. These features of knowledge, such as the factors in its causal history and its intimate connection with the truth out there, would explain its added value.

This aside, it is now time to turn to the specific arguments I have in mind concerning why ITE knowledge is more *valuable* (or, perhaps more fittingly, *better*²⁷) than mere true belief. The basic idea is that an instance of knowledge that *p* typically 'breeds' more truth (value) and less falsity (disvalue) than a corresponding instance of mere true belief that *p*. The two related arguments I have in mind are roughly as follows:

1. Knowledge based on information is conducive to the generation of more knowledge, hence more of something with epistemic value

²⁷This can be seen as a stylistic preference which also reflects the view that the intuition that knowledge has something over mere true belief could very well be explored without delving too deeply into sophisticated value-theoretic notions.

2. True belief formation is often followed by action based on that true belief. This action (which is prompted by the true belief), will often be followed by the need to form another true belief. If the original true belief was caused by an information source, then the next belief can also be caused by that information source. So such true beliefs are more instrumentally valuable because they lead you to situations where you can re-use what got you there in the first place to once again acquire a true belief. Thus information-caused beliefs offer a kind of ‘guarantee’.

To get a better hold of the issue, let us begin with a formulation of the swamping problem given by Pritchard, in terms of an inconsistent triad of claims [153, p. 5]:

(1) The epistemic value conferred on a belief by that belief having an epistemic property is instrumental epistemic value relative to the further epistemic good of true belief.

(2) If the value of a property possessed by an item is only instrumental value relative to a further good and that good is already present in that item, then it can confer no additional value.

(3) Knowledge that p is sometimes more epistemically valuable than mere true belief that p .

Now, since it follows from (1) and (2) that (3) is false, these claims are jointly inconsistent. So what modifications could be made to resolve this inconsistency? Claim (2) seems right, therefore I will take it as the first to remain. Claim (3) is the first to question. Whilst (3) is a claim which I would ultimately like to endorse, even if it were to be rejected, there are still other notions of value besides epistemic value which one could appeal to in order to support the claim that all in all knowledge has greater value than true belief.

For instance, knowledge can be said to have greater *practical* value than mere true belief. This type of response is taken up by Plato in the *Meno*, where it is noted that although a true belief about the correct way to Larissa is of just as much instrumental use as knowledge of the way to Larissa, someone who has knowledge about the way to Larissa might have certain practical advantages over someone who only has a true belief about the way to Larissa. For example, if the path starting off on the journey looks dubious, the person without knowledge might be tempted to turn back, whereas the other person will continue down the path certain with the knowledge that it is the correct path [151]. So despite the swamping problem, any greater value of knowledge over mere true belief need not be understood only in terms of instrumental value relative to the good of true belief, for there could be other benefits of having knowledge over mere true belief.

Given this, one option is to accept (1) and (2) whilst replacing (3) with a more general claim about the greater value of knowledge, “that knowledge is generally of greater all things considered value than mere true belief whilst simultaneously granting that from a purely epistemic point of view there is no additional value to be had” [153, p. 7].

But we shouldn’t be too hasty in settling for such a relaxing modification to (3). To begin with, the claim only requires that we explain why knowledge is *sometimes* more epistemically valuable than mere true belief, a less burdensome task than explaining why knowledge is *always* more epistemically valuable than mere true belief. Also, consider the following situation:

Suppose that someone comes to you and says that in a moment one of two scenarios will obtain: either one will have a true belief that p or one will have knowledge that p (where one does not know which proposition is at issue). Furthermore, it is stipulated that all the practical consequences are kept fixed in both scenarios, so there will be no practical benefit to choosing the one option over the other. Nevertheless, shouldn’t one choose knowledge rather than mere true belief? [153, p. 8]

Unlike those who defend (3), those who deny it outright or settle for a relaxed version as just discussed, seem unable to account for this intuition. So let us make it clear that this investigation into the distinctive value of knowledge is in terms of epistemic value. Predictably this leads to consideration of claim (1). This claim is susceptible to scrutiny by those who can argue that a value conferred on a belief by that belief having an epistemic property can be of a suitable value type other than *instrumental* epistemic value.

But rather than considering this type of strategy, I would like to expound an argument for the case that knowledge that p is sometimes more epistemically instrumentally valuable than mere true belief that p . I would say that where in this apparently inconsistent triad of claims this argument fits is a leak resulting from the fact that (1) and (2) do not speak sufficiently for knowledge to rule out (3).

Given our definition of knowledge, the knowledge that p , because it results from the information that p , is generally of more instrumental epistemic value than mere true belief that p , because it generally encompasses or leads to more true beliefs and less false beliefs, it is ‘truth conducive’. To help get this point across, I would first like to emphasise that as I see it, it seems that the swamping problem is limited in what it covers. The following quote exemplifies the general question asked: “our question is why a true belief that is knowledge has more value than a mere true belief. The answer has to specify a property of

that belief" [112]. But comparing the knowledge that p against the mere true belief that p is not just about comparing the true belief that p which is a constituent of knowledge against the mere true belief that p . Instead, it is about comparing the *knowledge* that p , including all it involves, against the mere true belief that p . Thus, one can argue that there is something, other than the constituent true belief that p , which is involved with the knowledge that p and which makes it, *ceteris paribus*, more valuable. In terms of information, whilst instances of information_W stand alone and need not be connected to each other, instances of information_S are often embedded in a network and necessarily connected so that one piece of information_S in the network leads to another and so on.

Of course, there are cases where knowledge is no better than mere true belief. For example, suppose that you are looking at a jar filled with jelly beans and correctly guess that there are 214 jelly beans in the jar. You would then come to have the mere true belief that the jar contains 214 jelly beans. On the other hand, if you were to open the jar and count each jelly bean, you would know that the jar contains 214 jelly beans. Whilst counting the number of jelly beans is a safer way to form the true belief, once the true belief is generated, it is no better for being knowledge than mere true belief. But the point remains that sometimes, if not often, knowledge that p is more truth conducive than mere true belief that p . Here are a couple of examples to help convey this idea. They involve both instances of empirical information and instances of *a priori* information.

In the first example, the knowledge that p in some sense necessarily requires the knowledge, hence true belief, that q ($Kp \supset Bq$). Suppose someone thinks about the question 'how many factors does the number 152 have?'. If they were to form the mere belief that '152 has 6 factors', their belief would be true, requiring no other true beliefs. If, on the other hand, they wanted to *know* that '152 has 6 factors', they would proceed by calculating the factors of 152. In doing so, it follows that they would also come to have the true belief that '1, 2, 4, 38, 76 and 152 are factors of 152'. So the knowledge here involves intertwined true beliefs; the knowledge that 152 has 6 factors, unlike the corresponding mere true belief, goes hand-in-hand with other true beliefs.

Similarly, another way that pieces of knowledge are connected can be found in cases where knowledge of a quantitative datum implies knowledge of a qualitative datum. For example, imagine the following situation. A customer enters a deli and asks the store attendant if they could see the heaviest piece of cheese available. The store attendant goes into the back room and sees two pieces of cheese, which are clearly the largest. But they are so similar in shape and size that the attendant cannot determine which of the two pieces is larger. Now take the following two sub-scenarios:

- The attendant guesses and forms the true belief that piece A is the largest. They bring piece A to the customer, who decides that they would like to buy the piece. Cheese price is proportional to weight, \$1.00 per 1kg. This piece of cheese happens to weight 4.05kg, so it would be priced at \$4.05. But since the attendant does not know the weight of the piece, then they have to either take a guess or go back and make the measurement.
- Since they cannot determine the largest piece by visual observation, the attendant uses a precise scale to measure the weights of the cheese pieces. Piece A weighs 4.05kg and piece B weighs 4.00kg. Therefore, with the knowledge that piece A weighs more than piece B, the attendant selects piece A. In this case, the knowledge that piece A is the heavier piece involves the knowledge that it weighs 4.05kg. This knowledge can then in turn be used to price the piece immediately upon the customer opting to buy it.

Here, the information-caused beliefs that piece A weighs 4.05kg and piece B weighs 4.00kg lead to the information-caused belief that piece A weighs more than piece B. So the knowledge that piece A weighs more than piece B is intertwined with knowledge of the weight of each piece.

In this way, the acquisition of one piece of information_S guarantees the acquisition of another piece of information_S since the latter is nested in the former. It follows from this that the knowledge that *A* based on the information that *A* encapsulates the knowledge that *B*. In the end, it was having the information-caused knowledge rather than the mere true belief that piece *A* is the largest that resulted in a more successful outcome for those concerned.

These two examples concern the first of the two arguments I have in mind. The epistemic value (EV) of the knowledge that *p* is being assessed in terms of not only the constituent true belief that *p*, but also any other true beliefs that it encompasses or leads to. Put simply, if the knowledge that *p* also encompasses or leads to the true belief that *q* and $EV(\text{Belief}(p)) + EV(\text{Belief}(q)) > EV(\text{Belief}(p))$, then the knowledge that *p* would overall have greater epistemic value than the mere true belief that *p*, in virtue of the fact that it generates a greater overall quantity of true belief. You get two beliefs for one as it were. Under these terms, whilst the knowledge that *p* would not be of greater value than the mere true belief that *p* *plus* the mere true belief that *q*, such a response would miss the point. For if there is possibly going to be a greater total value (in the way of extra true beliefs) associated with the knowledge of a *single* proposition as opposed to the mere true belief of *that single* proposition, then that is technically grounds for the claim that knowledge that *p* is of greater epistemic value than mere true belief that *p*.

Further to this, knowledge is not only going to result in more true beliefs, but also less false beliefs. As a simple example, take John, who is waiting at home for his friend Peter when he hears someone knocking on his front door. He believes that the person at the door is his friend Peter and hence forms the belief that the person at his door has a name that starts with ‘P’. As it turns out, the person at the door does have a name that starts with ‘P’, so his derived belief is true. However, the person who is at the front door is actually another one of John’s friends, namely Paul. Compared with the mere true belief which John has, that the person at his door has a name that starts with ‘P’, if John had had the knowledge of this fact instead, say, because he knew it was Paul due to a certain pattern to the knock that only Paul does, then John would have had two true beliefs instead of one true belief and one false belief.²⁸

Linda Zagzebski writes “in the sense most commonly discussed by reliabilists, truth conduciveness is a function of the number of true beliefs and the proportion of true to false beliefs generated by a process” [181, p. 465]. So couched in these terms, the point here is that the knowledge that p is more *truth conducive* than the mere true belief that p , and truth conduciveness is translatable to epistemic value.

With regards to the second argument, dynamic information systems provide a great place to investigate the superiority of information_S over information_W. The next example hones in on the idea that knowledge, unlike mere true belief, is involved in the subsequent generation of other true beliefs. Thus, whilst the mere true belief that p and the knowledge that p might have the same immediate instrumental value, often the instrumental value of knowledge is potentially greater overall. Here we compare the information_S provided by the information-bearing signal of a functioning GPS navigator with the information_W generated by reading a malfunctioning GPS navigator.

We begin by considering a short trip from point A to point B, which requires three turns; two left followed by one right ($\text{left}_1, \text{left}_2, \text{right}_3$). Bob has a malfunctional GPS navigator, which constantly responds that the correct turn to take is a left one. Let L_S stand for a signal to turn left and L_A stand for left being the correct action to take. It is clear that $L_S \sqsubset L_A$ fails and there are relevant alternatives in which $L_S \wedge \neg L_A$. Since the GPS navigator always says left, its signal does not have the required correlation and a signal to turn left does not carry the information that left is the correct turn to take. Bob enters his origin point A and destination point B into the GPS navigation system. As a consequence of his reliance on a non-information bearing signal, he ends up correctly turning left the first two times before incorrectly turning left the third time. John on the other hand has a fully functioning

²⁸This type of example was adapted from Bertrand Russell [159, p. 131]. It is another early example of arguments found in Russell’s work on epistemology that cover Gettier-notions several decades before Gettier’s paper.

GPS navigation system. He turns the correct way each time and successfully arrives at the destination.

Obviously, in driving from one side of town to another with the aid of a GPS system, it is better to have a GPS system which is fully functioning than one which often malfunctions. In the examples above, the former always signals the correct turn to take whereas the latter sometimes misleads. But why is John's knowledge that left_2 is the correct turn to take better than Bob's mere true belief that left_2 is the correct turn to take?

Goldman and Olsson (G&O) [86] defend reliabilism against the charge that it cannot account for the extra value knowledge has over true belief. Given the connection between ITE and the type of reliabilism they support, it is worth incorporating their defence into the discussion. They offer two solutions to the value problem for reliabilism, the *conditional probability* solution and the *value autonomization* solution. It is the former of these which we shall look at.

