This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
Perspectives:
A Relativistic Approach to the Theory of Information

Jeremy Michael Seligman

Ph.D.
University of Edinburgh
1990
Acknowledgements

Firstly to my dear mother and to the memory of my father, to whom this thesis is dedicated.

Robin Cooper has been a ‘model’ supervisor to me: patient, enthusiastic, insightful, flexible, encouraging and always available to discuss ideas and problems. I would certainly have lacked the courage to pursue the ideas discussed in this thesis if it was not for his constant support and interest. Thank you, Robin.

Ideas on the ‘perspective’ theme bubbled up from the basement of No 1 Buccleuch Place in the Winter of 1988. This was the venue for numerous heated debates on many fascinating topics. My thanks therefore to Nick Chater, Mike Oaksford and Mike Malloch for forcing me to think. Nick Chater and I started writing a number of unfinished papers at that time and one which we did finish. I am grateful to him for agreeing to the inclusion of part of it in Chapter 3.

The Centre for Cognitive Science in Edinburgh is such a good place for both formal and informal discussion, in seminar rooms and in pubs that I could not begin to trace the conversations that had an influence on my work. However, Udi Rahat, Mike Reape, Patrick Blackburn, Richard Cooper, Mike McPartlin and Max Volino have been present at many of them and have been good friends of mine these last three and a half years. Thanks especially to Patrick, Richard and Max who have recently read through reams of \LaTeX{} for me.

Thanks are also due to the HCRC and DYANA projects that I am employed on for letting me spend so much time on thinking about and writing this thesis.

For
Ernest Dudgeon Seligman
Abstract

This thesis is concerned with the elucidation of the structure of three basic cognitive functions.

Firstly, an organism must be able to make distinctions between different aspects of its environment if it is to respond selectively. This is classification.

Secondly, it must be able to anticipate conditions in other parts of its environment. If an organism at $x$ is to anticipate that the condition $c$ holds at $y$ then, at the very least, the information that $c$ holds must be accessible from $x$. Hence anticipation depends on a flow of information.

Thirdly, an organism must be able to recognize uniformities across different parts of the environment. This is individuation.

We propose that each of these functions can be understood in terms of a primitive ability of 'seeing' the world from a perspective. In contrast to the possession of a conceptual scheme, or mastery of a language of thought, a characteristic of an organism's ability to adopt perspectives is the additional ability to shift from one perspective to another.

In the thesis we first propose a theory of classification. Its usefulness in categorizing different classificatory systems, like taxonomies, state systems and attribute-value structures, is demonstrated in the Appendix.

We then study two approaches to characterizing the flow of information. One, due to Dretske (1981), is based on conditional probabilities. The other, due to Barwise and Perry (1983), is based on the Situation Theoretic idea of a constraint.

Our theory of perspectives takes ideas from both accounts: from Situation Theory, the distinction between information supported and information carried by a situation, and from Dretske, the implicit relativity to an information channel.

We give a rudimentary account of the individuation of objects as predictive regularities across situations. Properties of objects individuated in this way are characterized as shifts in perspective which preserve the predictive regularity.
Finally, we consider a more concrete model of information flow (called a *world system*, Rosenschein (1989)) in which environmental conditions are understood in terms of possible state distributions over locations and times. We generalize his model and show how *information channels* offer a more sensitive account of information flow than the one induced by the global notion of possibility. Information channels are then used to construct perspectives within a world system.
## Contents

1 Classifications ............................................. 6

2 Situations ................................................. 14
   2.1 Ontological Groundwork ................................. 17
      2.1.1 Parts of the Worlds ................................. 17
      2.1.2 Schemes of Individuation ......................... 19
   2.2 Basic Infons ........................................... 23
      2.2.1 Identity ........................................... 23
      2.2.2 Appropriateness .................................. 26
   2.3 Types and Propositions ................................ 28
      2.3.1 Propositions ....................................... 30
      2.3.2 Types ............................................... 31
      2.3.3 Argument Roles ..................................... 32
      2.3.4 Appropriateness .................................... 33
      2.3.5 Being of a Type .................................... 34
      2.3.6 Indeterminates and Anchors ....................... 35
   2.4 Situations and Infons .................................. 36
      2.4.1 More Primitive Types ............................... 36
      2.4.2 Complex Infos ....................................... 37
      2.4.3 Properties ........................................... 39
      2.4.4 When Situations Support Infons ................... 40

3 Constraints ................................................... 42
   3.1 The Flow of Information ................................. 44
CONTENTS

3.1.1 Communication Theory ........................................ 44
3.1.2 Information Carried and Information Content .................. 50
3.1.3 The Analog/Digital Distinction ............................... 53
3.2 Constraints in Situation Theory ................................. 55
3.3 Attunement .................................................................. 64
3.4 Digitalization, Individuation and Cognition ..................... 73
  3.4.1 From Kapok to Bits ............................................. 73
  3.4.2 From Digitalization to Cognition .............................. 76
  3.4.3 Constraints before Individuals? ............................... 84

4 Perspectives .................................................................. 86
  4.1 A Theory of Perspectives .......................................... 88
    4.1.1 Information Content ........................................... 89
    4.1.2 Inside Perspectives ............................................. 91
    4.1.3 Perspectives from the Outside .............................. 95
  4.2 Conditional Constraints ........................................... 100
  4.3 Informational Perspectives ........................................ 104
    4.3.1 The Source/Receiver Perspective ......................... 105
    4.3.2 Informational Perspectives in General .................... 108
  4.4 Logical Perspectives ................................................ 111
    4.4.1 The Classical Logic Perspective ......................... 111
    4.4.2 Information Flow and Modality ............................ 113
    4.4.3 Modal Orderings and the Direction of Information Flow .. 117

5 Individuals .................................................................. 119
  5.1 Individuation and Predication .................................... 119
  5.2 Predictive Coherence ............................................... 121
  5.3 Objects as Uniformities ............................................ 122
  5.4 Properties as Shifts ................................................ 124
  5.5 Basic Infons Revisited ............................................ 126

6 Worlds ..................................................................... 132
  6.1 World Systems ...................................................... 133
CONTENTS

6.1.1 Aggregate Locations ........................................... 136
6.1.2 Dynamic Locations ............................................. 139
6.1.3 Behaviour Over Time ......................................... 141
6.1.4 Complex Behaviour ........................................... 143
6.2 Situations and Situation Conditions .......................... 147
6.3 Information Channels ............................................ 157
6.4 Channel Conditions and Conditional Channels .............. 166
6.5 Perspectives on the World ...................................... 170

A Classification Systems .................................................. 178
A.1 Homogeneous Classification Systems .......................... 179
  A.1.1 Taxonomies ................................................... 179
  A.1.2 State Systems ............................................... 181
  A.1.3 Feature Systems ............................................ 184
  A.1.4 Attribute-Value Systems ................................... 186
A.2 Heterogeneous Classification Systems ......................... 192
  A.2.1 Attribute-Value Representations .......................... 192
  A.2.2 Type-Feature Systems ..................................... 194
  A.2.3 Feature Based Recognition ................................ 196
A.3 Classifying Words .................................................. 201

Bibliography .............................................................. 208
Introduction

This thesis is concerned with the elucidation of the structure of three basic cognitive functions.

- An organism must be able to make distinctions between different aspects of its environment if it is to respond selectively. This is classification.

- It must be able to anticipate conditions in other parts of its environment. If an organism at $x$ is to anticipate that the condition $c$ holds at $y$ then, at the very least, the information that $c$ holds must be accessible from $x$. Hence anticipation depends on a flow of information.

- An organism must be able to recognize uniformities across different parts of the environment. This is individuation.

We propose that each of these functions can be understood in terms of a primitive ability of 'seeing' the world from a perspective. In contrast to theories which suppose the existence of a conceptual scheme, or mastery of a language of thought, a characteristic of an organism's ability to adopt perspectives is the additional ability to shift from one perspective to another.

In the thesis we first propose a theory of classification. Its usefulness in categorizing different classificatory systems, like taxonomies, state systems and attribute-value structures, is demonstrated in the Appendix.

We then study two approaches to characterizing the flow of information. One, due to Dretske (1981), is based on conditional probabilities. The other, due to Barwise and Perry (1983) is based on the Situation Theoretic idea of a constraint.
Our theory of perspectives takes ideas from both accounts: from Situation Theory, the distinction between information supported and information carried by a situation, and from Dretske, the implicit relativity to an information channel.

We give a rudimentary account of the individuation of objects as predictive regularities across situations. Properties of objects individuated in this way are characterized as shifts in perspective which preserve the predictive regularity.

Finally, we consider a more concrete model of information flow (called a world system, Rosenschein (1989)) in which environmental conditions are understood in terms of possible state distributions over locations and times. We generalize his model and show how information channels offer a more sensitive account of information flow than the one induced by the global notion of possibility. Information channels are then used to construct perspectives within a world system.

In Chapter 1 we formulate a theory of classifications in fairly general terms. Given a fixed collection $D$ of unstructured objects (called tokens) to be classified, we define a class of classifications of $D$ called $D$-classifiers. A $D$-classifier simply consists of a collection of objects (called types) together with a 'classification' relation between tokens and types. We note some elementary properties of $D$-classifiers and some basic constructions. A wide range of examples are discussed in the Appendix.