According to this solution, a reliably produced belief gives a composite state of affairs with a certain epistemically valuable property that is missing in the case of mere true belief. The property in question is that a reliably formed belief makes it likely that one's future beliefs of a similar kind will also be true. More precisely, "under reliabilism, the probability [interpreted objectively] of having more true belief (of a similar kind) in the future is greater conditional on S's knowing that p than conditional on S's merely truly believing that p" [86, p. 28]. Letting F stand for 'future true beliefs of a similar kind', this can be expressed formally as:

$$\Pr(F|K_s p) > \Pr(F|B_s p \wedge p)$$

Fortunately, G&O also use a GPS example, so we can use it to continue:

Suppose you are driving to Larissa but are at loss as to which turns to take at various crossroads. On the way to Larissa there are two forks. If you choose correctly on both occasions, you will get to Larissa on time. If not, you will be late at best. Your only assistance in forming beliefs about the right ways to turn is the on-board computerized navigation system. We consider two situations differing only in that the navigation system is reliable in Situation 1 and unreliable in Situation 2. We assume that in both cases the navigation system tells you correctly how to turn at the first crossroads. In the first scenario this is to be expected, because the system is reliable. In the second it happens by chance.

Suppose the correct information at the first crossroads is ‘The best route to Larissa is to the right’. Hence in both situations you believe truly that the road to Larissa is to the right (p) after receiving the information. On the simple reliabilist account of knowledge, you have knowledge that p in Situation 1 but not in Situation 2. This difference also makes Situation 1 a more valuable situation (state of affairs) than Situation 2. The reason is that the conditional probability of getting the correct information at the second crossroads is greater conditional on the navigation system being reliable than conditional on the navigation system being unreliable. [86, p. 28]

The conditional probability approach evades the swamping issue because it does not rely on the reliable origin of the true belief conferring extra value onto the belief. Rather, it offers a reason for why the composite state of knowledge has more value than the composite state of mere true belief; as well as having the value of the constituent true belief, it has something else.

As G&O qualify, knowledge is more valuable than true belief under the conditional probability solution given that a number of empirical conditions are satisfied. These are *non-uniqueness*, *cross-temporal access*, *learning*, and *generality*. Non-uniqueness says that an encounter with a problem of a certain type will often be followed by encounters with problems of the same type at some point later on. The GPS scenario is a perfect example of this. Cross-temporal access says that once a method has been successfully employed in one encounter, it will usually be available to reuse down the track. The GPS navigation device satisfies cross-temporal access since it can clearly be re-used. Learning says that if you have successfully employed a method on one occasion you are likely to use it again on another occasion. The driver will keep on using the GPS navigation system at each turn. Finally, generality says that if a method is reliable on one occasion it will usually remain reliable for other similar situations. Again, in the GPS scenario, the functioning GPS device is reliable and will continue to work and provide navigation information.

Depending on the extent to which these conditions are satisfied, “the fact that S has knowledge on a given occasion makes it to some extent likely that S will acquire further true beliefs in the future” [86, p. 30]. This is not the case for unreliably produced belief. Whilst non-uniqueness and cross-temporal access are independent of the method’s reliability, generality certainly fails and learning is also affected. Although these conditions do not always hold, it is fair to say that they are normally satisfied. Given the aim of showing that knowledge is at least sometimes more valuable than mere true belief, that is enough.

There are a couple of ways that one can challenge the conditional probability solution.

One is to reject its central claim that $\Pr(F|K_sp) > \Pr(F|B_sp \wedge p)$. Horvath [110] presents such a challenge and Olsson [140] rebuts it in turn. I would say that there has been no convincing argument against the central claim so let us maintain that it is true.

Another option is to question whether the property identified by G&O actually adds more value to the composite state of knowledge. As we have seen, according to the conditional probability solution to the swamping problem, “reliabilist knowledge has surplus value in virtue of indicating not only the truth of the belief that was acquired but also the truth of future beliefs adopted in response to similar problems” [140, p. 112].

But such a quote can be misconstrued and the right reading of it is essential to understanding G&O’s position. Piller [146] questions whether the property identified by G&O actually adds more value to the composite state of knowledge. He works with the belief that according to the conditional probability solution “it would be a property of believing that p as a result of a reliable method that I will have a high chance of having true beliefs about similar matters in the future” [146, p. 130]. But, as Piller points out, it would seem that this value-grounding property is not a property of the belief itself but rather the method that it is formed by. It is not the knowledge constituting true belief itself that results in the prospect of future true belief. Rather, it is “the availability and appropriateness of the methods employed and the assumed fact that one will home in on such methods in forming one’s beliefs which assures us that the valuable property will be realized” [146, p. 128].

Piller rightly thinks that it is the method and G&O’s conditions (non-uniqueness, cross-temporal access, learning, and generality) that determine a higher chance of having future true beliefs; it is not the true belief itself. In the functioning GPS scenario, if for some reason the driver irrationally decided to form an isolated false belief contrary to the GPS navigation device’s information, it would not affect the likelihood of future true beliefs. In fact, even a headache (to use one of Piller’s examples), given the method and G&O’s conditions, indicates that the likelihood of future true beliefs is high.

But Piller is missing G&O’s underlying point. It is not that the knowledge *makes* future true beliefs more likely. Rather, it is that knowledge *indicates* that the likelihood of future true beliefs is high in a way that unreliably produced true belief does not. Olsson offers a really good explanation of this position. Interestingly, he does not invoke a form of instrumental or intrinsic value. He invokes what is termed ‘indicator value’, which is value belonging to something in virtue of its indicating something good. Rather than being some *ad hoc* concoction, this is a legitimate form of value identified by value theorists. As Zimmerman puts it:

Many philosophers write as if instrumental value is the only type of extrinsic value, but that is a mistake. Suppose, for instance, that the results of a certain medical test indicate that the patient is in good health, and suppose that this patient's having good health is intrinsically good. Then we may well want to say that the results are themselves (extrinsically) good. But notice that the results are of course not a means to good health; they are simply indicative of it. [182]

Likewise, reliably produced true beliefs are not a means to more true beliefs, they are simply indicative of it. It is in this sense that under the conditional probability solution knowledge has extra value in virtue of the fact that it indicates something good, namely more true beliefs.

Piller raises another objection to the conditional probability solution that misses the mark. He questions the evaluative significance of the fact that knowledge is likely to lead to more true belief. He characterises G&O's position with the following: "it is good that your belief is likely to be true because, if it is likely to be true, you will also get beliefs in the future, which are likely to be true" [146, p. 130]. But what is good about true belief in the first place? As Piller points out, the goodness of something cannot be explained by appealing to the fact that it leads to more of the same thing. He attempts to emphasise this with the following analogy:

'What is so great about having rabbits?' I ask my friend. 'Well', he answers 'if you have a couple of rabbits to start with, you will have even more rabbits in the future'. ... If all likely being true gives you is more of the same (other things likely to be true), we do not understand what good it is. [146, p. 130]

This is a bad objection and a bad analogy. G&O's conditional probability solution is grounded in the generally accepted conviction that truth is an epistemic good. They do not need to explicate this conviction as part of their argument.

Furthermore, given that truth is an epistemic good, we can play with the rabbit analogy and make it support the case that knowledge is better than mere true belief because it leads to more true belief. This would be even better if an argument were provided to show how the knowledge itself can play a part in determining a higher chance of having future true beliefs.

Suppose that rabbits are valued because they have the valuable property of cuteness. Therefore the more rabbits you have the more cuteness and the more of a good thing. Each

rabbit is as cute as another. Similarly, the true belief constituting the knowledge that p is just as true as the mere true belief that p . Now, take two female rabbits, rabbit A and rabbit B. Rabbit A is highly fertile and rabbit B's fertility is low. Although rabbit A and rabbit B are valued equally insofar as they are both rabbits and have the property of cuteness, there is a sense in which rabbit A is more valuable because she is more likely to progenerate more future rabbits than rabbit B. In epistemic terms, rabbit A is the knowledge and rabbit B is the mere true belief.

Of course, fertility not only indicates future rabbits, but plays a part in causing future rabbits. Finding ways in which the knowledge that p can play a part in causing future true beliefs in a way that the mere true belief that p cannot is an interesting task. I do have one suggestion that involves knowledge and beliefs in a stronger way than they are involved in the conditional probability solution. This is not intended to dismiss the conditional probability solution, but rather strengthen the case for a reliabilist account of the extra value of knowledge.

In order to explicate this idea (which also involves the four empirical regularities of non-uniqueness, cross-temporal access, learning and generality), take the knowledge that left₂ is the correct turn to take versus the corresponding mere true belief in the example that I gave a little earlier. After correctly turning left the second time, the driver must then turn right at the third turning point. Their arrival at the third turning point is at least in part due to the true belief that left₂. When this true belief constitutes knowledge, it can 'vouch for' its truth and the action that it leads to with the fact that any future beliefs required for similar actions will very likely also be true. You can trust these true beliefs because their 'pedigree' helps to ensure that future action, which they are partly responsible for, will also be successful. Mere true beliefs that are not based on information do not have this quality. Bob with the malfunctional GPS navigator ends up at the third turning point after acting upon the mere true beliefs that left₁ is correct and left₂ is correct but ends up failing at the third turn.

So an information-based true belief is more valuable when it fits in with a greater network of information and leads to more subsequent true beliefs based on information. In saying that a true belief has instrumental value because it results in successful action, it should also be judged on how well it consequently leads to other true beliefs and successful actions. True beliefs that are knowledge do better than mere true beliefs on this score.

To return to the coffee analogy, sometimes I will want a second coffee because the first one I had was so great. But if I have a bad coffee it ruins my day. If a great cup of coffee is to make me want another cup, then I would value it more if it is reliably produced and

thus will reliably lead me to another great cup of coffee rather than being unreliably produced and leading me to a bad coffee.

Recall the terms soundness and completeness for information systems from Section 5.2.2. It is clear that the malfunctioning GPS navigator is both unsound and incomplete whereas the functional GPS navigator is both sound and complete. Also recall from Section 5.2.2 the distinction between the information-bearing status of a particular signal in an information system and the reliability of the system as a whole. Clearly it is optimal for an information source to be maximally reliable. So not only is belief caused by an information-bearing signal better than mere true belief, but the more reliable the system (of which the signal is a part) the better.

Take the information system constituted by the collection of instruments on the dashboard of a car. Even if the in-built GPS navigation device were fully functioning it would be problematic if other instruments did not function. For example, suppose that you consult the functioning GPS navigator and it tells you to take a certain road with a speed limit of 50 km/h. Whilst on that road your malfunctioning speedometer misinforms you that your speed is 49 km/h, when it is in fact 60 km/h. As a result of this unreliability and lack of coordination amongst instruments, you receive a speeding fine. Thus whilst information-based belief is better than mere-true belief, information-based beliefs that are generated in a sound and complete information system are better than information-based beliefs that are generated in an unsound or incomplete system.

5.5 Conclusion

The informational epistemology investigated in this chapter has the conception of information developed throughout this thesis (particularly the previous chapter) playing a central role in founding an externalist epistemology. Objective semantic information is encapsulated by knowledge, which is defined as belief that is caused by and/or carries the right information. Whilst Dretske's work serves as a starting point for the development of an informational epistemology, certain examples test its limits and seriously raise the possibility that his account needs minor supplementation at the least. If this is not enough, then an impure informational epistemology involving non-informational conditions is an option.

The approach taken to determining relevant alternatives for information flow carries over to knowledge. If one adopts standard or attributor contextualism, then given some set of contextual relevant alternatives all knowledge statements will be evaluated against that set. The logic of such contextual knowledge is relatively straightforward and can be appropriately

represented with standard epistemic multi-modal logic. Dretske's picture of relevant alternatives is more complicated. It means the rejection of several properties, including closure, and calls for a rather unorthodox epistemic logic. An explication of this logic has been a central task of the last two chapters.

Finally, it was shown how an informational analysis of knowledge can explain and support the intuition that knowledge is better than mere true belief. Information-based true beliefs often go hand-in-hand with or lead to other information-based true beliefs in a way that mere true beliefs cannot. Thus the generation of an information-based true belief often results in a network with greater epistemic value.

Epilogue

I would like to finish with some brief summary thoughts and suggestions regarding further research. This discussion will pertain to both the chapters of this thesis and some general open questions.

This thesis can be divided into two main sections, each dealing with a topic I became interested in as a result of my initial research. The first half of this thesis (Chapters 2 and 3) was concerned with the development of a formal logical framework to represent propositional information and measure informativeness based on the notion of truthlikeness. The starting point for this work was Floridi's theory of strongly semantic information and his departure from traditional inverse probabilistic approaches to quantifying semantic information. With regards to the material in these two chapters, here are some further possible research tasks:

- The investigation into truthlikeness information measures covered quite a bit of territory. Still, there is always room to explore ways in which the idea of a truthlikeness information measure can be realised and fleshed out. As an example, one approach to further look into is that behind the value aggregate method introduced in Section 2.5.
- The work on agent-relative informativeness overlaps and can find application in the areas of formal epistemology and belief revision. Looking at ways to supplement belief revision for the aim of truthlikeness is one interesting project to look further into. Another concerns the applicability of tools and results from chapters 2 and 3 to the field of database merging/belief fusion.
- In general, combining information/truthlikeness measures with one or more of belief revision, decision theory or confirmation theory can be a fruitful way to come up with new research topics and questions.

The framework employed was based on propositional logic and mainly dealt with quantitative factors. This account could serve as a foundation for accounts that represent informa-

tion in a richer way beyond propositional logic representation. Also, the measurement methods considered dealt with a purely quantitative notion of informativeness/information yield. Beyond this there is the subject of information quality to consider and the measurement of qualitative factors. The incorporation of qualitative factors relates to the consideration of pragmatic aspects concerning the relevance and effectiveness of information to an agent.