Chapter 2 concentrates on classifications of situations. The ontological status of various Situation Theoretic entities is reviewed and their dependence on a scheme of individuation is noted. Special attention is given to basic infons which are the types of the classifier given by a scheme. A brief review of Situation Theory ensues in which the framework of types and propositions is regarded as providing a 'theorist's classification' of scheme-individuated entities like situations, infons and properties.

In Chapter 3 we ask how the classification of one part of the world can carry information about other parts. First we review the Communication Theoretic idea of the flow of information in probabilistic systems. Dretske's definition of information content is presented as well as his distinction between analog and digital modes of carrying information. Next we review the status of a Situation Theoretic view of constraints and especially the notion of a conditional constraint. We say how constraints provide an account of what information is carried by a situation and compare this with Dretske's
definition. Our conclusion is that they are internally equivalent: the only difference being the implicit relativity of Dretske's definition to a communication channel.

We then ask how the flow of information can be used by an organism to perform other cognitive functions like inference and individuation. The Situation Theoretic idea of attunement is defended from an objection by means of an extended example. Finally, we pursue a suggestion of Devlin's that Dretske's analog/digital distinction can be used to give an account of individuation and the related claim of Dretske's that the same distinction partially explains the difference between perception and cognition. We reject both suggestions as they stand, but note that a proper account of individuation might be forthcoming if the implicit relativity of Dretske's account could be more fully understood.

In Chapter 4 we present a theory of perspectives. A perspective consists of some parts of the world seen from a certain point of view. The theory brings together the idea of an abstract classification from Chapter 1 and the idea of a constraint's licensing information flow from Chapter 2. But the information flow is only local to the domain being classified by the perspective, so something of the conditionality of conditional constraints or the relativity of information flow to a channel is incorporated. Different kinds of perspective are identified according to the strength of the connection between their informational and classification roles and we define a comparison between perspectives called a shift. We then discuss the connection with conditional constraints, suggesting the 'background' parameter should really be referring to a perspective rather than to a type or to a situation.

Next, two-device and many-device communication systems are modelled probabilistically and are shown to define perspectives. These models produce perspectives which satisfy certain criteria precisely when information is flowing in the system. For contrast we consider perspectives defined from the notion of logical consequence within standard semantic interpretations of classical logics. According to the previous criteria there is no information flow within these 'logical' perspectives. Finally we model various modal logics as perspectives and make a connection between the accessibility relations of modal semantics and the direction of information flow.
In Chapter 5, we return to re-consider the prospect of defining individuals as uniformities across situations. The individuation of objects is characterized relative to a perspective. The intuitive motivation for this is that a sequence of events seen from one perspective may appear to be a coherent persisting individual, but, seen from another perspective they might make no sense at all. We define a notion of predictive regularity within a perspective and define an object to be a sequence of situations that gives rise to a predictive regularity. Properties are identified with shifts between perspectives and the condition under which an object has a property within a perspective is also defined. Finally, within a collection of perspectives, called a perspectival domain the notion of a basic info is revisited and the logical consequences of this notion in its new setting are explored.

In Chapter 6 we discuss a more concrete model of the flow of information used by Rosenschein (1987, 1989) to provide a non-symbolic account of information processing by robots. His model is based on primitive notions of location, time and state rather than object or property. The model is used to develop a methodology for robot design called 'situated-automata theory'. The basic idea is that a location's being in a certain state at a certain time constrains the possible states that other locations may be in, or, more globally, which of a set of possible worlds the actual world is. Content can be ascribed to the location by interpreting any one of a number of epistemic logics using the appropriate Kripke style semantics.

We generalize Rosenschein's 'world system' model to provide a notion of state and information content for arbitrary aggregates of locations which may be moving with respect to the reference frame over some time period. Rosenschein's definition of information content generalizes with it. We define a classification of parts of the world (situations) into types. The notion of a situation condition is introduced as a Situation Theoretic version of Rosenschein's 'world conditions' (relations between worlds and times). Satisfaction of a situation condition is relative to a 'point of view situation'. Next we model information channels which are connections between situations along which information can flow. The information channel approach is contrasted with Rosenschein's 'possible worlds' account of implication. Finally, we show that perspectives can be constructed in our generalization of Rosenschein's model. Our example is that of straightforward
INTRODUCTION

visual perspective.

In the Appendix we return to the subject of classification. D-classifiers are used as to form various kinds of classification system including taxonomies, state systems, and various feature based systems. These example are studied independently at first and then related to each other. We show how to relate taxonomic systems to feature systems, form attribute-value representations of feature systems and state systems, and define principles of inheritance for feature-based recognition systems. Finally, we consider a specific example: syntactic dependency in linguistics. We model dependency grammars as classification systems.
Chapter 1

Classifications

The ability to classify is an essential cognitive skill. Without it an organism would be no better off in the world than a lump of rock: selective response to an environment would be impossible. Somehow, a difference must be made between one aspect of the environment and another. By 'the ability to classify' we mean no more than the capacity to make such distinctions.

If we are to understand this fundamental capacity we must first understand what is achieved when classification occurs. What is a classification? In this chapter we follow certain proposals towards a theory of classifications. We suppose, firstly, that a classification has something to classify: that there is some domain which has parts which are different from other parts. The parts of the domain being classified we call 'tokens'. Tokens are taken to be completely unstructured particulars with only one characterizing property: they are different from other tokens.\(^1\)

Given a domain of tokens to be classified, we suppose that a classification of the domain amounts to the making of distinctions between the tokens. If a token is classified then, at most, it is distinguished from some other tokens. It need not be distinguished from all other tokens, nor need it be distinguished from any tokens at all. We capture this idea

\(^1\)Some defence of this 'atomistic' viewpoint may be thought necessary. It might be objected that the environment in which organisms live is essentially undifferentiated. There are no primitive divisions of the world into convenient cognitive sized boxes. This is probably true, but it does not argue against our approach. To claim that the world is 'undifferentiated' is not to claim that it has no parts. We might say that empty space is undifferentiated, but by that we do not mean that it is a single indivisible point. We just mean that there is no difference between one part of empty space and another.
by saying that a classification consists in the gathering together of tokens into types. Tokens of a type are distinguished from other tokens, namely those which are not of that type.

Tokens are only individuated by their not being other tokens. Types are only individuated by their classificatory role: the way in which they gather together tokens. It is this notion of 'classificatory role' which we are attempting to capture in our theory of classifications. Unlike token-identity, type-identity is intrinsically bound up with the particular kind of classification being used.

In what follows we present a theory of classifications in which minimal assumptions are made about the classificatory role of types. The general notion of classification is linked with other concepts like similarity and indistinguishability and the question of when two classifications are the same, or equivalent, is addressed. Under a strong notion of equivalence (isomorphism) we study the algebraic properties of classifications.

We start by fixing a domain $D$ of tokens to be classified. A classification of $D$ is given by a class of objects, called types, and a binary relation between tokens and types.

**Definition 1.1** Given a collection of objects, $D$, which we call the token domain, a $D$-classifier, $T$, consists of a collection, $|T|$, of types and a relation, $\gamma$ between elements of $D$ and elements of $T$. For $d \in D$ and $t \in |T|$, we say that $d$ is classified by or is of type $t$ iff $d : t$ (dropping the subscript in context).
CHAPTER 1. CLASSIFICATIONS

Each element \(d\) of \(D\) can be classified by many or no elements of \(|T|\). If \(d : t_1\) then \(d\) is classified as being of type \(t_1\). If, in addition, \(d : t_2\) then \(d\) is classified as being both of type \(t_1\) and of type \(t_2\). The total \(T\)-classification of \(d\) is given by its type class, \(d = \{t \in |T| \mid d : t\}\). If \(d = \emptyset\) then \(d\) is said to be unclassified by \(T\).

There is no need for \(T\) to classify all of \(D\). The collection of tokens that are classified by \(T\) is called the domain of \(T\), written \(D_T\). For any particular type \(t \in T\), the collection of tokens that it classifies is called its token class (or extension), \(t = \{d \in D \mid d : t\}\).

In general, the collections of types and tokens involved in the above definitions are classes rather than sets. Classifiers which limit these collections to being sets have some special properties and so we make the following definitions.

**Definition 1.2** Let \(T\) be a \(D\)-classifier

1. \(T\) is set-based iff the token classes of every type are sets, i.e. \(\{t \mid t \in |T|\} \subseteq \text{pow } D\).
2. \(T\) is set-complete iff every set of \(D\) tokens is the token class of some type, i.e. \(\text{pow } D \subseteq \{t \mid t \in |T|\}\).
3. \(T\) is set-like iff for each \(a \in \text{pow } D\), \(\{t \in |T| \mid t = a\}\) is a set.

Whereas tokens are regarded as concrete objects, types are regarded as abstract objects. For this reason our theory is developed with respect to a particular class \(D\) of tokens.

The elements of \(D\) are fixed and unstructured. There is no sense in which we could replace one token by another and achieve a 'similar' classification, since we have no general notion of a 'classifier'. We have only defined the notion of a '\(D\)-classifier' for a particular \(D\). It should be noted, however, that any \(D\)-classifier is also an \(E\)-classifier for any collection \(E\) of \(D\) tokens.