The second half of this thesis (Chapters 4 and 5) built upon the work of Dretske in developing a framework for (semantic) environmental information and knowledge. The investigation conducted over these two chapters resulted in some important points and clarifications regarding the logic of information flow and knowledge. With regards to the material in these chapters, aside from perennial philosophical debates such as the issue of information/knowledge closure, here are some further possible research tasks:

- If information is understood as information_S , then this adds an extra factor to the quantification of semantic information. If quantification is to just basically distinguish between true content and false content, then perhaps there could be a variable in measuring true content that factored in the reliability of the true content. On the other hand, if there is to be a sharp distinction between information (information_S), true semantic content (information_W) and misinformation, then measures would have to take into account the fact that information and misinformation are not direct opposites. Something not being information does not imply that it is misinformation and accordingly it would be possible for a true statement that is not information to have neither a positive information measure nor a positive misinformation measure.
- With regards to information and epistemology, bringing specific accounts of information into epistemological topics beyond the analysis of propositional knowledge is one possible undertaking to look into. Social epistemology and the epistemology of science are two areas that come to mind.

Another endeavour concerns further developing a response to the value problem for knowledge in terms of an informational epistemology.

This thesis has both addressed old questions and raised some new ones. Further research avenues left by or related to the investigations conducted and ways in which the results of this thesis can be applied to and combined with results from other research fields remain to be explored.

Appendices

Appendix A

Quantifying Semantic Information

Theorem A.0.1. $\text{misinfo}_1() = \text{misinfo}_2()$

Proof. • $w =_{df}$ number of states contained in a statement

• $t =_{df}$ number of true atomic states in the collection of states contained in a statement

• $f =_{df}$ number of false atomic states in the collection of states contained in a statement

• $x =_{df}$ the weight assigned to each basic state

• $n =_{df}$ the number of basic states per state

For example, using the weather framework above, given the statement $h \wedge \neg r$ $w = 2$ (w_1, w_2), $t = 5$, $f = 1$, $x = \frac{1}{3}$ and $n = 3$. Given these definitions, it follows that:

$$\begin{aligned} t + f &= w \times n \\ \therefore t &= (w \times n) - f \quad (\text{Eq. 1}) \end{aligned}$$

$$n \times x = 1 \quad (\text{Eq. 2})$$

The equivalence of these two approaches translates to the following formal equation:

$$\begin{aligned}
\frac{t \times x}{w} &= 1 - \frac{f \times x}{w} \\
\frac{t \times x}{w} + \frac{f \times x}{w} &= 1 \\
\frac{((w \times n) - f) \times x}{w} + \frac{f \times x}{w} &= 1 \quad (\text{Using Eq. 1}) \\
\frac{(((w \times n) - f) \times x) + f \times x}{w} &= 1 \\
\frac{w \times nx - fx + fx}{w} &= 1 \\
\frac{w \times nx}{w} &= 1 \\
\frac{w}{w} &= 1 \quad (\text{Using Eq. 2}) \\
1 &= 1
\end{aligned}$$

□

Theorem A.0.2. $\frac{\sum_{i=1}^n \text{Pr}^*(p_i, A)}{n}$ is equivalent to the standard method for the Tichy/Oddie measure.

Proof. The states in W_A combined give t true atomic element valuations relative to the actual state. The standard method can be defined as

$$\text{info}_{TO}(A) = \frac{t}{n|W_A|}$$

Since n is a common factor in both the standard definition and this new definition, it will suffice to show that $\sum_{i=1}^n \text{Pr}^*(p_i, A) = \frac{t}{|W_A|}$:

- $\sum_{i=1}^n \text{Pr}^*(p_i, A)$ reduces to $\frac{\text{Pr}(\pm p_1 \wedge A)}{\text{Pr}(A)} + \dots + \frac{\text{Pr}(\pm p_n \wedge A)}{\text{Pr}(A)}$, where $\text{Pr}(A) = \frac{|W_A|}{2^n}$
- Therefore $\sum_{i=1}^n \text{Pr}^*(p_i, A)$ reduces to $\frac{2^n \text{Pr}(\pm p_1 \wedge A) + \dots + 2^n \text{Pr}(\pm p_n \wedge A)}{|W_A|}$
- Therefore we need to show that $t = 2^n \text{Pr}(\pm p_1 \wedge A) + \dots + 2^n \text{Pr}(\pm p_n \wedge A)$:
 - For each p_i , let t_i stand for the number of true atomic element valuations for p_i in W_A . So $t = t_1 + t_2 + \dots + t_n$.
 - Therefore we need to show that $t_1 + \dots + t_n = 2^n \text{Pr}(\pm p_1 \wedge A) + \dots + 2^n \text{Pr}(\pm p_n \wedge A)$, which is the case since each t_i is equal to $2^n \text{Pr}(p_i \wedge A)$: $2^n \text{Pr}(p_i \wedge A) = 2^n \times \frac{t_i}{2^n} = t_i$

□

A.1 Adequacy Condition Proofs for the Value Aggregate Method

Adequacy condition results for the value aggregate method.

- Let W_A stand for the set of states satisfying statement A .
- w_T denotes the actual state.
- $S_* = T$.
- Consult Section 2.5.1 for further terminology.

Theorem A.1.1. (M1) $0 \leq \text{Tr}(A, T) \leq 1$.

Proof. $X = \text{lineup}(\text{arraystate}(W_A))$. The lowest possible value for $\text{sum}(X)$ is 0, when each item of X is the lowest valued state ($2^n \times 0 = 0$). The highest possible value for $\text{sum}(X)$ is 1, when each item of X is the highest valued state ($2^n \times \frac{n}{n \times 2^n} = 1$). $\square$

Theorem A.1.2. (M2) $\text{Tr}(A, T) = 1$ iff $A = T$.

Proof. $X = \text{lineup}(\text{arraystate}(W_A))$. $\text{sum}(X) = 1$ iff each item of X is the highest valued state (*i.e.*, w_T). Each item of X is the highest valued state iff the statement being measured is a state description of the actual state. $\square$

Theorem A.1.3. (M3) All true statements do not have the same degree of truthlikeness, all false statements do not have the same degree of truthlikeness.

Proof. Evident with the results from Table 2.18. $\square$

Theorem A.1.4. (M4) Among true statements, truthlikeness covaries with logical strength:

- (a) If A and B are true statements and $A \vdash B$, then $\text{Tr}(B, T) \leq \text{Tr}(A, T)$.
- (b) If A and B are true statements and $A \vdash B$ and $B \not\vdash A$, then $\text{Tr}(B, T) < \text{Tr}(A, T)$.

Proof. We will show that (b) holds, since this entails (a).

- $X_1^A = \text{arraystates}(W_A)$
- $X_1^B = \text{arraystates}(W_B)$

- $X_2^A = \text{lineup}(X_1^A)$
- $X_2^B = \text{lineup}(X_1^B)$

Now

- where n is the number of propositional variables, the values of the states will be drawn from the following: $\frac{0}{(n \times 2^n)}, \frac{1}{(n \times 2^n)}, \frac{2}{(n \times 2^n)} \cdots \frac{n}{(n \times 2^n)}$.
- Since A and B are true, $w_T \in W_A$ and $w_T \in W_B$.
- since $A \vdash B$ and $B \not\vdash A$, $W_A \subset W_B$

So X_2^A is going to contain at least 1 more instance of w_T than X_2^B . Suppose that X_2^B does just have 1 less instance of w_T and that this is replaced by an instance of the second highest valued state (which will have a value of $\frac{n-1}{n \times 2^n}$). This is the upper limit, the closest X_2^B will come to having a sum value greater than X_2^A , so it will suffice to show that in this case $\text{sum}(X_2^B) < \text{sum}(X_2^A)$.

The term $X[i]$ denotes position i of array X . Let m be the position at which $X_2^A[m] \neq X_2^B[m]$. So $1 \leq m \leq 2^n - 1$ and for each $i \in \{x \mid 1 \leq x \leq m-1\}$, $X_2^A[i] = X_2^B[i]$, so the value of $\text{sum}()$ is equal for both arrays up to point m .

After and including point m , whilst the remaining elements of X_2^A sum up to $(2^n - m) \times \frac{n}{(n \times 2^n)}$, the remaining elements of X_2^B sum up to $\frac{n-1}{(n \times 2^n)} + ((2^n - m) - 1) \times \frac{n}{(n \times 2^n)}$.

So the final thing to show is that $(2^n - m) \times \frac{n}{(n \times 2^n)} > \frac{n-1}{(n \times 2^n)} + ((2^n - m) - 1) \times \frac{n}{(n \times 2^n)}$

$$\begin{aligned} \frac{n(2^n - m)}{(n \times 2^n)} &> \frac{(n-1) + n(2^n - m - 1)}{(n \times 2^n)} \\ n2^n - nm &> n2^n - nm - 1 \end{aligned}$$

□

Theorem A.1.5. (M5) Among false statements, truthlikeness does not covary with logical strength; there are false statements A and B such that $A \vdash B$ but $\text{Tr}(A, T) < \text{Tr}(B, T)$.

Proof. Evident with the results from table 2.18.

□

Theorem A.1.6. (M7) If A is a false statement, then $\text{Tr}(T \vee A, T) > \text{Tr}(A, T)$

Proof. Since A is false, it follows that $w_T \notin W_A$. The set of states corresponding to $T \vee A$ is $W_A \cup \{w_T\}$. Let

- $X_2 = \text{lineup}(\text{arraystates}((W_A)))$
- $X_{2'} = \text{lineup}(\text{arraystates}((W_A \cup \{w_T\})))$

So we need to show that $\text{sum}(X_{2'}) > \text{sum}(X_2)$.

Say the highest valued element of W_A is w_a and that w_a takes up the last n positions in X_2 . The addition of w_T results in it replacing w_a for the last $n - 1$ positions. Since $\text{val}(w_T) > \text{val}(w_a)$, $\text{sum}(X_{2'}) > \text{sum}(X_2)$. $\square$

Theorem A.1.7. (M8) Assume $j \notin \mathbf{I}_A$. Then $\text{Tr}(A \vee S_j, T) > \text{Tr}(A, T)$ iff $\Delta_{*j} < \Delta_{\min}(A, T)$

- $X_2 = \text{lineup}(\text{arraystates}((W_A)))$
- $X_{2'} = \text{lineup}(\text{arraystates}((W_A \cup \{w_j\})))$

Proof. First we show that

$$\Delta_{*j} < \Delta_{\min}(A, T) \Rightarrow \text{Tr}(A \vee S_j, T) > \text{Tr}(A, T)$$

If the antecedent here holds, then $(\forall w)(w \in W_A \supset \text{val}(w_j) > \text{val}(w))$. Therefore the sum of $X_{2'}$ will have a greater value than the sum of X_2 .

Second we show that

$$\text{Tr}(A \vee S_j, T) > \text{Tr}(A, T) \Rightarrow \Delta_{*j} < \Delta_{\min}(A, T)$$

via the contraposition

$$\Delta_{*j} > \Delta_{\min}(A, T) \Rightarrow \text{Tr}(A \vee S_j, T) < \text{Tr}(A, T)$$

If the antecedent here holds, then $(\exists w)(w \in W_A \wedge \text{val}(w_j) < \text{val}(w))$. Therefore the sum of $X_{2'}$ will have a lower value than the sum of X_2 , since one instance of $w \in W_A$ is taken off from the $\text{sum}()$ calculations and replaced by the lower valued w_j , resulting overall in a lower value.

□

Theorem A.1.8. (M9) Let $\Delta_{*j} < \Delta_{*i}$. Then $\text{Tr}(S_j \vee S_i, T)$ decreases when Δ_{*i} increases

Proof. There are two states, w_j which corresponds to S_j and w_i which corresponds to S_i , with the value of w_j being greater than the value of w_i . Let $X_2 = \text{lineup}(\text{arraystates}(\{w_j, w_i\}))$, so $\text{Tr}(S_j \vee S_i) = \text{sum}(X_2)$. Let $X_{2'} = \text{lineup}(\text{arraystates}(\{w_j, w_{w'}\}))$, where $w_{w'}$ replaces w_i when Δ_{*i} increases. Since w_i is replaced by a lower $w_{w'}$, then $\text{sum}(X_{2'}) < \text{sum}(X_2)$. □

Theorem A.1.9. (M10) Some false statements may be more truthlike than some true statements.

Proof. Evident with the results from table 2.18. □

Theorem A.1.10. (M12) If $\Delta_{*j} < \Delta_{*i} < \Delta_{*k}$, then $\text{Tr}(S_j \vee S_i \vee S_k, T)$ increases when Δ_{*i} decreases.

Proof. This is straightforward; any increase for $\text{Tr}(S_i)$ means an increase in the value of w_i , which means an overall total increase in $\text{sum}()$. □

Theorem A.1.11. (M13) $\text{Tr}(A, T)$ is minimal, if A consists of the Δ -complements of T .