Our attitude to types is very different. Anything, concrete or abstract, can be a type of a \(D\)-classifiers by being associated with a class of tokens. However, the constitution of a particular type is not always important to its role as a classifier of tokens. What is important about a type in a \(D\)-classifier is determined by its 'classificatory role'. We will not precisely define the notion of a classificatory role, but its sense should become clearer through the use we make of it in this chapter.

A useful way of thinking about the role of types in a classifier is as units of information about tokens of \(D\). The smaller the token class of a type, the more informative is
CHAPTER 1. CLASSIFICATIONS

the fact that a certain token is of that type. Conversely, the bigger the type class of a token, the more information we have about it. This observation gives us two information orderings, one on the types and one on the tokens:

\[ t_1 \prec t_2 \iff t_1 \subseteq t_2 \]

\[ d_1 \sqsubseteq d_2 \iff d_1 = d_2 \]

In other words, \( t_1 \prec t_2 \) if \( d : t_2 \) tells us more about \( d \) than \( d : t_1 \), and \( d_1 \sqsubseteq d_2 \) if we have more information about \( d_2 \) than about \( d_1 \).

Both these relations are pre-orders (reflexive and transitive) but not necessarily partial orders. We may have exactly the same information about two distinct tokens \( d_1 \) and \( d_2 \), i.e. \( d_1 = d_2 \). There is then no way of telling the tokens apart: the classification is incapable of distinguishing them. If this is the case then we say that \( d_1 \) and \( d_2 \) are indistinguishable and write \( d_1 \approx d_2 \).

\[ x \sqsubseteq y \text{ and } y \sqsubseteq x \text{ then } x \approx y \]

A classification also gives us a notion of similarity between tokens. If tokens \( d_1 \) and \( d_2 \) are of the same type (i.e. \( \mathcal{A} \cap \mathcal{B} \neq \emptyset \)) then they are similar, at least in that regard and we write \( d_1 \sim d_2 \). If \( d_1 \not\sim d_2 \) then we say that \( d_1 \) and \( d_2 \) are completely distinguishable.

Both \( \approx \) and \( \sim \) (and hence \( \neq \)) are symmetric. \( \approx \) is also reflexive and transitive (an equivalence relation), but \( \sim \) and \( \neq \) are, in general, neither reflexive nor irreflexive nor transitive.

Another interpretation of the notion of 'classificatory role' is given by the principle of extensionality:

EXT For \( t_1, t_2 \in \mathcal{T} \), if \( t_1 = t_2 \) then \( t_1 = t_2 \).

A D-classifier is said to be extensional if it satisfies the condition EXT. However, we do not restrict our attention to extensional D-classifiers as there may be more to the classificatory role of a type in a D-classifier than its extension. Instead, we arrive at a broader conception of 'classificatory role' indirectly, by considering morphisms of various kinds between D-classifiers.
CHAPTER 1. CLASSIFICATIONS

Figure 1.2: Degrees of similarity inside a $D$-classifier

Definition 1.3 A map $f : |T_1| \rightarrow |T_2|$ is a $D$-classifier morphism between $D$-classifiers $T_1$ and $T_2$ iff for all $d \in D$, $t \in |T_1|$, $d \sim t$ iff $d \sim f(t)$.

The classificatory role of a type is captured by its image under morphisms in some class of morphisms. Different classes of morphisms give different kinds of classification, each with a corresponding notion of classificatory role. We will see plenty of examples of this in later sections of this chapter.

Proposition 1.4 $T$ is extensional iff there is only one morphism from $T$ to $T$.

Proof: If $T$ is extensional and $f$ is a morphism from $T$ to $T$ then for all $t \in |T|$, $ft = t$ so, by EXT, $t = ft$, i.e. $f$ is the identity morphism on $T$. Conversely, if $T$ is not extensional then there are distinct types $t_1, t_2 \in |T|$ such that $t_1 \neq t_2$. Let $f : |T| \rightarrow |T|$ be defined by $ft = t(t \neq t_1)$ and $ft_1 = t_2$. Then $f$ is a morphism from $T$ to $T$ which is distinct from the identity morphism. Hence $T$ has at least two morphisms.

One class of morphisms of interest is the class of bijective morphisms, or isomorphisms. We say that $D$-classifiers $T_1$ and $T_2$ are isomorphic iff there is an isomorphism from $T_1$ to $T_2$. This gives us a very strong equivalence relation on the class of $D$-classifiers.

For the remainder of this section we shall only be concerned with the properties of $D$-classifiers up to isomorphism; i.e. with those properties that do not distinguish between isomorphic $D$-classifiers.

Associated with the notion of isomorphism are notions of embedding, substructure and restriction which are defined as follows.
CHAPTER 1. CLASSIFICATIONS

Definition 1.5 Embedding, subclassifier and restriction. \( T_1 \) can be embedded in \( T_2 \), written \( T_1 \subseteq T_2 \), iff there is an injection \( f: |T_1| \rightarrow |T_2| \) such that for all \( d \in D, \ t \in |T_1|, \ d :\tau_1 t \iff f(d) :\tau_2 f(t) \). If, in addition, \( |T_1| \subseteq |T_2| \) then \( T_1 \) is a subclassifier of \( T_2 \), written \( T_1 \triangleleft T_2 \). If \( T \) is a \( D \)-classifier then the subclassifier of \( T \) with types \( X \subseteq |T| \) is called the restriction of \( T \) to \( X \), written \( T \upharpoonright X \).

If a \( D \)-classifier, \( T_1 \), is embeddable in another \( D \)-classifier, \( T_2 \), then there is a sense in which \( T_1 \) is represented inside \( T_2 \): all the distinctions that can be made with \( T_1 \) can also be made with \( T_2 \). A natural question that arises is as to whether, given a class \( X \) of \( D \)-classifiers, there is a single \( D \)-classifier in which all \( T \in X \) can be embedded. For set-based extensional \( D \)-classifiers we can build a 'canonical' \( D \)-classifier \( PD \) which has this property. \( PD \) is constructed by taking sets of tokens as types with the membership relation as classification, i.e. \( |PD| = pow D \) and \( d :PD t \iff d \in t \).

Proposition 1.6 Every set-based, extensional \( D \)-classifier \( T \) can be embedded in \( PD \). If \( T \) is also set-complete then it is isomorphic to \( PD \).

The following are two constructions of classifiers which will be used later in this appendix.

Definition 1.7 Given a \( D \)-classifier, \( T \), and a set \( X \) of subsets of \( |T| \) we define the result of lifting \( T \) by the set \( X \) to be the \( D \)-classifier \( T \upharpoonright X \) where \( |T \upharpoonright X| = X \) and, for \( d \in D, x \in X, \ d :X x \iff x \subseteq d^T \).

Given a \( D_1 \)-classifier, \( T_1 \), and an \( D_2 \)-classifier, \( T_2 \), we define the product of \( T_1 \) and \( T_2 \) to be the \( D_1 \times D_2 \)-classifier \( T_1 \times T_2 \) where \(|T_1 \times T_2| = |T_1| \times |T_2| \) and, for \( d_1 \in D_1, d_2 \in D_2, t_1 \in |T_1| \) and \( t_2 \in |T_2|, \ (d_1, d_2) :T_1 \times T_2 (t_1, t_2) \iff d_1 :T_1 t_1 \) and \( d_2 :T_2 t_2 \).

For the remainder of this section we will study the algebraic properties of set-based and set-like \( D \)-classifiers.

Definition 1.8 A \( D \)-classifier \( T \) is finite (countable) iff it is set-based and set-like and for each \( a \in pow D, \ \{ t \in |T| \mid t = a \} \) is finite (countable).

Let \( \mathcal{D}_f \) be the class of isomorphism classes of finite \( D \)-classifiers. We write the isomorphism class of \( T \) as \( [T] \). We define an ordering \( \leq \) on \( \mathcal{D}_f \) by
We would like to determine the characteristic properties of \((\mathcal{D}_\omega, \leq)\), but this structure is difficult to study directly. Instead we construct an order isomorphic structure whose properties are easier to inspect.

Let \(\mathcal{M}_\omega\) be the class of class operators from \(\text{pow}D\) to \(\omega\). Elements of \(\mathcal{M}_\omega\) are, in effect, assignments of numbers to sets of \(D\) tokens. Let \(\leq\) be the pointwise ordering on \(\mathcal{M}_\omega\), vis.

\[
f \leq g \quad \text{iff} \quad \forall a \in \text{pow}D, \quad fa \leq ga
\]

**Proposition 1.9** (\(\mathcal{D}_\omega, \leq\)) is order isomorphic to \((\mathcal{M}_\omega, \leq)\).

**Proof:** For each finite \(D\)-classifier, \(T\), define \(\mu_T \in \mathcal{M}_\omega\) by

\[
\mu_T a = \|\{ f \in [T] \mid f = a\}\|
\]

where, for set \(x\), \(\|x\|\) is the cardinal of \(x\). The operator \(\mu_T\) assigns to each set \(a\) of \(D\) tokens, the number of types which classify that set, i.e. the number of types which have \(a\) as their token class. For each \(f \in \mathcal{M}_\omega\) define the finite \(D\)-classifier, \(\gamma_f\), by

\[
|\gamma_f| = \bigcup_{a \in \text{pow}D} \{(a,n) \mid 0 < n \leq fa\}
\]

for each \(d \in D\), \(d \gamma_f (a,n) \quad \text{iff} \quad d \in a\).