Proof. If A is the Δ -complement of T then it is the state such that each atom is false. Therefore $\text{Tr}(A, T)$ is a minimal 0. □

A.2 Translation Invariance

The issue of translation invariance affects approaches to truthlikeness such as the ones covered here (what can be labelled syntactic approaches), which are not translation invariant. Such approaches proceed

in terms of syntactic surrogates for their semantic correlates – sentences for propositions, predicates for properties, distributive normal forms for partitions of the logical space, and the like. The question naturally arises, then, whether we obtain the same measures if all the syntactic items are translated into an essentially equivalent language – one capable of expressing the same propositions and properties with a different set of primitive predicates. [139]

David Miller showed this not to be the case. Suppose we add the following two new weather conditions, minnesotan (m) and arizonan (a) to the weather example, such that:

- minnesotan $=_{df}$ hot if and only if rainy ($h \equiv r$)
- arizonan $=_{df}$ hot if and only if windy ($h \equiv w$)

Now it seems that the language $h-m-a$ -ese can describe the same set of weather states as those based on the language $h-r-w$ -ese:

	$h-r-w$ -ese	$h-m-a$ -ese
T	$h \wedge r \wedge w$	$h \wedge m \wedge a$
A	$\neg h \wedge r \wedge w$	$\neg h \wedge \neg m \wedge \neg a$
B	$\neg h \wedge \neg r \wedge w$	$\neg h \wedge m \wedge \neg a$
C	$\neg h \wedge \neg r \wedge \neg w$	$\neg h \wedge m \wedge a$

But although $\neg h \wedge r \wedge w$ and $\neg h \wedge \neg m \wedge \neg a$ are logically equivalent

If T is the truth about the weather then theory A , in $h-r-w$ -ese, seems to make just one error concerning the original weather states, while B makes two and C makes three. However, if we express these two theories in $h-m-a$ -ese however, then this is reversed: A appears to make three errors and B still makes two and C makes only one error. But that means the account makes truthlikeness, unlike truth, radically language-relative. [139]

I would say that translation invariance is not an issue for information as it is for truthlikeness. Firstly, there is the simple case to be made that if information is placed at the language level, then language-relative information measurements could legitimately be used in certain accounts.

Secondly, a response aligned with the propositional analysis of information established in Section 1.3.2 of Chapter 1 is another way to deal with the issue.

For the truthlikeness enterprise, which is concerned with things like the formulation of statements in a scientific theory, the measurements of truthlikeness are being made on the actual statements themselves, on the language of a theory. But if measurements of information are ultimately being made on the propositions behind the statements, we may

say that whilst $\neg h \wedge r \wedge w$ and $\neg h \wedge \neg m \wedge \neg a$ express different statements, they express the same information. Whilst measurements on statements are variant, measurements on information are invariant, explicated as follows.

Take the following propositions: $X, Y, Z, X \equiv Y, X \equiv Z$. These propositions can be mapped to the statements in question:

- $h \Rightarrow X$
- $r \Rightarrow Y$
- $w \Rightarrow Z$
- $m \Rightarrow X \equiv Y$
- $a \Rightarrow X \equiv Z$

Now, the truth is represented by the proposition $X \wedge Y \wedge Z$ or equivalently by $X \wedge (X \equiv Y) \wedge (X \equiv Z)$. A measurement of truthlikeness on the statement $\neg h \wedge r \wedge w$ is ultimately a measurement on the proposition $\neg X \wedge Y \wedge Z$. A measurement of truthlikeness on the statement $\neg h \wedge \neg m \wedge \neg a$ is ultimately a measurement on the proposition $\neg X \wedge \neg(X \equiv Y) \wedge \neg(X \equiv Z)$. Since $\neg X \wedge Y \wedge Z$ and $\neg X \wedge \neg(X \equiv Y) \wedge \neg(X \equiv Z)$ are the same proposition, the information-as-truthlikeness measurement is going to be the same.

The above mapping is not unique though. An alternative mapping is:

- $h \Rightarrow X$
- $r \Rightarrow X \equiv Y$
- $w \Rightarrow X \equiv Z$
- $m \Rightarrow Y$
- $a \Rightarrow Z$

The task of determining a/the correct mapping is a scientific/epistemological one, which does not impinge upon the ontological nature of information. Quite simply, under this propositional analysis if two statements map to the same proposition, then they yield the same amount of information, regardless of the actual nature of this mapping. $\neg h \wedge r \wedge w$ and $\neg h \wedge \neg m \wedge \neg a$ have the same information measure because they map to the same proposition, whether that proposition be represented by $\neg X \wedge Y \wedge Z$ or $\neg X \wedge (X \equiv Y) \wedge \neg(X \equiv Z)$.

A.3 Formula-Based Approaches

The main approaches to truthlikeness information measurement looked at in Chapter 2 deal with the models of a formula. Another way to approach a method for truthlikeness information measurement is to operate on the formulas themselves. Following is a method that will serve as an illustrative sketch of one way to go about implementing this idea.

To begin with, statements are first converted to conjunctive normal form. Although this means that logically equivalent statements are treated the same, they are still being dealt with directly and there is no analysis whatsoever of the models that they correspond to.

A logical statement is in conjunctive normal form (CNF) when it consists of a conjunction of disjunctive clauses, with each disjunct in each conjunction being a literal (either an atom or a negated atom). $A \wedge B \wedge \neg C$ and $(A \vee \neg B) \wedge (A \vee C)$ are two examples of statements in CNF. Also, here the normalised statements are *fully* converted to CNF, meaning that all redundancies are eliminated. Here is the procedure:

- use equivalences to remove $\leftrightarrow$ and $\rightarrow$.
- use De Morgan's laws to push negation signs immediately before atomic statements.
- eliminate double negations.
- use the distributive law $A \vee (B \wedge C) \equiv (A \vee B) \wedge (A \vee C)$ to effect the conversion to CNF.
- if a conjunct contains both A and $\neg A$ (*i.e.*, $A \vee \neg A$), remove it.
- use absorption law $A \wedge (A \vee B) \equiv A$ so that for any literal that occurs as a sole disjunct (*i.e.*, occurs in a disjunctive clause of which it is the only member), remove any others disjunctive clauses that contain that literal.
- use rule of unsatisfiability $A \vee \perp \equiv A$ so that for any literal A that occurs as a sole disjunct (*i.e.*, occurs in a disjunctive clause of which it is the only member), remove its negation from any disjunction it occurs in. So $A \wedge (B \vee \neg A)$ becomes $A \wedge B$.

Given a normalised statement, how can its information yield be measured? Let us start with the simplest type of statement in CNF, one consisting of a conjunction of literals. Each literal has a value of magnitude 1. At this stage there are two ways to go about this. The first method (Method 1) involves just positively incrementing the total information yield of the

statement as each true literal is encountered. The second method (Method 2) involves also decrementing the total information yield of the statement as each false literal is encountered.

Take the statement $h \wedge r \wedge \neg w$. It contains 2 true literals and 1 false literal. Going by Method 1, the total information yield of this statement is 2. Going by Method 2, where the misinformation instances of a statement count against its total information yield, the total information yield of this statement is 1.

This is all very simple and gives the right results. $h \wedge r \wedge w$ has a maximal information yield of 3. $h \wedge r \wedge \neg w$ and $h \wedge r$ have the same information yield of 2 going by Method 1, but going by method 2 the former is ‘punished’ for its assertion of misinformation and goes down to 1.

The introduction of proper disjunctions makes things more complicated. How can the calculation method reflect the fact that $h \vee r$ yields less information than other statements such as h and $h \wedge r$. The more disjuncts a disjunction contains, the less information it yields; information yield is inversely related to the number of disjuncts. Hence, where n is the number of disjuncts in a disjunction, a simple yet potentially suitable measure of the maximum information yield of the disjunction is $\frac{1}{n}$. With this way, each literal in the disjunction has a value of magnitude $\frac{1}{n^2}$.

To illustrate all of this, take the statement $h \vee r \vee w$. The total information yield of this statement is $3 \times \frac{1}{9} = \frac{1}{3}$. Given a disjunction with false literals, the results depend of whether Method 1 or Method 2 is adopted. For the statement $h \vee r \vee \neg w$, Method 1 gives a result of $\frac{1}{3} + \frac{1}{3} = \frac{2}{3}$ and Method 2 gives a result of $\frac{1}{3} + \frac{1}{3} - \frac{1}{3} = \frac{1}{3}$.

Table A.1 lists results using Method 1 and results using Method 2. Clearly these methods are not as refined as the model-based approaches we have looked at, with the variation in assignments of measure values being significantly cruder. Method 1 is clearly just about aggregating truth. Interestingly, the results for Method 2 can be ranked in the same order as the results for the Tichy/Oddie approach in Table 2.11.

Method 1				Method 2			
#	Statement (A)	T/F	info(A)	#	Statement (A)	T/F	info(A)
1	$h \wedge r \wedge w$	T	3	1	$h \wedge r \wedge w$	T	3
2	$h \wedge r \wedge \neg w$	F	2	2	$h \wedge r$	T	2
3	$h \wedge r$	T	2	3	$h \wedge (r \vee w)$	T	1.5
4	$h \wedge (r \vee w)$	T	1.5	4	$h \wedge (\neg r \vee w)$	T	1
5	$h \wedge (\neg r \vee w)$	T	1.25	5	$(h \wedge r) \vee w$	T	1
6	$h \wedge \neg r \wedge \neg w$	F	1	6	h	T	1
7	$h \wedge \neg r$	F	1	7	$h \wedge r \wedge \neg w$	F	1
8	$h \wedge \neg h$	F	1	8	$h \vee r$	T	0.5
9	h	T	1	9	$(h \wedge \neg r) \vee w$	T	0.5
10	$(h \wedge r) \vee w$	T	1	10	$h \vee r \vee w$	T	0.33
11	$(h \wedge \neg r) \vee w$	T	0.75	11	$h \vee r \vee \neg w$	T	0.11
12	$h \vee r$	T	0.5	12	$h \vee \neg r$	T	0
13	$h \vee r \vee w$	T	0.33	13	$h \vee \neg h$	T	0
14	$(h \vee \neg w) \wedge \neg r$	F	0.25	14	$h \wedge \neg r$	F	0
15	$h \vee \neg r$	T	0.25	15	$h \wedge \neg h$	F	0
16	$h \vee \neg h$	T	0.25	16	$h \vee \neg r \vee \neg w$	T	-0.11
17	$h \vee r \vee \neg w$	T	0.22	17	$\neg h \vee \neg r \vee \neg w$	F	-0.33
18	$h \vee \neg r \vee \neg w$	T	0.11	18	$h \wedge \neg r \wedge \neg w$	F	-1
19	$\neg h \wedge \neg r \wedge \neg w$	F	0	19	$(h \vee \neg w) \wedge \neg r$	F	-1
20	$\neg h$	F	0	20	$\neg h$	F	-1
21	$\neg h \wedge \neg r$	F	0	21	$\neg h \wedge \neg r$	F	-2
22	$\neg h \vee \neg r \vee \neg w$	F	0	22	$\neg h \wedge \neg r \wedge \neg w$	F	-3

Table A.1: Information yield using formula-based approach.

Appendix B

Agent-Relative Informativeness

B.1 Combining Information Measurement and Belief Revision

B.1.1 True Content and False Input

Theorem B.1.1. For $\text{info}_{VA}()$ and $\text{info}_{TO}()$, if D is true and A is false then either of (1) $\text{info}(D + A) < \text{info}(D)$ or (2) $\text{info}(D + A) > \text{info}(D)$ are possible.

Proof. (1) is obvious. For (2) take the example where $D = w_1 \vee w_2 \vee w_4 \vee w_6 \vee w_7 \vee w_8$ and $A = w_2$, so that $D + A = w_2$. $\square$

Theorem B.1.2. For $\text{info}_{VA}()$ and $\text{info}_{TO}()$, if D is true and A is false then either of (1) $\text{info}(D * A) < \text{info}(D)$ or (2) $\text{info}(D * A) > \text{info}(D)$ are possible.

Proof. (1) is obvious. For (2) let $D = \neg h \vee \neg r \vee w$ and $A = h \wedge r \wedge \neg w$ so that $D * A = h \wedge r \wedge \neg w$. $\square$

Note: When D is true and A is false then $D \div A = D$.

B.1.2 False Content and False Input

Theorem B.1.3. For $\text{info}_{VA}()$ and $\text{info}_{TO}()$, if D is false and A is false then either of (1) $\text{info}(D + A) < \text{info}(D)$ or (2) $\text{info}(D + A) > \text{info}(D)$ are possible.

Proof. (1) is obvious. For (2) let $D = (h \wedge r \wedge \neg w) \vee (\neg h \wedge \neg r \wedge \neg w)$ and $A = h \wedge r \wedge \neg w$, so that $D + A = h \wedge r \wedge \neg w$. $\square$

Theorem B.1.4. For $\text{info}_{VA}()$ and $\text{info}_{TO}()$, if D is false and A is false then either of (1) $\text{info}(D * A) < \text{info}(D)$ or (2) $\text{info}(D * A) > \text{info}(D)$ are possible.

Proof. (1) is obvious. For (2) take the example where $D = \neg h \wedge \neg r \wedge \neg w$ and $A = h \wedge r \wedge \neg w$. $\square$

Theorem B.1.5. For $\text{info}_{VA}()$ and $\text{info}_{TO}()$, if D is false and A is false then either of (1) $\text{info}(D \div A) > \text{info}(D)$ or $\text{info}(D \div A) < \text{info}(D)$ are possible.