Now let \(\mu : \mathcal{D}_\omega \rightarrow \mathcal{M}_\omega\) take \([T]\) to \(\mu_T\) (which is well-defined) and let \(\gamma : \mathcal{M}_\omega \rightarrow \mathcal{D}_\omega\) take \(f\) to \([\gamma_f]\). We claim that \(\mu\) and \(\gamma\) are inverses and therefore bijections. In addition

\[
\mu T \leq \mu T' \quad \text{iff} \quad T \leq T'
\]

and so \(\mu\) is an order isomorphism between \(\mathcal{D}_\omega\) and \(\mathcal{M}_\omega\).
CHAPTER 1. CLASSIFICATIONS

Note that the proof gives us a way of constructing a canonical representative of the isomorphism class of any finite $D$-classifier, $T$, namely $\gamma_T$.

Proposition 1.10 For $(\mathcal{M}_\omega, \leq)$ (and so $(\mathcal{D}_\omega, \leq)$),

1. $\leq$ is a partial order
2. $\leq$ is a distributive lattice
3. $\leq$ has a smallest element

Proof: Each property follows from the pointwise ordering of $\mathcal{M}_\omega$ and the respective property of the natural numbers. In particular, for $f_1, f_2 \in \mathcal{M}_\omega$, $f_1 \wedge f_2$ and $f_1 \vee f_2$ are given by

\[ f_1 \wedge f_2(a) = \min(f_1(a), f_2(a)) \]
\[ f_1 \vee f_2(a) = \max(f_1(a), f_2(a)) \]

The smallest element of $\leq$ is the operator which assigns 0 to each set of $D$ tokens.

Note that the structure is not a complete lattice since there is no biggest element. However, the class $\mathcal{D}_{\omega^+}$ of isomorphism classes of countable $D$-classifiers does form a complete lattice under $\leq$. $(\mathcal{D}_{\omega^+}, \leq)$ can be shown to be order isomorphic to $(\mathcal{M}_{\omega^+}, \leq)$ where $\mathcal{M}_{\omega^+}$ is the class of operators from $\text{pow}D$ to $\omega \cup \{\omega\}$. The biggest element of $\mathcal{M}_{\omega^+}$ is just the operator with constant value $\omega$.

Despite $(\mathcal{M}_{\omega^+}, \leq)$ being a complete lattice, there is no obvious choice for a complement operation to make it a boolean algebra. However, the substructures $(\mathcal{M}_n, \leq)$ where $\mathcal{M}_n$ is the class of $\{0, 1, \ldots, n-1\}$-valued operators on $\text{pow}D$ are complete distributive lattices and have natural complement operations, taking $f \in \mathcal{M}_n$ to $\neg f \in \mathcal{M}_n$ given by $\neg f(a) = (n - 1) - f(a)$, which turns the structures into boolean algebras. The structure $(\mathcal{M}_\omega, \leq)$ is order isomorphic to $(\mathcal{D}_{\text{ext}}, \leq)$, where $\mathcal{D}_{\text{ext}}$ is the class of isomorphism classes of extensional $D$-classifiers.

The exact structure of classes of uncountable $D$-classifiers is dependent on what is said about the existence of various infinite cardinals.

An extensive range of examples of $D$-classifiers is provided in the Appendix.
Chapter 2

Situations

The theory of classifications presented in Chapter 1 is intended to be fairly general. Nothing is said about the token domain except that it consists of distinct tokens. Our examples (in the Appendix) concern the classification of physical and biological objects, like leaves and hedgehogs, closed physical systems like a container of fluid, events like the descent of a meteor, abstract objects like triangles, linguistic objects like words and even syntactic structures. A complete account of cognition should certainly explain how it is possible to interact with and manipulate all of these kinds of entity. However, we will be primarily concerned with an organism’s ability to classify its immediate sensory environment; that is, with perception.

Of course, what is meant by an ‘environment’ is far from clear. For this we turn to Situation Theory as conceived and developed by Barwise and Perry. One of the original applications of Situation Theory (henceforth ST) was to provide a semantic theory of perception reports (Barwise 1981) and it is in this application that the notion of a ‘situation’ is most intuitively presented.

"When I look around I cannot see a single thing-in-itself, some sort of ideal physical object stripped of its properties and relations with other objects. What I do see is a scene, a complex of objects having properties and bearing relations to one another. The properties and relations are every bit as important to what I see as the idealized thing-in-itself. In fact, what really counts is the whole complex of objects-having-properties-and-bearing-relations which constitutes the scene. ... Scenes are visually per-
CHAPTER 2. SITUATIONS

The ontology implicit in traditional model theory is that of a world consisting of individual objects having properties and bearing relations to each other. When applied to mathematical languages, this ontology is intuitive and precise as, in general, mathematical discourse concerns a specific domain of abstract objects, be they groups, fields, numbers or whatever. It is even applicable across mathematical domains as the success of set theory in providing a foundation for mathematics has shown. Model theory is therefore a good generalization of how mathematical structures are classified.

The perceptual world, however, has at least one difference from mathematical worlds which make it difficult to adopt a similar ontology: it is essentially partial. An organism never perceives the world, only part of it. This is because the perceiving is being done by something (the organism) which is firmly embedded in the world it perceives. Nevertheless, there is at least a possibility that the world as a whole can be modeled as consisting of objects having properties and bearing relations to each other in such a way that the parts of the world which are perceived have some direct relationship to the whole. Partial model theory (for example Langholm 1988) can be seen as a way of explaining this relationship. A part of the world is seen as consisting of objects having properties and bearing relations to each other; it is just that some of the relations and properties may not be determined for the objects of the part's domain.

We take ST to be an alternative response to the problem of partiality. Instead of individuating a part of the world by the objects it contains and the properties they have in it, ST introduces a new ontological category: the situation. It is situations that we perceive rather than objects or even collections of objects, properties and relations. This position leaves us with certain fundamental questions to be answered as to the relationship between situations, properties, relations and objects.

Traditional model theory has a precise conception of the relationship between relations,

1 Barwise 1981, p.390
2 Although the nature of the set theoretic universe seems indefinitely extendible with new cardinality axioms. Moreover the tension between constructive and classical mathematics and the desire to understand structures which are not intuitively conceived of as sets would suggest a more pluralist attitude. A more modern outlook would speak of different universes of mathematical objects. See, for example, Aczel 1989 and Barwise 1989 ch 14.
CHAPTER 2. SITUATIONS

properties and objects. Relations have a number associated with them called their arity: properties are relations with arity 1. Given n objects in sequence, \(a_1, \ldots, a_n\), and an n-ary relation, \(r\), it is determinately true or false as to whether or not \(a_1, \ldots, a_n\) stand in the relation \(r\). This is justified, from a ST point of view, by re-interpreting the concerns of (partial) model theory as being about a single situation. When we admit more than one situation the global notion of truth is no longer sufficient.

To resolve this difficulty ST introduces a further ontological category: the fact. Loosely speaking, given a relation and an appropriate sequence of objects, a situation can determine one of two facts: either the positive fact that the objects stand in the relation or the negative fact that they do not stand in the relation. Either way, the situation is said to support the fact so determined. However, the situation may fail to support either fact.

The logical content of this proposal (at the level of facts) is similar to that of partial model theory. In fact, we could capture certain aspects of situations using partial models: given a situation, each n-ary relation has a positive and a negative extension in the set of sequences of n objects. Facts could then be modeled as tuples \((r, a_1, \ldots, a_n, +/\)−) which are supported by a partial model just in case \(a_1, \ldots, a_n\) is in the \(+/\)−ve extension of \(r\). What is missing from this approach is an explanation of how situations, facts and relations can themselves have properties and stand in relations with each other, i.e. that which makes them existing objects in their own right.

Hidden behind this preliminary story there are a host of questions about situations, facts and the relation of supporting which have to be answered. For example, is every fact supported by some situation? Is there anything more to a situation than the set of facts that it supports? Is the fact determined by a relation and a sequence of objects independent of the situation? Can the same fact be determined by different relations and objects?

ST sets out to provide a coherent (and hopefully correct) answer to these kinds of questions. In doing so it introduces further ontological categories, such as types and propositions as well as certain other categories such as assignments, argument roles, parameters and anchors, whose ontological status is unclear. In this chapter we will

---

3Essentially, this is the proposal of van Benthem in Fenstad et al. (1987).
present a brief overview of the theory with little attention paid to formal details. Our account will not be even partially definitive, let alone complete, since there are many points ('branch points' in Barwise's terminology) to which different proponents of ST have responded differently. We are most sympathetic to the attitude towards ST presented in Barwise 1989b which is closer to one end of the spectrum of attitudes than representative of the middle. Hopefully, at least some significant proportion of what is said here overlaps with what most proponents of the theory take the consensus to be!

Our account is based directly on Barwise's (1987 and 1989c). A model (in non-well-founded set theory) for much of the theory presented here has been proposed by Fernando (1989) and a first order axiomatization has been developed by Westerstål (1989).

2.1 Ontological Groundwork

In this section we consider some of the basic ontological issues which must be decided by, or at least addressed by, a full theory of situations. Firstly, there are a number of questions about the purely objective claims that ST makes about the world. These are issues which are independent of any classification. Secondly, we consider the status of those entities whose existence or individuation is somehow relative to a 'scheme of individuation'. Thirdly, we discuss the ontological status of those entities which are derivative on a scheme of individuation or ancillary in some other respect.