Proof. (1) is obvious. For $\text{info}(D \div A) < \text{info}(D)$ take the following example:

- $D = h \wedge \neg r \wedge \neg w = w_4$
- $A = \neg w_1 \wedge \neg w_8$
- $D \vdash A$
- $D \div A = D \vee D * (w_1 \vee w_8) = w_4 \vee w_8$
- $\text{info}(w_4 \vee w_8) < \text{info}(w_4)$

$\square$

B.2 True Inputs Guaranteed to Increase Information Yield

Theorem B.2.1. Take a logical space with n atoms and a true atom p . We have:

1. $\text{info}_{TO}(D + p) - \text{info}_{TO}(D) = \frac{1}{2^n}$
2. $\text{info}_{TO}(D * p) - \text{info}_{TO}(D) = \frac{1}{n}$
3. $\text{info}_{TO}(D + \neg p) - \text{info}_{TO}(D) = -\frac{1}{2^n}$
4. $\text{info}_{TO}(D * \neg p) - \text{info}_{TO}(D) = -\frac{1}{n}$

Proof. Since D is a c-statement with m conjuncts, it is satisfied in 2^{n-m} states. $\text{info}_{TO}(D) = 1 - (\frac{d}{2^{n-m}} \times \frac{1}{n}) = 1 - \frac{d}{n2^{n-m}}$, where d is the number of differences relative to the actual state.

1. The expansion of D with a true atom p will halve the set of satisfying models/states.

The new value for $\text{info}_{TO}()$ will be $\text{info}_{TO}(D + p) = 1 - (\frac{2d_2}{n2^{n-m}})$, where $d_2 = (d - \frac{2^{n-m}}{2}) \times \frac{1}{2} = \frac{2d-2^{n-m}}{4}$. So $\text{info}_{TO}(D + p) = 1 - (\frac{2d-2^{n-m}}{2n2^{n-m}})$.

Now to show that $(1 - (\frac{2d-2^{n-m}}{2n2^{n-m}})) - (1 - (\frac{d}{n2^{n-m}})) = \frac{1}{2n}$:

$$\begin{aligned} (1 - (\frac{2d-2^{n-m}}{2n2^{n-m}})) - (1 - (\frac{d}{n2^{n-m}})) &= \frac{1}{2n} \\ (\frac{d}{n2^{n-m}}) - (\frac{2d-2^{n-m}}{2n2^{n-m}}) &= \frac{1}{2n} \\ \frac{2d-(2d-2^{n-m})}{2n2^{n-m}} &= \frac{1}{2n} \\ \frac{2^{n-m}}{2n2^{n-m}} &= \frac{1}{2n} \\ \frac{1}{2n} &= \frac{1}{2n} \end{aligned}$$

2. The revision of D with a true atom p will replace the set of satisfying models/states with the corresponding set of models in which all that is changed is that the valuation for p is now true instead of false. The new value for $\text{info}_{TO}()$ will be $\text{info}_{TO}(D * p) = 1 - (\frac{d_2}{2^{n-m}} \times \frac{1}{n})$, where $d_2 = d - 2^{n-m}$. So $\text{info}_{TO}(D * p) = 1 - (\frac{d-2^{n-m}}{n2^{n-m}})$

Now to show that $(1 - (\frac{d-2^{n-m}}{n2^{n-m}})) - (1 - (\frac{d}{n2^{n-m}})) = \frac{1}{n}$:

$$\begin{aligned} (1 - (\frac{d-2^{n-m}}{n2^{n-m}})) - (1 - (\frac{d}{n2^{n-m}})) &= \frac{1}{n} \\ (\frac{d}{n2^{n-m}}) - (\frac{d-2^{n-m}}{n2^{n-m}}) &= \frac{1}{n} \\ \frac{2^{n-m}}{n2^{n-m}} &= \frac{1}{n} \\ \frac{1}{n} &= \frac{1}{n} \end{aligned}$$

Similar converse proofs apply for 3 and 4.

□

Theorem B.2.2. With D being any statement and A being a c-statement that is completely false, $\text{info}(D + A) < \text{info}(D)$ does not universally hold.

Proof. Since this failed for $\text{info}_{VA}()$ in Theorem 3.2.11, we will just look at $\text{info}_{TO}()$. Here is one example:

- $D = w_2 \vee w_7$
- $A = \neg w$
- $D + A = w_2$
- $\text{info}_{TO}(w_2) = 0.6667 > \text{info}_{TO}(w_2 \vee w_7) = 0.5$

□

Theorem B.2.3. With D being any statement and A being a c-statement that is completely false, $\text{info}(D * A) < \text{info}(D)$ does not universally hold.

Proof. Since this failed for $\text{info}_{VA}()$ in Theorem 3.2.11, we will just look at $\text{info}_{TO}()$. Here is one example:

Take a 6-proposition space with true state description $p_1 \wedge p_2 \wedge p_3 \wedge p_4 \wedge p_5 \wedge p_6$. We have:

- $D = p_1 \wedge ((p_2 \wedge \neg p_3 \wedge \neg p_4 \wedge \neg p_5 \wedge \neg p_6) \vee (\neg p_2 \wedge p_3 \wedge p_4 \wedge p_5 \wedge p_6))$.
- $A = \neg p_1 \wedge \neg p_2$
- $D * A = \neg p_1 \wedge \neg p_2 \wedge p_3 \wedge p_4 \wedge p_5 \wedge p_6$
- $\text{info}_{TO}(D * A) = 0.6667 > \text{info}_{TO}(D) = 0.583$

□

Theorem B.2.4. With D being any statement and A being a false c-statement

1. $\text{info}_{TO}(D \div A) > \text{info}_{TO}(D)$
2. $\text{info}_{VA}(D \div A) > \text{info}_{VA}(D)$ does not universally hold

Proof. For $\text{info}_{TO}()$ this result is essentially the converse of the corresponding result in Theorem 3.2.14.

For $\text{info}_{VA}()$ we have the following counterexample. Take a 10-proposition space with true state description $p_1 \wedge p_2 \wedge p_3 \wedge p_4 \wedge p_5 \wedge p_6 \wedge p_7 \wedge p_8 \wedge p_9 \wedge p_{10}$. We have:

- $D = \neg p_1 \wedge \neg p_2$
- $A = \neg p_2$
- $\text{info}_{VA}(\neg p_1 \wedge \neg p_2) = 0.7 > \text{info}_{VA}(\neg p_1) = 0.675$

□

State	<i>h</i>	<i>r</i>	<i>w</i>	State	<i>h</i>	<i>r</i>	<i>w</i>	State	<i>h</i>	<i>r</i>	<i>w</i>
w ₁	T	T	T	w ₁₀	B	T	T	w ₁₉	F	T	T
w ₂	T	T	B	w ₁₁	B	T	B	w ₂₀	F	T	B
w ₃	T	T	F	w ₁₂	B	T	F	w ₂₁	F	T	F
w ₄	T	B	T	w ₁₃	B	B	T	w ₂₂	F	B	T
w ₅	T	B	B	w ₁₄	B	B	B	w ₂₃	F	B	B
w ₆	T	B	F	w ₁₅	B	B	F	w ₂₄	F	B	F
w ₇	T	F	T	w ₁₆	B	F	T	w ₂₅	F	F	T
w ₈	T	F	B	w ₁₇	B	F	B	w ₂₆	F	F	B
w ₉	T	F	F	w ₁₈	B	F	F	w ₂₇	F	F	F

Table B.1: *LP* Truth Table for 3-Proposition Logical Space

B.3 Paraconsistent Approaches

Theorem B.3.1. Like its classical counterpart, this distance function satisfies the metric conditions:

1. $\Delta(w_i, w_j) \geq 0$
2. $\Delta(w_i, w_j) = 0$ iff $w_i = w_j$
3. $\Delta(w_i, w_j) = \Delta(w_j, w_i)$
4. $\Delta(w_i, w_k) \leq \Delta(w_i, w_j) + \Delta(w_j, w_k)$

Proof. 1 - 3 are straightforward. For 4, we adopt and modify a technique from [158, p. 27]

We start off by representing each state/model as the set of literals it satisfies. For example, take w to be a state in which atom p_1 is assigned a value of T, atom p_2 is assigned a value of F and atom p_3 is assigned a value of B. Then its representation would be $\{p_1, \neg p_2, p_3, \neg p_3\}$.

Let w_i and w_j be two models with such corresponding sets w_i and w_j . Denote the operation $(w_i - w_j) \cup (w_j - w_i)$ with $D(w_i, w_j)$. An appropriate way to measure the distance between w_i and w_j is by using the value of $|D(w_i, w_j)|$. We have: $\Delta(w_i, w_j) = |D(w_i, w_j)|/2$.

Where S is some set of literals we use the following shortcut notation

- $p_i \in_* S =_{df} p_i \in S \wedge \neg p_i \notin S$
- $\neg p_i \in_* S =_{df} \neg p_i \in S \wedge p_i \notin S$
- $(p_i, \neg p_i) \in_* S =_{df} p_i \in S \wedge \neg p_i \in S$

- $(p_i, \neg p_i) \notin_* S =_{df} p_i \notin S \wedge \neg p_i \notin S$

With these details established, here are the points for a proof of 4:

- Let w_i, w_j and W_k stand for the sets of literals that correspond to the states/models w_i, w_j, w_k respectively.
- Define $X_{ik} = D(w_i, W_k)$, $X_{ij} = D(w_i, w_j)$ and $X_{jk} = D(w_j, W_k)$.
- We show that $\Delta(w_i, w_k) \leq \Delta(w_i, w_j) + \Delta(w_j, w_k)$ by showing that $|X_{ik}| \leq |X_{ij}| + |X_{jk}|$ by proving that $X_{ik} \subseteq X_{ij} \cup X_{jk}$. It suffices to demonstrate that in the following two cases the contribution p_i makes to $|X_{ik}|$ never exceeds the contribution it makes to $|X_{ij}| + |X_{jk}|$. These demonstrations transfer to other permutations:
 - If $p_i \in_* w_i$ and $\neg p_i \in_* W_k$ then $(p_i, \neg p_i) \in_* X_{ik}$. In this case, the contribution p_i makes to $|X_{ik}|$ is 2. In each of the following three possible values for w_j , the contribution p_i makes to $|X_{ij}| + |X_{jk}|$ is also 2:
 1. $p_i \in_* w_j$ - if this is the case then $(p_i, \neg p_i) \in_* X_{jk}$ and $(p_i, \neg p_i) \notin_* X_{ij}$
 2. $\neg p_i \in_* w_j$ - if this is the case then $(p_i, \neg p_i) \in_* X_{ij}$ and $(p_i, \neg p_i) \notin_* X_{jk}$
 3. $(p_i, \neg p_i) \in_* w_j$ - if this is the case then $\neg p_i \in_* X_{ij}$ and $p_i \in_* X_{jk}$
 - If $(p_i, \neg p_i) \in_* w_i$ and $\neg p_i \in_* W_k$ then $p_i \in X_{ik}$. In each of the following three possible values for w_j , the contribution p_i makes to $|X_{ij}| + |X_{jk}|$ is greater than or equal to 1:
 1. $p_i \in_* w_j$ - if this is the case then $\neg p_i \in_* X_{ij}$ and $(p_i, \neg p_i) \in_* X_{jk}$
 2. $\neg p_i \in_* w_j$ - if this is the case then $p_i \in_* X_{ij}$ and $(p_i, \neg p_i) \notin_* X_{jk}$
 3. $(p_i, \neg p_i) \in_* w_j$ - if this is the case then $p_i \in_* X_{jk}$ and $(p_i, \neg p_i) \notin_* X_{ij}$

□

Theorem B.3.2. $(D \div A) + \neg A$ is equivalent to $(D + \neg A) \div A$

Proof. In the trivial case where $A \notin D$, $(D \div A) = D$, so $(D \div A) + \neg A$ reduces to $D + \neg A$. It is also the case that $A \notin (D + \neg A)$, so $(D + \neg A) \div A = D + \neg A$.

When $A \in D$ and $\neg A \in D$, both operations reduce to $D \div A$.

Otherwise, we have the following.

Let X stand for the set of states corresponding to D . When $A \in D$ the set of states Y corresponding to $D + \neg A$ will all be paraconsistent states and $Y \subset X$. These states will remain present in $(D + \neg A) \div A$. Let Z_1 stand for the set of states with closest distance to $D + \neg A$ at which A fails: $(D + \neg A) \div A = Y \cup Z_1$

Let Z_2 stand for the set of states with closest distance to X at which A fails.

- $(D + \neg A) \div A = (D \cap \neg A) \cup Z_1$
- $(D \div A) + \neg A = (D \cup Z_2) \cap \neg A = (\neg A \cap D) \cup (\neg A \cap Z_2)$
- $\neg A \cap Z_2 = Z_2$ (since each world in Z_2 fails to satisfy A it must satisfy $\neg A$ so $Z_2 \subseteq \neg A$)
- $\therefore (D \div A) + \neg A = (\neg A \cap D) \cup Z_2$

Therefore showing that $Z_1 = Z_2$ will suffice to prove this equivalence.

$D + \neg A$ will result in the removal of some states in which A holds, and these states will have some T/F valuations. Their corresponding states in which the T/F valuations are replaced by B valuations will remain. These remaining states will ensure that the set of states with closest distance to Y at which A fails (Z_1) will remain the same as the set of states with closest distance to X at which A fails (Z_2), since the closeness of these states is due to the B valuations of the remaining states. For example, say the models for D are w_{13} (B, B, T) and w_{14} (B, B, B). The closest state at which w fails is w_{15} (B, B, F), whose closeness is due to (B, B, B) rather than (B, B, T). So when (B, B, T) is removed from the centre due to $D + \neg A$, the position of (B, B, F) will not change.

Example B.1

Let $D = h \wedge (r \vee w)$, which corresponds to states $\{w_1 - w_8, w_{10} - w_{17}\}$. Here are its sphere rankings:

sphere	Δ	set of states
0	0	$\{w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8, w_{10}, w_{11}, w_{12}, w_{13}, w_{14}, w_{15}, w_{16}, w_{17}\}$
1	0.5	$\{w_9, w_{18}, w_{19}, w_{20}, w_{21}, w_{22}, w_{23}, w_{24}, w_{25}, w_{26}\}$
2	1	$\{w_{27}\}$

Let $A = r \vee w$. The set of closest states at which A fails is $\{w_9, w_{18}\}$. $(D \div r \vee w) + \neg(r \vee w) = \{w_5, w_6, w_8, w_9, w_{14}, w_{15}, w_{17}, w_{18}\}$.

The set of states corresponding to $D + \neg A$ is $\{w_5, w_6, w_8, w_{14}, w_{15}, w_{17}\}$. Here are its sphere rankings:

sphere	Δ	set of states
0	0	$\{w_5, w_6, w_8, w_{14}, w_{15}, w_{17}, \}$
1	0.5	$\{w_2, w_3, w_4, w_7, w_9, w_{11}, w_{12}, w_{13}, w_{16}, w_{18}, w_{23}, w_{24}, w_{26}, \}$
2	1	$\{w_1, w_{10}, w_{20}, w_{21}, w_{22}, w_{25}, w_{27}, \}$

The set of closest states at which $r \vee w$ fails is still $\{w_9, w_{18}\}$. All of the other states in level 1, including the newly added, are states in which $r \vee w$ holds. The states w_9 and w_{18} remain because the states they are closet to at level 0 remain.

$$\text{So } (D + \neg(r \vee w)) \div r \vee w = \{w_5, w_6, w_8, w_9, w_{14}, w_{15}, w_{17}, w_{18}\}$$

In both cases, w_9 and w_{18} remain at the closest distance because of the states $\{w_6, w_8, w_{15}, w_{17}\}$, which remain in the centre sphere.

□

□

Theorem B.3.3. if $A \wedge \neg A \in \mathbf{D}$ then $\mathbf{D} \div A \wedge \neg A = \{\mathbf{D} \div A\} \cup \{\mathbf{D} \div \neg A\}$

Proof. Firstly, the conditional qualification in this theorem is required for if $A \in \mathbf{D}$ and $\neg A \notin \mathbf{D}$ then:

- $\mathbf{D} \div A \wedge \neg A = \mathbf{D}$.
- $\mathbf{D} \div \neg A = \mathbf{D}$.
- $\mathbf{D} \div A = \mathbf{D} \cup X$, where X is the closest set of states where A fails.
- $\mathbf{D} \neq \mathbf{D} \cup X \cup \mathbf{D}$.

Similarly for $\neg A \in \mathbf{D}$ and $A \notin \mathbf{D}$

If $A \notin \mathbf{D}$ and $\neg A \notin \mathbf{D}$ then the result for all operations is $\mathbf{D}$ and the equation trivially holds.

Otherwise, here are the points for a proof. Each state in the core will be a paraconsistent state. What we have to do is show that the closest sphere at which $A \wedge \neg A$ fails is also one in which there will be states at which A fails and states at which $\neg A$ fails:

- Put A and $\neg A$ into CNF
- If A is an atom p , then there is at least one state that is one step away in the next sphere where $v(p) = F$ and p fails (because of the minimal distance between B and F). But this also means that there is a state in that sphere where $v(p) = T$ and $\neg p$ fails.
- More generally, say A is of the form $p_1 \vee p_2 \vee \dots \vee p_n$ and $\neg A$ is of the form $\neg p_1 \wedge \neg p_2 \wedge \dots \wedge \neg p_n$. Then all states in the centre sphere will be such that (1) at least one will have just one B valuation over $p_1, p_2, \dots, p_n$ (2) no valuation over $p_1, p_2, \dots, p_n$ is T. To find a state in the closest sphere where $p_1 \vee p_2 \vee \dots \vee p_n$ fails, simply get these inner sphere states with one B valuation over $p_1, p_2, \dots, p_n$ and change the B to an F. To find a state in the closest sphere where $\neg p_1 \wedge \neg p_2 \wedge \dots \wedge \neg p_n$ fails, simply get these inner sphere states with one B valuation over $p_1, p_2, \dots, p_n$ and change the B to an T. The result of both procedures will end up in the same sphere.

□

Theorem B.3.4. The set of states that the Conjunction Removal Consolidation Procedure steps 1-3 result in will contain a subset of classical states.

Proof. Let M stand for all the models that correspond to a statement A . Let w_B stand for the paraconsistent state in which every atom has a B valuation. Define the degree of inconsistency of a statement A as the lowest number of atomic B valuations out of any state in M :

$$\text{inconsistency}(A) = \min(\{\Delta(w_B, w) | w \in M\})$$

- Given this definition classically satisfiable statements have a zero degree of inconsistency.
- Given an inconsistent statement A with a degree of inconsistency x , the closest sphere at which there will be a classical state is at a distance of x spheres away.
- As can be seen from the proof of Theorem B.3.3, every contraction of a contradictory element A or $\neg A$ results in the addition of at least one state to the centre sphere with a degree of contradiction $x - 1$.

- The statement A , with a degree of inconsistency x , has at least x contradictory pairs. In cases such as Example 3.18 there are technically two contradictory pairs $\{(h \vee r, \neg h \wedge \neg r), (r \vee w, \neg r \wedge \neg w)\}$, though since $\neg r$ is common to both the degree of inconsistency equals to 1.
- Therefore successive contraction until all contradictions are removed will result in statements with a zero degree of inconsistency (i.e. classical states) being in the centre sphere.

□

Appendix C

Environmental Information and Information Flow

C.1 Probabilistic Information

Theorem C.1.1. $\Pr(B|A) > \Pr(B|\neg A)$ iff $\Pr(B|A) > \Pr(B)$

Proof.

$$\Pr(B|A) = \frac{\Pr(B \wedge A)}{\Pr(A)}$$
$$\Pr(B|\neg A) = \frac{\Pr(B \wedge \neg A)}{\Pr(\neg A)}$$

$$\begin{aligned}\Pr(B|A) &> \Pr(B|\neg A) \\ \frac{\Pr(B \wedge A)}{\Pr(A)} &> \frac{\Pr(B \wedge \neg A)}{\Pr(\neg A)} \\ \Pr(B \wedge A)\Pr(\neg A) &> \Pr(B \wedge \neg A)\Pr(A) \\ \Pr(B \wedge A) \times (1 - \Pr(A)) &> \Pr(B \wedge \neg A) \times \Pr(A) \\ \Pr(B \wedge A) - \Pr(B \wedge A)\Pr(A) &> \Pr(A)\Pr(B \wedge \neg A) \\ \Pr(B \wedge A) &> \Pr(A)(\Pr(B \wedge \neg A) + \Pr(B \wedge A)) \\ \Pr(B \wedge A) &> \Pr(A)\Pr(B) \\ \Pr(B|A) &> \Pr(B)\end{aligned}$$

□

C.2 The Arrow of Information Flow

Given the specific use of inverse conditional probabilities, it seems that with accounts such as Dretske's, if A carries the information that B , then B is prior to or concurrent with A . Is it possible for A to carry the information that B if A precedes B ? Expanding the domain to include cases where A (present) carries the information that B (future) is something to consider. In terms of Dretske's definition, technically this would simply mean employing conditional probabilities. The modal logical framework is already neutral with regards to the direction of information flow.

Here is a basic example of how this would all work, starting with a case of event followed by signal. It is the weekend and you are informed by a friend that an event occurred on Tuesday or Wednesday last week. You then find newspapers for both of those days and discover that it was the Wednesday. This is all very straightforward; the article in the Wednesday newspaper carries the information that the event occurred on the Wednesday.

In a case of signal followed by event on the same weekend, you are informed that you will have a meeting with someone on either Tuesday or Wednesday. If by the weekend's end the person must do something else, they will do it on the Tuesday and your meeting will be on the Wednesday. If not, your meeting will be on the Tuesday. By Sunday's end you learn that the person has been asked to attend a function on Tuesday. So this signal carries the information that your meeting will be on Wednesday.

Knowledge about future events is tricky. Hawthorne questions applying Dretske's counterfactual conclusive reasons to scenarios such as the following:

I think on Monday that I am going to meet you on Wednesday on the basis of familiar sorts of reasons. I meet you on Wednesday. It is natural to say, in many such cases, that my belief on Monday was a piece of knowledge. Consider now the counterfactual: if you hadn't met me on Wednesday, I would not have had those reasons on Monday for thinking that you would meet me on Wednesday. I do not find that counterfactual very compelling. ... So my reasons turn out not to be conclusive, since they fail Dretske's counterfactual test. [97, p. 37]

Dretske replies:

... I think the counterfactual condition accords with common and widespread intuitions about knowledge of the future. In my more careful moments, I prefer

to say (on Monday) not that I know that I will be in my office on Wednesday, but that I know that is where I intend or plan to be on Wednesday. I retreat to this more cautious claim because I realize that my reasons for thinking I will be in my office on Wednesday are reasons I would have on Monday even if a Tuesday accident were to prevent me from being there on Wednesday. As long as I might not make it to the office, I do not know I will be there. How could I? One doesn't, generally speaking, have conclusive reasons for believing that something will happen. Death may be an exception, but I doubt whether taxes are. I'm a bit of a skeptic about the future. [57, p. 44]

Let us frame further discussion of this issue in terms of information. Contrary to his comments above, couldn't Dretske's idea of relevant alternatives be applied here? It seems that the uncertainty exhibited in his comment "as long as I might not make it to the office, I do not know I will be there" could just as equally well be applied to information/knowledge claims about the past/present. For example, could we not say 'as long as it might be the case that the clock has freakishly malfunctioned, I do not know that the time is such-and-such'? As discussed earlier, Dretske adopts a more realistic, common sense way of dealing with things by setting a conditional probability of one against a background of certain conditions. Likewise, couldn't the same thing be done for future events such as the one described in Hawthorne's example? Yes, something could happen, but if extraneous possibilities were treated as irrelevant, then the required conditional probability of one could be determined and information flow claims made. Perhaps Dretske does have another reason in mind that only applies to future events. At any rate, if a general contextualist approach is adopted then given a relatively relaxed context the familiar sorts of reasons are enough to carry the information that I am going to meet you on Wednesday.

One potential problem that comes to mind is the truth value of contingent future events. In the case of a clock, as long as it is actually functioning correctly and the time is such-and-such, then the clock is providing information, despite there being irrelevant possibilities that something could have/could happen. So a person can come to know that it is 5pm by basing their true belief on this information; importantly, the proposition in question is true. But in the Wednesday meeting case above, if the truth value of the proposition in question was to be considered undetermined (think Lukasiewicz's third truth value for future contingent events), then the belief fails one of the conditions for knowledge, namely the truth of the proposition in question.

On a related note, another issue could be the indeterminacy of the actual state, which is always a relevant alternative. With the clock example, the relevant alternative that is the actual state is one in which the clock is working. With the Wednesday meeting case, if what

actually happens (either a meeting or no meeting) is undetermined, then what is the relevant alternative corresponding to the actual state?

To deal with both of these issues one strategy is to pre-set the actual state to one in which the meeting does occur and the proposition is true. If it turns out that the meeting does not go ahead, then things change. Because of this, perhaps contextualist informational knowledge about the future should be treated as a special kind of tentative knowledge.

Bibliography

- [1] F. Adams. Information and knowledge a la floridi. *Metaphilosophy*, 41(3):331–344, 2010.
- [2] Fred Adams. The informational turn in philosophy. *Minds and Machines*, 13:471–501, 2003.
- [3] Fred Adams. Knowledge. In Luciano Floridi, editor, *The Blackwell Guide to the Philosophy of Computing and Information*, pages 228–236. Blackwell, 2004.
- [4] Fred Adams. Tracking theories of knowledge. *Veritas*, 50:1–35, 2005. Draft available online at <http://revistaseletronicas.pucrs.br/ojs/index.php/veritas/article/viewFile/1813/1343>.
- [5] Fred Adams, John A. Barker, and Julia Figurelli. Towards closure on closure. *Synthese*, 2011. Online First.
- [6] Jonathan Adler. Epistemological problems of testimony. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Winter 2010 edition, 2010.
- [7] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *J. Symb. Log.*, 50(2):510–530, 1985.
- [8] Patrick Allo. Being informative: Information as information handling. Online paper. URL = <<http://homepages.vub.ac.be/~pallo/papers/wspi-final.pdf>>.
- [9] Patrick. Allo. Formalising the ‘no information without data-representation’ principle. In A. Briggie, K. Waelbers, and P.A.E. Brey, editors, *Current Issues in Computing and Philosophy*. IOS Press, 2008.
- [10] Patrick Allo. A classical prejudice? *Knowledge, Technology & Policy*, 23(1–2):25–40, 2010.
- [11] Patrick Allo and Edwin Mares. Informational semantics as a third alternative? *Erkenntnis*, 2011. Online First.