2.1.1 Parts of the Worlds

*The world is made up of parts.* ST supposes that there is an objective world which is not a single indivisible entity, but which has parts which possess objective identity. We call these *parts-of-the-world*. Our assumptions about the token domain of a classification are similar in nature to ST's assumption about the world. In fact we will apply the notion of a classification to classify parts of the world. A difference is that ST accepts additional ontological categories than 'parts-of-the-world' like objects, relations, facts and so on. These too can be tokens of a classifier. However, ST makes a further commitment to the objective structure of the world, namely that *the parts-of-the-world are partially ordered*. The symbol \( \sqsubseteq \) is used to denote the 'part of' relation on parts-of-the-world.
Apart from the claim that $\mathcal{D}$ is a partial order, there is little consensus as to other properties it might have. We discuss some possibilities below.

1. *There is a maximal part-of-the-world.* This is fairly uncontroversial. The existence of a maximal part-of-the-world is necessary if there is to be a designated actual world which is our world, this world. It seems intuitive that the actual world is maximal. Other maximal parts-of-the-world are also called, somewhat confusingly, *worlds.* Perhaps the term 'part-of-reality' should be preferred to 'part-of-the-world'; but it is best not to make too many ontological distinctions as there are plenty already. One simplifying assumption which would make this distinction unnecessary is

2. *There is a greatest part-of-the-world.* This is controversial. The assumption that there is a (necessarily unique) greatest part-of-the-world would make things simpler since then every part-of-the-world would be just that - a part of the world (the actual world). However, the existence of parts-of-the-world which are not part of this world but nevertheless are somehow in this world (as objects of our thoughts, the reference for fictional discourse, a realist basis for notions of probability and necessity and so on) has its supporters, including Barwise. If there is no greatest world, then another possibility contra 1. and 2. is

3. *There are no maximal parts-of-the-world.* To me this is an attractive possibility, although I believe no one else shares this view. On this assumption the (one and only) world would not be a part-of-the-world at all. Modal distinctions between parts-of-the-world would be explained in terms of the way parts-of-the-world are classified rather than by a realist notion of distinct existing possibilities.

4. *Every part-of-the-world is part of a maximal part-of-the-world.* This assumption would seem to go along with 1. and the negation of 2. (It is of course implied by 2. and inconsistent with 3.) On this view the 'world' would consist of a collection of 'possible' worlds (i.e. maximal parts), one of which is actual. As Barwise points out in his (1989c) this conception of a 'possible world' need not be that similar to

---

4 Though in some applications $\mathcal{D}$ may be just a pre-order. See Chapter 6.

5 There are arguments for having many different actual worlds. See Chapter 6.
Lewis's possible worlds since they could be very small (e.g., an isolated point in the $\aleph$ order would be a world). A related assumption is

5. Every part-of-the-world is part of precisely one maximal part-of-the-world. On this assumption the possible worlds would be entirely distinct—pure modal alternatives in the sense of Lewis (1973).

6. $\mathfrak{P}$ is directed, i.e., for any two parts-of-the-world there is a part-of-the-world of which both are part. This is not the popular view, mainly because it is inconsistent with the $\mathfrak{P}$-$\mathfrak{Q}$-$\mathfrak{O}$-$\mathfrak{M}$ combination. It is implied by 2, 3, and 5.

7. As above, reversing $\mathfrak{P}$. Questions about minimal parts-of-the-world, the existence of minimal parts of every part-of-the-world, lower bounds and so on, have not attracted as much attention as their dual counterparts. Nevertheless, we mention them here because of their relevance to Chapter 6, where they play a part in providing conditions on the construction of worlds.

In summary, not much is known about worlds and their parts. One reason for this is that we only interact with the world via a scheme of individuation.

2.1.2 Schemes of Individuation

The notion of a scheme of individuation is recent but its role seems central to ST. In brief, it is a mode of interaction with the world on which are founded all the central ontological categories of ST: situations, objects, properties, relations and facts. Part of the problem is that the applications of ST are concerned with different kinds of 'agent' doing the interacting. Barwise (1989c) lists eight potential applications for ST: frogs, vision, data bases, robot design, mathematics, Japanese speakers, cognitive science and situation theorists. Each application is concerned with a different kind of interaction with the world. The way we study a frog's individuation of the world may bear little relation to the way we study the common linguistic individuation of a community of Japanese speakers. Our concern is mainly for frogs and vision although with regard to the extensions needed for cognitive science in general. With this in mind we will extract some of the consensual attitudes to schemes of individuation.
CHAPTER 2. SITUATIONS

An organism's primary connection with the world is its being in the world. Being in the world, an organism is in perceptual contact with certain parts-of-the-world and not with others. If it is to be active in the world, and certainly if it is to be cognitive, an organism must classify parts-of-the-world that it comes into contact with. We use 'classify' in precisely the same sense as we have used it in Chapter 1. In other words, the organism must perceive 'differences' in its environment. The organism's environment may include parts-of-the-world that are remote from it, so we are using 'contact' in a very broad sense. Moving our attention from frogs to humans, we might say that a human can classify parts-of-the-world which are very remote in both space and time (Beijing, Socrates and the Andromeda Galaxy, for example) with her thoughts.

The parts-of-the-world that an organism classifies are called the situations of the organism's scheme of individuation, which at base is a parts-of-the-world-classifier. The types of this classifier are called infons. If a situation \( s \) is classified by \( \sigma \) in scheme \( I \), then \( s \) is said to support \( \sigma \) in \( I \), written

\[
\models_I \sigma
\]

The \( \models \) relation is one of two ways in which information can be located at a situation. The other will be studied in Chapter 3. The notion of a 'fact' is defined in terms of this classification: a fact is an infon which is supported by some actual situation (a part of the actual world).

In general, certain parts-of-the-world remain unclassified by a scheme. These parts would then fail to be situations. Barwise illustrates the distinction between parts-of-the-world and situations in the following analogy:

"Let us imagine a child's toy, a simple modification of toys on the market. Imagine a large sheet of paper with a scene drawn in red, blue and purple. This scene will be our analogue for the world. The toy comes with two tubes that the players are required to use when examining the scene. One tube has a piece of red cellophane on the far end, the other tube has blue. Think of these tubes as rival cognitive schemes of individuation. When the child
looks at the scene with the red tube, she sees only the purple and blue parts of the scene. When she examines the scene with the other tube, she sees the red and purple. Now fix one of these tubes, say the red one. We suppose that the child must stand in a fixed place in front of the scene, one close enough that only a small circular portion of the scene is visible through the tube at any one time. These circular portions, with their blue and purple lines, play the role of the child's (focus) situations in this analogy. They are there, independent of the child, but it is the tube(s) and cellophane that make them the (relevant) situations to consider, and not some other class.\footnote{Barwise 1989 ch 10 p232. Our parenthessen.}

The notion of a scheme of individuation gives the category of situations just the right ontological status: both objective yet relative to a classification. It allows the possibility that different schemes can be classifying the same parts-of-the-world even if they do so in different ways (the purple lines are visible through both tubes). It is this notion of 'objective relativity' which we will explore farther in Chapters 4 onward.

A possibility that is opened by the distinction between parts of the world and situations is that no \textit{maximal} parts (if they exist) are classified by a certain scheme of individuation. If we are concerned with frogs and vision then this possibility seems eminently reasonable. Barwise and Etchemendy (1987) have used the gap between situations and worlds to provide a solution to various semantic paradoxes and Barwise (1989 ch 14) uses the same distinction to suggest an explanation for Skolem's 'paradoxes' in set theory. So for the concerns of mathematics it also seems reasonable. Situation theorists too are likely to require this distinction if ST is to avoid similar paradoxes. For cognitive science in general, it may be less reasonable. After all, the previous section was all about maximal parts-of-the-world. Perhaps there is a proof here for position 3!}

The terminology of Chapter 1 will be applied to a scheme's classification of situations. For example, situations are ordered by the information ordering \( \subseteq \) and when \( s \subseteq s' \) and \( s' \subseteq s \), \( s \) will be said to be (informationally) indistinguishable from \( s' \). The infon set \( \mathcal{S} \) of a situation has significance in ST as models of the theory tend to identify \( s \) with \( \mathcal{S} \). Here we distinguish them, but sometimes refer to sets of infons like \( \mathcal{S} \) as 'abstract situations' (in line with Barwise and Perry 1983). The early approaches to ST tended to assume
CHAPTER 2. SITUATIONS

that if \( s_1 \preceq s_2 \) then \( s_1 \subseteq s_2 \): the persistence of information. In more recent accounts (Barwise 1989c), this assumption is called into question, and 'persistence' is defined for particular infons. An infon is said to be persistent if, whenever it is supported by a situation \( s \), it is also supported by every situation which \( s \) is part of.

Once classified, situations provide the 'raw material' for further classification. Barwise and Perry (1983) consider objects, properties, relations, locations and facts to be uniformities across situations. An organism individuates these things in a mutually constrained way in order to make sense of the 'undifferentiated kapok' (as Perry has called the stuff of the world).\(^8\)

The reason why situations are "metaphysically prior to facts" which are "metaphysically on a par with objects, properties and relations" (Barwise 1989 ch 10 p. 232) is a subtle one. Why could it not be the case that situations are extracted from the 'kapok' together with the rest in a 'mutually constrained' way? The answer is that there is an asymmetric ontological dependency between facts (say) and situations. Situations exist independently of any scheme of individuation since they are just parts-of-the-world. All a scheme of individuation does is to 'highlight' certain parts and not others. Facts, on the other hand, have no existence independent from a scheme of individuation: they have no correlate in the kapok.