- [12] Horacio Arlo-Costa. The logic of conditionals. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Stanford University, spring 2009 edition, 2009.
- [13] Jie. Bao. Towards a theory of semantic communication. Technical report, Rensselaer Polytechnic Institute, 2011. URL = <http://www.cs.rpi.edu/~baojie/pub/2011-03-28_nsw_tr.pdf>.
- [14] Y. Bar-Hillel and R. Carnap. An outline of a theory of semantic information. Research Laboratory of Electronics Technical Report 247, Massachusetts Institute of Technology, 1952. URL = <<http://dspace.mit.edu/bitstream/handle/1721.1/4821/RLE-TR-247-03150899.pdf?sequence=1>>.
- [15] Y. Bar-Hillel and R. Carnap. Semantic information. *The British Journal for the Philosophy of Science*, 4(14):147–157, 1953. URL = <<http://hum.uchicago.edu/~jagoldsm/Courses/2008HistoryOfLx/barhillelCarnap.pdf>>.
- [16] J. Barwise and J. Seligman. *Information Flow: The Logic of Distributed Systems*. Cambridge University Press, Cambridge, MA, 1997.
- [17] P. Baumann. Information, closure, and knowledge: On jager’s objection to dretske. *Erkenntnis*, 64:403–408, 2006.
- [18] Johan Van Benthem, Patrick Girard, and Olivier Roy. Everything else being equal: A modal logic for ceteris paribus preferences. *Journal of Philosophical Logic*, 38(1):83–125, 2009.
- [19] Sven Bernecker and Fred I. Dretske. *Knowledge: Readings in Contemporary Epistemology*. Oxford University Press, 2000.
- [20] Tim Black. Contextualism in epistemology. In James Fieser and Bradley Dowden, editors, *Internet Encyclopedia of Philosophy*. July 15, 2006 edition, 2006. URL = <<http://http://www.iep.utm.edu/contextu/>>.
- [21] P. Borlund. The concept of relevance in ir. *Journal of the American Society for Information Science and Technology*, 54(10):913–925, 2003.
- [22] Manuel Bremer and Daniel Cohnitz. *Information and Information Flow: An Introduction*. Ontos Verlag, Germany, 2004.
- [23] Manuel E. Bremer. Do logical truths carry information? *Minds and Machines*, 13(4):567–575, 2003.
- [24] Jeffrey Bub. Quantum entanglement and information. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Winter 2010 edition, 2010.

- [25] M. Burgin. *Theory of Information: Fundamentality, Diversity and Unification*. World Scientific, Singapore, 2010.
- [26] R. Carnap. *Logical Foundations of Probability*. University of Chicago Press, Chicago, 1950.
- [27] G. Cevolani. Verisimilitude and strongly semantic information. *Etica & Politica*, 13(2):159–179, 2011.
- [28] G Cevolani, V Crupi, and R. Festa. Verisimilitude and belief change for conjunctive theories. *Erkenntnis*, 75(2):183–202, 2011.
- [29] Gustavo Cevolani, Vincenzo Crupi, and Roberto Festa. The whole truth about linda: probability, verisimilitude, and a paradox of conjunction. In Marcello D’Agostino et al., editor, *New Essays in Logic and Philosophy of Science*, pages 603–615. College Publications, 2010.
- [30] G.J. Chaitin. *Algorithmic Information Theory*. Cambridge University Press, Cambridge, 1987.
- [31] David Chapman. Information, meaning and context. In Magnus Ramage and David Chapman, editors, *Perspectives on Information*, pages 36–50. Routledge, 2011.
- [32] Brian F. Chellas. *Basic Conditional Logic*, volume 4. 1975.
- [33] Brian F. Chellas. *Modal Logic: An Introduction*. Cambridge University Press, Cambridge, 1980.
- [34] Laurence Cholvy and Stephanie Roussel. Towards agent-oriented relevant information. *Lecture Notes in Computer Science*, 5253:22–31, 2008. URL = <<http://www.springerlink.com/content/b156r01742512q64/>>.
- [35] C. A. J. Coady. *Testimony: A Philosophical Study*. Oxford University Press, 1992.
- [36] Jonathan Cohen and Aaron Meskin. An objective counterfactual theory of information. *Australasian Journal of Philosophy*, 84(3):333–352, 2006.
- [37] Jonathan Cohen and Aaron Meskin. Counterfactuals, probabilities, and information: Response to critics. *Australasian Journal of Philosophy*, 86(4):635–642, 2008.
- [38] T. R. Colburn. Information, thought, and knowledge. *Proceedings of the World Multi-conference on Systemics, Cybernetics and Informatics (Orlando FL, July 23-26,2000)*, X:467–471, 2000.
- [39] John M. Collins. Epistemic closure principles. In James Fieser and Bradley Dowden, editors, *Internet Encyclopedia of Philosophy*. June 13, 2006 edition, 2006. URL = <<http://www.iep.utm.edu/epis-clo/>>.

- [40] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, second edition, 2006.
- [41] Mukesh Dalal. Investigations into a theory of knowledge base revision. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, pages 475–479, 1988.
- [42] Marian David. Truth as the primary epistemic goal: A working hypothesis. In Matthias Steup and Ernest Sosa, editors, *Contemporary Debates in Epistemology*, pages 296–310. Wiley-Blackwell, 2003.
- [43] Hilmi Demir. Counterfactuals vs. conditional probabilities: A critical analysis of the counterfactual theory of information. *Australasian Journal of Philosophy*, 86(1):45–60, 2008.
- [44] Hilmi Demir. The counterfactual theory of information revisited. *Australasian Journal of Philosophy*, pages 1–3, 2011. Online First.
- [45] Keith DeRose. *The Case for Contextualism*. Oxford University Press, 2009.
- [46] K. J. Devlin. *Logic and Information*. Cambridge University Press, Cambridge, 1991.
- [47] Igor Douven. The lottery paradox and our epistemic goal. *Pacific Philosophical Quarterly*, 89(2):204–225, 2008.
- [48] Anthony Doyle. Is knowledge information-produced belief? a defense of dretske against some critics. *The Southern Journal of Philosophy*, XXIII:33–46, 1985.
- [49] Fred Dretske. Epistemic operators. *Journal of Philosophy*, 67:1007–1023, 1970.
- [50] Fred Dretske. Conclusive reasons. *Australasian Journal of Philosophy*, 49:1–22, 1971.
- [51] Fred Dretske. *Knowledge and the Flow of Information*. MIT Press, Cambridge, MA, 1981.
- [52] Fred. Dretske. Precis of knowledge and the flow of information. *Behavioral and Brain Sciences*, 6(1):55–63, 1983.
- [53] Fred Dretske. Information theory and epistemology. In E. Craig, editor, *Routledge Encyclopedia of Philosophy*. Routledge, 1998.
- [54] Fred Dretske. Skepticism: What perception teaches. In Steven Luper, editor, *The Sceptics: Contemporary Essays*, pages 105–118. Aldershot: Ashgate Publishing, 2003.
- [55] Fred Dretske. Externalism and modest contextualism. *Erkenntnis*, 61:173–186, 2004.
- [56] Fred Dretske. The case against closure. In Matthias Steup and Ernest Sosa, editors, *Contemporary Debates in Epistemology*. Blackwell, 2005.

- [57] Fred Dretske. Reply to hawthorne. In Matthias Steup and Ernest Sosa, editors, *Contemporary Debates in Epistemology*. Blackwell, 2005.
- [58] Fred. Dretske. Information and closure. *Erkenntnis*, 64:409–413, 2006.
- [59] Fred Dretske. Epistemology and information. In Johan. Adriaans, Pieter.; van Benthem, editor, *Handbook on the Philosophy of Information*. Elsevier, 2008. Draft available online at http://www.illc.uva.nl/HPI/Draft_Epistemology_and_Information.pdf.
- [60] Fred I. Dretske. A cognitive cul-de-sac. *Mind*, 91(361):109–111, 1982.
- [61] Marcello D.’Agostino and Luciano Floridi. The enduring scandal of deduction. *Synthese*, 167(2), 2009.
- [62] J.M. Dunn. Information in computer science. In Johan. Adriaans, Pieter.; van Benthem, editor, *Handbook on the Philosophy of Information*. Elsevier, 2008. Draft available online at http://www.illc.uva.nl/HPI/Draft_Information_in_Computer_Science.pdf.
- [63] Satu Eloranta, Raul Hakli, Olli Niinivaara, and Matti Nykanen. Accommodative belief revision. In *JELIA ’08*, pages 180–191, 2008.
- [64] L. Floridi. Open problems in the philosophy of information. *Metaphilosophy*, 35(4):554–582, 2004.
- [65] L. Floridi. Outline of a theory of strongly semantic information. *Minds and Machines*, 14(2):197–221, 2004.
- [66] L. Floridi. Is semantic information meaningful data? *Philosophy and Phenomenological Research*, 70(2):351–370, 2005.
- [67] L. Floridi. The logic of being informed. *Logique et Analyse*, 49(196):433–460, 2006.
- [68] L. Floridi. In defence of the veridical nature of information. *European Journal of Analytic Philosophy*, 3(1):31–41, 2007.
- [69] L. Floridi. Understanding epistemic relevance. *Erkenntnis*, 69(1):69–92, 2008.
- [70] L. Floridi. The philosophy of information: Ten years later. *Metaphilosophy*, 41(3):402–419, 2010.
- [71] L. Floridi. Semantic information and the correctness theory of truth. *Erkenntnis*, 74(2):147–175, 2010.
- [72] Luciano Floridi. On the logical unsolvability of the gettier problem. *Synthese*, 142(1):61–79, 2004.

- [73] Luciano Floridi. Semantic conceptions of information. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Stanford University, summer 2009 edition, 2009. URL = <<http://plato.stanford.edu/archives/sum2009/entries/information-semantic/>>.
- [74] Luciano Floridi. *Information: A Very Short Introduction*. Oxford University Press, 2010.
- [75] Luciano Floridi. The philosophy of information as a conceptual framework. *Knowledge, Technology and Policy*, 23(1-2):253–281, 2010.
- [76] Luciano Floridi. *The Philosophy of Information*. Oxford University Press, New York, 2011.
- [77] Luciano Floridi. Semantic information and the network theory of account. *Synthese*, 184(3):431–454, 2012.
- [78] Richard Foley. Dretske’s “information-theoretic” account of knowledge. *Synthese*, 70(2):159–184, 1987.
- [79] C.J. Fox. *Information and Misinformation: An Investigation of the Notions of Information, Misinformation, Informing, and Misinforming*. Greenwood Press, Westport CT, U.S.A, 1983.
- [80] P. Gardenfors. On the logic of relevance. *Synthese*, 37:351–367, 1978.
- [81] P. Gardenfors. *Knowledge in Flux. Modeling the Dynamics of Epistemic States*. The MIT Press, Cambridge, Massachusetts, 1988.
- [82] Edmund Gettier. Is justified true belief knowledge? *Analysis*, 23:121–123, 1963. Available online at <http://www.ditext.com/gettier/gettier.html>.
- [83] Peter Godfrey-Smith and Kim Sterelny. Biological information. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2008 edition, 2008.
- [84] Alvin Goldman. What is justified belief? In *Justification and Knowledge*, pages 1–25. Boston: D. Reidel, 1979.
- [85] Alvin Goldman. Reliabilism. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2011 edition, 2011.
- [86] Alvin I. Goldman and Erik J. Olsson. Reliabilism and the value of knowledge. In D. Pritchard, A. Millar, and A. Haddock, editors, *Epistemic Value*. Oxford University Press, Oxford, 2008. Preprint available at <http://fas-philosophy.rutgers.edu/goldman/Reliabilism%20and%20the%20Value%20of%20Knowledge.pdf>.