As an analogy, imagine the image of an arrow projected onto the wall of a dark room. The light illuminates parts of the wall, which give existence to the arrow as a uniformity across the dark surface. If we turn off the projector then the arrow disappears. The parts of the wall that constitutes the arrow are still there; they are just no longer distinguished from other parts of the wall. But there is no way we can keep the arrow and not illuminate the part of the wall that it is a uniformity across.

It is relatively easy to see how infons are 'uniformities' across situations. A particular infon will in general appear in the infon set of many situations, so that the notion of 'uniformity' for infons is given by what we have been calling its 'classificatory role'. It is less easy to see what kind of 'uniformity' objects and relations are. There

\(^8\)Saying what 'makes sense' is not part of the ontological considerations we are engaged in. Any division of the world into objects and relations would do, but for computational and biological reasons some are possible (good, even) whilst others are not.
are two approaches to clarifying this issue. Either we can study the 'logical' properties of infons, objects and relations, and try to deduce the relationship between objects and the situations they are uniformities across from this. Or we can study the 'genesis' of objects in terms of some notion of 'uniformity' and try to use this to deduce their logical properties. In this chapter we present some of the main aspects of ST that are relevant to the former project. Later, in Chapters 3 and 5, we will attempt to say something about the latter project.

2.2 Basic Infons

A scheme of individuation classifies parts of the world using 'basic infons'. It is this classification which captures an organism's primary understanding of its environment. Other, more complex, classifications are dependent on this one for their connection to the world. In this section we discuss what is meant by a 'basic infon' in ST, in preparation for a more formal treatment in Section 2.4. A number of the issues we discuss are also relevant to Section 2.3.

A positive basic infon, denoted $\langle r, A, + \rangle$, consists of a collection of objects $A$ standing in the relation $r$ to each other. A negative basic infon, denoted $\langle r, A, - \rangle$, consists of a collection of objects $A$ not standing in relation $r$ to each other. For example, if we consider the relation of increasing size then the elephants in the top row of Figure 2.1 do not stand in this relation whilst those in the bottom row do. Of course it does not matter whether or not the elephants are actually standing in a row: what counts is the order that we consider them in. This observation has motivated the almost universal practice of modelling relations as sets of sequences of objects. Usually a relation will have a fixed arity $n$ and so is modeled as a subset of $\Pi_{i \leq n} A_i$ for sorts $A_1, \ldots, A_n$. For example, $n$-ary relations between elephants are modeled as subsets of $\text{Elephants}^n$. In such a model a collection of objects $A$ stand in a relation $r$ just in case $(a_1, \ldots, a_n) \in r$ where $A = \{a_1, \ldots, a_n\}$.

2.2.1 Identity

The ordering of arguments resolves the ambiguity in the idea of a collection of objects standing in a relation, but it can be too restrictive. For example, if we name the first
two elephants in the top row of Figure 2.1 Elenor and Humphrey then we certainly have the information that Elenor is bigger than Humphrey. We may model this by having 
\[ \text{bigger-than}_1 \subseteq \text{Elephants}^2 \]
and
\[ \{(\text{bigger-than}_1, \text{Elenor}, \text{Humphrey}, +)\} \iff (\text{Elenor}, \text{Humphrey}) \in \text{bigger-than}_1. \]

Equally we could choose to have
\[ \{(\text{bigger-than}_2, \text{Elenor}, \text{Humphrey}, +)\} \iff (\text{Humphrey}, \text{Elenor}) \in \text{bigger-than}_2. \]

What is to choose between the two models? If we take English word order to determine the ordering of the elephants then we might say that the relation \((bigger-thane)_1\) models the relation we call `bigger than', whereas \((bigger-thane)_2\) models the relation we call `smaller than'. But if these relations really are converses of each other then it seems wrong to say that \((bigger-than_1, \text{Elenor}, \text{Humphrey}, +)\) and \((bigger-than_2, \text{Elenor}, \text{Humphrey}, +)\) are different pieces of information.

A similar point can be made with the single relation 'the same size as'. Whilst it is uncontroversial that elephant \(x\) is the same size as elephant \(y\) iff elephant \(y\) is the same size as elephant \(x\), is the information that Elenor is not the same size as Egbert different from the information that Egbert is not the same size as Elenor, i.e. is \((\text{the-same-size-as}, \text{Elenor, Egbert, -}) = (\text{the-same-size-as}, \text{Egbert, Elenor, -})\)? We would probably say that is was, but there are less clear cases about which intuitions vary. Although
the relation 'the same size as' may have its symmetry built in, does this extend to the
transitivity of 'bigger than'?

Whilst sometimes wanting to identify the information picked out by different English
sentences, we must be careful not to go to far. For example, if information is modeled
by sets of possible worlds the above problems are resolved since Elenor is the same size
as Egbert in the same possible worlds as Egbert is the same size as Elenor. Consequent-
ly, they are the same thought. Less fortunately, the model identifies all necessarily
equivalent thoughts. The thought that the water in the pan is boiling is taken to be
identical to the thought that the water is the pan is at 100° C. Moreover, the model
allows only one necessarily true thought, giving rise to the notorious problem of 'logical
omniscience': if you know one such truth you know them all. Mathematical knowledge
is the celebrated counter-example.

The problem re-occurs for any attempt to identify an infon with the set of situations
which support it. Although there may be no infons supported by all situations, there are
bound to be some which we would like to distinguish between, but which are supported
by the same situations. But what is the alternative? On the one hand we can individuate
infons by the syntax of the expressions we use to name them; and so separate infons we
wish to identify. On the other hand we can individuate them according the situations
that support them; and so identify infons we would like to keep separate. The two
extremes both seem inadequate.

Approaching the problem from the syntactic end, we note that linguistic theories have a
potential solution in the notion of thematic role. The use of thematic roles is most clear
for verbs of action. If Elenor kissed Humphrey we might say that Elenor and Humphrey
stood in the 'kissing' relation with Elenor adopting the agent role and Humphrey adopt-
ing the patient role. The assignment of roles removes the ambiguity as to who was kissing
whom without appealing to some ordering to do so. The reliance on thematic roles has
been much criticized because of the difficulty of finding a sufficient number of roles to
analyse all sentences whilst retaining their use in formulating linguistic generalizations.
Nevertheless, there are grounds for using roles merely as a solution to the puzzles of the
identity of information that we have been discussing. A less pejorative approach is to
assign roles on a relation specific basis: we may say that the relation of 'kissing' comes
with two roles, 'the kisser' and 'the kissed'. In our example, Elenor would be adopting the role of kisser and Humphrey the role of the kissed.

Situation theory adopts a neutral position with respect to the possibility of framing generalizations using roles. It is assumed that every relation \( r \) comes with a set of argument roles \( R \) and an infon is formed by assigning objects to those roles. Whether or not roles are shared between relations (like 'agent' and 'patient' are) is left open. The notation for basic infons which we have been using is taken to be shorthand for more explicit notation specifying the roles to which objects are assigned. Thus the information that Elenor kissed Humphrey would be represented by \( [(\text{kiss}, \text{kisser} : \text{Elenor}, \text{kissed} : \text{Humphrey}, +)] \) with the understanding that the argument role / argument pairs are unordered. In other words, we have as a convention of our notation that

\[
\{r, r_1 : a_1, \ldots, r_n : a_n, i\} = \{r, r_{a_1} : a_{a_1}, \ldots, r_{a_n} : a_{a_n}, i\}
\]

for all positional permutations \( \pi \). Technically there is no magic being used here. There is nothing to stop us modelling the roles of an n-ary relation by the the numbers 1, \ldots, \( n \), so that the assignments to those roles are nothing more than the sequences we started with. The importance of this move is merely that no significance is attached to any ordering of the roles. If there is any merit to generalisations across roles then we can reintroduce thematic roles as types of argument roles, e.g. kisser is of type agent.

It should be noted that the use of argument roles is only a partial solution to the problems of individuating information. We could not use this approach to identify the information that Elenor is the same size as Egbert with the information that Egbert is the same size as Elenor.

### 2.2.2 Appropriateness

There are a number of ways in which a sequence of words of a language can fail to express a determinate thought. Purely syntactic factors can be responsible for producing ambiguity or nonsense, but semantic factors can also be relevant. Chomsky's example of "Colourless green ideas sleep furiously" is a notorious illustration of this point. Although syntactically well formed, it is claimed that this expression fails to express a thought since ideas cannot be green, sleeping cannot be done furiously and ideas cannot sleep. It is a category mistake to suppose that a genuine thought is expressed.
CHAPTER 2. SITUATIONS

The resolution of this problem (for relations) is usually attained by classifying objects into various sorts. Assignments to the argument roles of relations are restricted by sorts so that only objects of the specified sorts can be considered as either standing or not standing in the relation. For example, the relation bigger-than presumably only holds between two objects that have some spatial extent. One of our sorts might be $S$, defined as the sort of object that has spatial extent. We would then insist that $\langle\text{bigger-than}, a, b, i\rangle$ only describes an infon if $a$ and $b$ are both of sort $S$. Elenor and Humphrey are all objects of sort $S$, so $\langle\text{bigger-than}, Elenor, Humphrey, +\rangle$ and $\langle\text{bigger-than}, Humphrey, Elenor, +\rangle$ both name genuine infons, whereas $\langle\text{bigger-than}, Elenor, Tuesday, +\rangle$ does not since Tuesday is not of sort $S$.

Unfortunately the use of sorts for resolving category mistakes has similar defects to the use of thematic roles for resolving the identity problems of the previous section. A good guess at the sorts of object which can be compared by the relation 'bigger than' is the sort of objects that have spatial extent. But it is somehow wrong to say that Elenor is bigger than the garage door. Even though garage doors have spatial extent, it is unclear how to compare the size of one with the size of a elephant. One possible solution to this problem would be to concentrate on sorts which do ensure that category mistakes are avoided. Given any two elephants we can say that one is bigger than another and we can also say that one window is bigger than another. Perhaps we could provide an exhaustive list of sort pairs which are each sufficient for avoiding category mistakes when related by bigger-than: elephants and elephants, windows and windows, giraffes and giraffes, giraffes and elephants, elephants and tigers, etc. Another solution would be to say that we are really talking about different relations in each of these cases: there is a relation bigger-than($S_1, S_2$) for each of the pairs of sorts $S_1, S_2$ in the previous list. Just as with thematic roles, we are trading the generalization for accuracy.

In ST category mistakes are avoided by supposing that a relationship of appropriateness exists between an infon's relation and the assignment of objects to its argument roles. The expression $\langle r, r_1 : a_1, \ldots, r_n : a_n, i\rangle$ only names an infon if the assignment $r_1 : a_1, \ldots, r_n : a_n$ is appropriate for the relation $r$. A theory of sorts may be added to constrain the 'appropriate for' relation, but it is not essential that this be done.

Apart from the apparent complexity of an account of appropriateness using sorts, the
situation theoretic approach has the advantage that other kinds of restriction on meaningfulness can be incorporated. Intuitions vary, but one might claim it is nonsense (rather than just false) to say that Elenor is bigger than herself. No sort can be attributed to Elenor which explains this lack of appropriateness. Although somewhat ad hoc, a similar approach can be used to block semantic paradoxes. If $R$ is Russell's property (the property of being non-self-applicative) then we can avoid paradox by claiming that $R$ is not appropriate for the property $R$. As before, one may supplement this account with an explanation on why some assignments are appropriate and others are not, but the theory of infons can be developed independently.

2.3 Types and Propositions

What follows in the next two sections could be described as a situation theorist’s classification of the world as individuated according to some scheme. It is a complex classification because there are many different sorts of entity: individuals, infons, situations, assignments, types, propositions, indeterminates, anchors, properties and constraints. All except the last of these will be defined in a long mutually inductive definition.

The theorist’s classificatory framework consists of types and propositions. The types classify ST entities and the propositions internalize this classification: making entities of the form $(a: T)$, where $a$ is an ST entity and $T$ is a type. Propositions, being ST entities may be classified themselves. For example $(a: T)$ is of type PROP, the type of propositions. Even types can be of a type (e.g. the type TYPE of all types). This part of the theory really has nothing to do with Situation Theory. All that is needed is a consistent type theory that is powerful enough to cope with ST entities.

In this classificatory framework, a particular scheme of individuation is classified. Associated with a scheme are infons, individual objects and properties. For each kind of entity there is a type. For example, SIT is the type of situations and INFON is the type of infons in the scheme. Similarly the 'supports' relation is thought of as a binary type HOLDS-IN which classifies pairs of situations and infons. There is no reason why several different schemes of individuation should not be classified by the theorist at the same time. This could be achieved by having types SIT$(i)$, INFON$(i)$ and HOLDS-IN$(i)$ for each scheme $i$. But the presentation here is restricted to one scheme.
On our view there is a real distinction between types and propositions, on the one hand, and situations and infons, on the other. The ontology of ST is properly limited by the mode of interaction between an 'individuator' (frog, linguistic community or whatever) and the world. In other words, the subject matter of ST should be the structure of this interaction, which in this context means the structure of a scheme of individuation. Whereas the types SIT and INFON refer to real ontological categories, the types TYPE and PROP do not.

The treatment of ST from which our presentation derives was an ambitious one. In it every ST entity was to be treated as a 'first class citizen'. The motivation for this approach seems to be the view that ST should be able to describe a situation theorist's scheme of individuation as well as those of frogs, Japanese speakers and so on. But a situation theorist is concerned with relations like the 'supports' and 'part of' relations rather than the more mundane relations individuated by frogs. In this case there should be some 'reflection' principles providing infons of what we might call the meta-scheme for each proposition in the theory. For example we would want a relation supports such that

\[
\langle \text{supports}, s, o \rangle \text{ is a fact iff } s \models o.
\]

With such principles in place, propositions and types could be abandoned and replaced by a theorist's scheme of individuation rather than the somewhat artificial classification being employed here. However, it is far from clear that this can be done in a consistent way. Instead, we will stick with an informal presentation in terms of propositions and types.

Despite a largely egalitarian attitude to ST objects, there is an important distinction between parametric and non-parametric objects which cuts across the type-proposition/situation-property-infon distinction. An object is parametric if some part of its structure is undetermined. Since the catalogue of objects we are about to list are all determined by a scheme of individuation, we can conceive of parametric objects as partial information about other (less parametric) objects. For any such object there are in general many ways in which it relates to more determined or even totally non-parametric objects. These relations are captured in the theory by anchors; functions which map parametric...
objects to less parametric objects in a way that preserves the information that was already determined.

2.3.1 Propositions

Propositions are bearers of truth in the classical sense: they are either true or false. The most simple is the atomic proposition, which consists of an assignment $a$ and a type $T$. It is written as

$$(a : T)$$

The proposition $(a : T)$ is true if and only if $a$ is of type $T$. Which assignments are of which types is either scheme relative or else structurally determined by the following definitions. In this section we will be principally concerned with the second case.

It should be noted that not all assignments form propositions with all types. There are two conditions which must be fulfilled:

1. $a$ must be an assignment for $T$. Every type possesses a number of argument roles, which can be assigned ST objects. An assignment $a$ for $T$ is just a function which is defined over all the argument roles of $T$.

2. $a$ must be appropriate for $T$. Which assignments for $T$ are appropriate is again either scheme relative or structurally determined.

So to fully describe the collection of atomic propositions we must say what types there are, what their argument roles are, the conditions under which an assignment is appropriate for a type, and the conditions under which it is of that type.

As an abbreviation we write $(a_1 \ldots a_n(T))$ for $(a : T)$ where $a$ assigns $a_1 \ldots a_n$ to argument roles of $T$ which are clear from the context.

In addition to the atomic propositions, there are \{$\land, \lor, \neg$\}-combinations of propositions with the usual truth conditions. Since we are talking about the objects of ST rather the construction of the wffs of some language, we must specify more than truth conditions. For example, the full theory must say whether $\land$ and $\lor$ are commutative and associative and $\neg$ involutive ($\neg\neg p = p$). In at least the case of propositions, it is likely that we would want a boolean algebra over the atomic propositions, with $\land$, $\lor$ and $\neg$ interpreted in the usual way.
CHAPTER 2. SITUATIONS

<table>
<thead>
<tr>
<th>Type</th>
<th>Argument-Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBJ</td>
<td>OBJ-INST</td>
</tr>
<tr>
<td>PROP</td>
<td>PROP-INST</td>
</tr>
<tr>
<td>TYPE</td>
<td>TYPE-INST</td>
</tr>
<tr>
<td>APPROP</td>
<td>APPROP-ASS, APPROP-TYPE</td>
</tr>
<tr>
<td>ASSIGN</td>
<td>ASSIGN-ASS, APPROP-TYPE</td>
</tr>
</tbody>
</table>

Table 2.1: Some primitive types

Parametric propositions are similarly defined from atomic parametric propositions consisting of parametric assignments or types. There are no restrictions analogous to appropriateness; \((a : T)\) is a parametric proposition for any (possibly parametric) assignment \(a\) and (possibly parametric) type \(T\). However, there is also nothing analogous to an assignment being of a type, and so no concept of truth for parametric propositions. As with all parametric objects, parametric propositions only serve to determine non-parametric objects, either by quantification, abstraction, anchoring or as part of a constraint.

2.3.2 Types

There are two kinds of types: primitive types, fully described by the theory, and complex types, formed by abstraction over propositions.

The primitive types (together with their argument roles) which we will discuss in this section are listed in Table 1. Others will be introduced in the next section.

Complex types are formed by abstraction over propositions. Since parametric propositions are just propositions part of whose structure is undetermined, we can regard them as types: an assignment for such a type is of that type if it anchors the parametric proposition to a true non-parametric proposition. The argument roles of a complex type formed in this way are just those parts of the parametric proposition which are undetermined (i.e. the indeterminates).

For parametric complex types, we must distinguish between those indeterminates which we are abstracting over, and those which are to be left as parameters. So a (possibly parametric) complex type is determined by a parametric proposition \(p\) together with a
list $X$ of indeterminates of $p$, and is written

$$[X|p]$$

$[X|p]$ has one argument role for each indeterminate in $X$. They are written as $\text{Arg}([X|p], i)$, for $i = 1$ to the length of $X$, abbreviated to $\text{Arg}_i$ when there is no chance of confusion.