- [87] G. Graham. *The Internet: A Philosophical Inquiry*. Routledge, London, 1999.
- [88] Peter J. Graham. Conveying information. *Synthese*, (123):365–392, 2000.
- [89] Peter J. Graham. Transferring knowledge. *Noûs*, 34(1):131–152, 2000.
- [90] H. P. Grice. Meaning. *The Philosophical Review*, 64:377–388, 1957.
- [91] H. P. Grice. *Studies in the Way of Words*. Harvard University Press, 1989.
- [92] A. Grove. Two modellings for theory change. *Journal of Philosophical Logic*, 17:157–170, 1988.
- [93] Cevolani Gustavo and Franceso Calandra. Approaching the truth via belief change in propositional languages. In *EPSA Epistemology and Methodology of Science: Launch of the European Philosophy of Science Association*, pages 47–62, 2010.
- [94] Sven Ove. Hansson. A survey of non-prioritised belief revision. *Erkenntnis*, 50(2-3):413–427, 1999.
- [95] Sven Ove. Hansson. Ten philosophical problems in belief revision. *Journal of Logic and Computation*, 13(1):37–49, 2003.
- [96] Sven Ove. Hansson. Logic of belief revision. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Stanford University, spring 2009 edition, 2009. URL = <<http://plato.stanford.edu/archives/spr2009/entries/logic-belief-revision/>>.
- [97] John Hawthorne. The case for closure. In Matthias Steup and Ernest Sosa, editors, *Contemporary Debates in Epistemology*. Blackwell, 2005.
- [98] M. Heller. Relevant alternatives and closure. *Australasian Journal of Philosophy*, 77(2):196–208, 1999.
- [99] David Hemp. The kk (knowing that one knows) principle. In James Fieser and Bradley Dowden, editors, *Internet Encyclopedia of Philosophy*. October 15, 2006 edition, 2006. URL = <<http://www.iep.utm.edu/kk-princ/>>.
- [100] Vincent Hendricks and John Symons. Epistemic logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2009 edition, 2009.
- [101] Vincent F. Hendricks. *Mainstream and Formal Epistemology*. Cambridge University Press, 2006.
- [102] Stephen Hetherington. Gettier problems. In James Fieser and Bradley Dowden, editors, *The Internet Encyclopedia of Philosophy*. 2005. URL = <<http://www.iep.utm.edu/gettier/>>.

- [103] J. Hintikka. On semantic information. In P. Hintikka, J.; Suppes, editor, *Information and Inference*, Dordrecht, Holland, 1970. D. Reidel.
- [104] J. Hintikka. Surface information and depth information. In P. Hintikka, J.; Suppes, editor, *Information and Inference*, Dordrecht, Holland, 1970. D. Reidel.
- [105] J. Hintikka. *The semantics of questions and the questions of semantics*. North-Holland, North-Holland, 1976.
- [106] J. Hintikka. Some varieties of information. *Information Processing & Management*, 20(1-2):175–181, 1984.
- [107] D. Hitchcock. Relevance. *Argumentation*, 6(2):251–270, 1992.
- [108] Alan Hjek. Interpretations of probability. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2010 edition, 2010.
- [109] Wesley H. Holliday. Epistemic logic and relevant alternatives. In M. Slavkovik, editor, *Proceedings of the 15th Student Session of the European Summer School in Logic, Language, and Information*, pages 4–16. 2010. copy available at <http://marija.gforge.uni.lu/proceedings.pdf>.
- [110] Joachim Horvath. Why the conditional probability solution to the swamping problem fails. *Grazer Philosophische Studien*, 79(1):115–120, 2009.
- [111] Christoph. Jager. Skepticism, information, and closure: Dretske’s theory of knowledge. *Erkenntnis*, 61:187–201, 2004.
- [112] Christoph Jager and Wayne A. Davis. Reliabilism and the extra value of knowledge. *Philosophical Studies*, 157(1):93–105, 2010.
- [113] William James. *Pragmatism and the Meaning of Truth*. Harvard UP, 1975.
- [114] Christoph Kelp. A minimalist approach to epistemology, 2007. Ph.D. Thesis, Department of Philosophy, University of Stirling.
- [115] J. Keynes. *A Treatise on Probability*. Macmillan, London, 1921.
- [116] Peter Klein. Epistemology. In E. Craig, editor, *Routledge Encyclopedia of Philosophy*. 2005.
- [117] H. Kyberg. *Probability and the Logic of Rational Belief*. Wesleyan University Press, Middletown CT, 1961.
- [118] G. Lakemeyer. Relevance from an epistemic perspective. *Artificial Intelligence*, 97:137–167, 1997.

- [119] Lissa Legee. How does the venus flytrap digest flies? <http://www.scientificamerican.com/article.cfm?id=how-does-the-venus-flytra>, August 2002. Scientific American ‘Ask the Experts’ response.
- [120] Keith Lehrer and Thomas Paxson Jr. Knowledge: Undefeated justified true belief. *Journal of Philosophy*, 66(8):225–237, 1969.
- [121] J.J. Levesque. A logic of implicit and explicit belief. In *Proceedings of the National Conference on Artificial Intelligence*, pages 198–202. 1984.
- [122] Isaac. Levi. *Gambling With Truth: An Essay on Induction and the Aims of Science*. Harper & Row, New York, 1967.
- [123] D. Lewis. *Counterfactuals*. Blackwell, Oxford, 1973.
- [124] David Lewis. Elusive knowledge. In Steven Luper, editor, *The Skeptics: Contemporary Essays*, pages 119–136. Aldershot: Ashgate Publishing, 2003.
- [125] Franck Lihoreau. Relevant alternatives contextualism and ordinary contingent knowledge. *Disputatio*, 2(24):281–294, 2008.
- [126] Barry Loewer. Information and belief. *Behavioral and Brain Sciences*, 6(1):75–76, 1983.
- [127] Steven Luper. The epistemic closure principle. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Stanford University, fall 2010 edition, 2010. URL = <<http://plato.stanford.edu/archives/fall2010/entries/closure-epistemic/>>.
- [128] Michael P. Lynch. Minimilism and the value of truth. *The Philosophical Quarterly*, 54(217):497–517, 2004.
- [129] J. Christopher Maloney. Drestske on knowledge and information. *Analysis*, 43:25–28, 1983.
- [130] Edwin Mares. Relevant logic and the theory of information. *Synthese*, 109(3):345–360, 1997.
- [131] D. Markinson. Screened revision. *Theoria*, 63(1-2):14–23, 1997.
- [132] John. Mingers. Prefiguring floridi’s theory of semantic information. Working Paper Series 235, Kent Business School, 2010. URL = <<http://www.kent.ac.uk/kbs/documents/research/working-papers/2010/235-mingers.pdf>>.
- [133] I. Niiniluoto. *Truthlikeness*. D. Reidel, Dordrecht, Holland, 1987.

- [134] I. Niiniluoto. Review of ‘likeness to the truth’. *The Journal of Symbolic Logic*, 54(1):296–297, 1989.
- [135] I. Niiniluoto. Theory change, truthlikeness, and belief revision. In *EPSA Epistemology and Methodology of Science: Launch of the European Philosophy of Science Association*, pages 189–199, 2010.
- [136] Ilkka Niiniluoto. Scientific progress. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Stanford University, summer 2011 edition edition, 2011. URL = <<http://plato.stanford.edu/archives/sum2011/entries/scientific-progress/>>.
- [137] Robert Nozick. *Philosophical Explanations*. Harvard University Press, 1981.
- [138] G. Oddie. *Likeness to Truth*. D. Reidel, Dordrecht, Holland, 1986.
- [139] Graham Oddie. Truthlikeness. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Stanford University, fall 2008 edition, 2008. URL = <<http://plato.stanford.edu/archives/fall2008/entries/truthlikeness/>>.
- [140] Erik J. Olsson. In defense of the conditional probability solution to the swamping problem. *Grazer Philosophische Studien*, 79(1):93–114, 2009.
- [141] Eric Pacuit. Neighborhood semantics for modal logic: An introduction, 2007. Lecture notes for ESSLLI 2007 course. URL = <<http://ai.stanford.edu/~epacuit/classes/esslli/nbhdesslli.pdf>>.
- [142] John Perry and David Israel. What is information? In P. Hanson, editor, *Information, Language and Cognition*, Oxford, 1990. Blackwell.
- [143] John Perry and David Israel. Information and architecture. In John Barwise, M Gawron, G Pltokin, and S Tutiya, editors, *Situation Theory and Its Architecture*, Stanford University, 1991. Stanford University.
- [144] John R. Pierce. *An Introduction to Information Theory: Symbols, Signals and Noise*. Dover Publications, New York, second edition, 1980.
- [145] Christian Piller. Valuing knowledge: A deontological approach. *Ethical Theory and Moral Practice*, 12(4):413–428, 2008.
- [146] Christian Piller. Reliabilist responses to the value of knowledge problem. *Grazer Philosophische Studien*, 79(1):121–135, 2009.
- [147] Ted Poston. Internalism and externalism in epistemology. In James Fieser and Bradley Dowden, editors, *Internet Encyclopedia of Philosophy*. 2008. URL = <<http://www.iep.utm.edu/epis-clo/>>.

- [148] Graham Priest. *An Introduction to Non-Classical Logic*. Cambridge University Press, Cambridge, United Kingdom, 2003.
- [149] Graham Priest and Koji Tanaka. Paraconsistent logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Stanford University, summer 2009 edition, 2009. URL = <<http://plato.stanford.edu/archives/sum2009/entries/logic-paraconsistent/#ManyValLog>>.
- [150] Duncan Pritchard. Knowing the answer, understanding and epistemic value. *Grazer Philosophische Studien*, 77:325–339, 2008.
- [151] Duncan Pritchard. The value of knowledge. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2008 edition, 2009. URL = <<http://plato.stanford.edu/archives/fall2008/entries/knowledge-value/>>.
- [152] Duncan Pritchard. The value of knowledge. *The Harvard Review of Philosophy*, 16(1), 2009.
- [153] Duncan. Pritchard. What is the swamping problem? In A. Reisner and A. Steglich-Petersen, editors, *Reasons for Belief*, pages 244–259. Cambridge University Press, 2011.
- [154] Manuel Rebuschi and Franck Lihoreau. Contextual epistemic logic. In Helge Ruckert Degremont, Cedric. Keiff Laurent, editor, *Dialogues, Logics and Other Strange Things. Essays in Honour of Shahid Rahman*. College Publications, 2008.
- [155] M. Reitz, Joan. Information ethics. In *Online Dictionary for Library and Information Science*. 2010.
- [156] Greg Restall. Information flow and relevant logics. In J. Seligman and D. Westerstahl, editors, *Logic, Language and Computation*, pages 463–477. CSLI, Stanford, 1996.
- [157] Greg Restall and John Slaney. Realistic belief revision, January 1995. URL = <<http://consequently.org/papers/rbr.pdf>>.
- [158] Odinaldo Rodrigues, Dov Gabbay, and Alessandra Russo. Belief revision. In D.M Gabbay et al., editor, *Handbook of Philosophical Logic*, volume 16, pages 1–114. Springer, 2010.
- [159] Bertrand. Russell. *The Problems of Philosophy*. Oxford University Press, 1997. Originally Published 1912.
- [160] Patrick Rysiew. Epistemic contextualism. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Winter 2011 edition, 2011.
- [161] Andrea Scarantino. Shell games, information, and counterfactuals. *Australasian Journal of Philosophy*, 86(4):629–634, 2008.

- [162] Andrea Scarantino and Gualtiero Piccinini. Information without truth. *Metaphilosophy*, 41(3):313–330, 2010.
- [163] S. Sequoiah-Grayson. The metaphilosophy of information. *Minds and Machines*, 17(3):331–344, 2007.
- [164] Sebastian Sequoiah-Grayson. The scandal of deduction. *Journal of Philosophical Logic*, 37(1), 2008.
- [165] Nicholas. Shackel. Shutting dretske’s door. *Erkenntnis*, 64:393–401, 2006.
- [166] C. E. Shannon. *Collected Papers*. IEEE Press, New York, 1993. edited by N. J. A. Sloane and A. D. Wyner.
- [167] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, 1949.
- [168] Claude Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656, 1948. URL = <<http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf>>.
- [169] Matthias Steup. Knowledge and skepticism. In Matthias Steup and Ernest Sosa, editors, *Contemporary Debates in Epistemology*. Blackwell, 2005.
- [170] Matthias Steup. The analysis of knowledge. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2012 edition, 2006.
- [171] Matthias Steup. Epistemology. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Winter 2011 edition, 2011.
- [172] G. C. Stine. Skepticism, relevant alternatives, and deductive closure. *Philosophical Studies*, 29:249–261, 1976.
- [173] A. Tversky and D. Kahneman. Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological Review*, 90(2):293–315, 1983.
- [174] Alasdair Urquhart. Many-valued logic. In Gabbay D. and Guenther F., editors, *Handbook of Philosophical Logic*, pages 71–116. D. Reidel Publishing Company, iii edition, 1986.
- [175] Warren Weaver. The mathematics of communication. *Scientific American*, 181(1):11–15, 1949.
- [176] Timothy Williamson. Probability and danger, 2009. The Amherst Lecture in Philosophy.

- [177] M. Winslett. *Updating Logical Databases*. Cambridge University Press, Cambridge, United Kingdom, 1990.
- [178] Wayne Wobcke. An information-based theory of conditionals. *Notre Dame Journal of Formal Logic*, 41:95–141, 2000.
- [179] Seth Yalcin. Information and probability. Chapter draft of dissertation. URL = <http://web.me.com/sethyalcin/web/work_files/yalcin.ch4.note.pdf>.
- [180] Palle Yourgrau. Knowledge and relevant alternatives. *Synthese*, 55:175–190, 1983.
- [181] Linda. Zagzebski. Virtues of the mind. In Ernest Sosa and Jaegwon Kim, editors, *Epistemology: An Anthology*, pages 457–467. Blackwell, 2000.
- [182] Michael J. Zimmerman. Intrinsic vs. extrinsic value. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Winter 2010 edition, 2010.
- [183] S.D. Zwart. *Refined Verisimilitude*. Kluwer, Dordrecht, Holland, 2001.