It is important that the indeterminates $X$ are not part of the complex type $[X|p]$ even though they appear in our name for them. In particular, when all the indeterminates of $p$ are included in $X$, $[X|p]$ is non-parametric. This suggests that abstraction over types by basic indeterminates admits of $\alpha$-conversion, i.e. for basic indeterminates $x_1 \ldots x_n$, $y_1 \ldots y_m, [x_1 \ldots x_n|p(x_1 \ldots x_n)] = [y_1 \ldots y_m|p(y_1 \ldots y_m)]$.

$\alpha$-reduction enables us to define boolean combinations of unary types in terms of boolean combinations of propositions:

- $T_1 \land T_2 = \equiv [z|((a_{xT_1} : T_1) \land (a_{xT_2} : T_2))]$
- $T_1 \lor T_2 = \equiv [z|((a_{xT_1} : T_1) \lor (a_{xT_2} : T_2))]$
- $\neg T = \equiv [z|\neg(a_{xT} : T)]$

where $a_{xT}$ is the assignment of $x$ to the (unique) argument role of $T$. We cannot extend this definition to types with multiple argument roles since they would be ambiguous with respect to permutations of the assignment. This problem would be removed by modifying our conception of the argument roles of a type - an option we will consider in the next section.

We may wish to place additional constraints on the formation of complex types by abstraction in order to ensure that various common identities hold between types. For example, although the commutativity of $\land$ and $\lor$ over types is determined by their commutativity over propositions, their associativity is not. Similarly it is undecided as to whether $\neg$ is involutive. Many questions like these must be answered in a fully formalised ST.

### 2.3.3 Argument Roles

The discussion of argument roles in Section 2.2.1 applies as much to types as it does to relations. In ST a neutral position is taken in which the argument roles of a type just
CHAPTER 2. SITUATIONS

form a set. However, we list some possible elaborations.

- A collection of argument roles is specified at the level of the scheme of individuation (although there may be some which are shared by all schemes). A function determines a set of argument roles from this collection for each property. Properties may share argument roles. In this case the corresponding linguistic correlate would be something like thematic role (see, for example, Engdahl 1989).

- As before, except that argument roles are not shared between properties.

- As before, except that the set of argument roles of a property is determined relative to a situation (see Barwise 1989c). Argument roles for types are invariant across situations.

- As before, except that the argument roles assigned to a property can be labelled and labels can be shared by different properties. A subcase of this would be where the argument roles are numbered, i.e. labelled with numerals.

- As before, except that indeterminates can be labelled, so that some of the effects of the previous alternative can be obtained without explicitly labelling argument roles (see Cooper 1987 and 1989).

- As before, except the set of argument roles assigned to a property is ordered (either partially, or totally, or whatever).

- Argument roles are abandoned and assignments are just sequences of objects (see Plotkin 1988). Equivalently, argument roles are numerals and each property is assigned an initial sequence of numerals (e.g. \{1, \ldots, n\}).

A more constrained approach, like the last, enables a wider definition of boolean combinations of properties, but loses out on the additional uniformities which can be captured by a more flexible approach, like the first.

2.3.4 Appropriateness

We have described the types of ST objects and their argument roles. This allows us to know of an assignment \(a\), whether or not it is an assignment for a type \(T\). We
have determined when condition 1 (of section 2.3.1) holds, but whether or not condition 2 holds remains unspecified by the theory so far. To fully describe the collection of propositions, we must say for each type \( T \) what the conditions are for an assignment \( a \) for \( T \) to be appropriate.

- Any assignment \( a \) for OBJ, PROP or TYPE is appropriate.
- An assignment \( a \) for APPROP is appropriate if and only if \( a \) assigns a type to APPROP-TYPE and an assignment for that type to APPROP-ASS.
- An assignment \( a \) for ASSIGN is appropriate if and only if \( a \) assigns a type to ASSIGN-TYPE and an assignment to ASSIGN-ASS.
- An assignment \( a \) for the complex type \([X\triangleright p]\) is appropriate if and only if the anchor \( a_\alpha \) (which maps the indeterminates \( x_i \) in \( X \) to the objects which \( a \) assigns to \( \text{Arg}_\alpha \)) anchors \( p \) to a proposition.
- No assignment is appropriate for a parametric type.
- A parametric assignment is not appropriate for any type (parametric or not).

As was mentioned in Section 2.2.2, appropriateness conditions can be used to block obvious inconsistencies. For example, not being of the type \( T \) is not sufficient for being of the type \( \lnot T \). In particular, we can block the direct road to Russell's paradox given by the argument

Let \( R = [x] \lnot (x : x) \)

\( R \) is of type \( R \) if \( \lnot (R : R) \) is true

\[ \begin{align*}
\text{if } (R : R) \text{ is false} & \\
\text{if } R \text{ is not of type } R & 
\end{align*} \]

by denying that the assignment of \( R \) to its only argument role is appropriate.

2.3.5 Being of a Type

The collection of propositions generated by the primitive types introduced so far is fully described by the above definitions, but the theory also determines which of these
propositions are true. To do this we need to say which appropriate assignments for a type are of that type.

- All (appropriate) assignments for OBJ are of type OBJ.
- An (appropriate) assignment for PROP is of type PROP if and only if it assigns a proposition to PROP-INST.
- An (appropriate) assignment for TYPE is of type TYPE if and only if it assigns a type to TYPE-INST.
- An appropriate assignment for APPROP is of type APPROP if and only if the assignment it assigns to APPROP-ASS is appropriate for the type it assigns to APPROP-TYPE.
- An appropriate assignment \( a \) for the complex type \([X|p]\) is of type \([X|p]\) if and only if the anchor \( \alpha \) anchors \([X|p]\) to a true proposition.
- No assignment is of a parametric type.
- A parametric assignment is not of any type.

Note that the following binary complex type correctly 'internalises' the property of being of a type.

\[
\text{OF-TYPE} = \{ x_1, x_2 \mid (x_1 : \text{TYPE}) \& (x_2 : \text{TYPE}) \}
\]

For any type \( T \), an assignment \( a \) for \( T \) is appropriate for \( T \) if and only if the assignment \( a' \) assigning \( a \) to \( \text{Arg}_1 \) of OF-TYPE and \( T \) to \( \text{Arg}_2 \) of OF-TYPE is appropriate. Also, \( a \) is of type \( T \) if and only if \( a' \) is of type OF-TYPE.

2.3.6 Indeterminates and Anchors

Indeterminates come in two kinds, basic and complex. The basic ones are unstructured and it is assumed that there is a plentiful supply of them. They are the building blocks of parametric objects. There is no primitive type for indeterminates, nor any specifically parametric object\(^6\). This ensures that parametric objects are of parametric types and

\(^6\)Except in Devlin (1988). Also, on some accounts (e.g. Barwise 1987), indeterminates have non-parametric correlates called the reifications of indeterminates. These could have a primitive type associated with them.
CHAPTER 2. SITUATIONS

non-parametric objects are of non-parametric types, in keeping with the intuition that
parametric objects are not a new kind of object but just partial information about non-
parametric objects. Nevertheless, there is a sense in which indeterminates can range over
a particular type. For any unary type $T$, there is a complex indeterminate, $x[(x : T)]$,
which can only be anchored to an object of type $T$.

Parametric objects are related to non-parametric objects by anchors. A function $\alpha$ is
an anchor for the basic indeterminate $x$ if and only if it is defined on $x$ and maps $x$ to
some (non-parametric) object. A complex indeterminate $x[P]$ is formed from a basic indeterminate $x$ and a parametric proposition $P$ which has $x$ as a parameter.

Complex indeterminates function much like basic indeterminates except for an addi-
tional restriction on their anchors. A function $\alpha$ is only an anchor for $x[P]$ if $\alpha$ is
declared on all the parameters of $P$ and $P$ is anchored by $\alpha$ to a true proposition.

A function is an anchor for a parametric object $p$ if and only if it is an anchor for all
the parameters$^{10}$ of $p$.

2.4 Situations and Infons

In this section we show how the constructions of type and propositions apply to scheme-
relative objects like situations and infons. There is an unfortunate terminological confu-
sion in the literature between the terms 'infon' and 'soa' (standing for 'state of affairs').
In what follows we use 'infon' exclusively. However, it should be noted that the infons
of Section 2.2 were non-parametric: situations only support non-parametric infons.

2.4.1 More Primitive Types

The primitive types in Table 2 are needed in extending the theory to situations and
infons.

Our scheme of individuation determines collections of individuals, situations, primitive
properties (and possibly other primitive objects like locations and times). Of these

$^{10}$Whether an indeterminate used in the construction of $p$ is a parameter of $p$ or not depends on
similar criteria to those determining whether a variable is free or bound in a formula.
CHAPTER 2. SITUATIONS

<table>
<thead>
<tr>
<th>Type</th>
<th>Argument-Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>IND</td>
<td>IND-INST</td>
</tr>
<tr>
<td>SIT</td>
<td>SIT-INST</td>
</tr>
<tr>
<td>INFON</td>
<td>INFON-INST</td>
</tr>
<tr>
<td>PROPERTY</td>
<td>PROPERTY-INST</td>
</tr>
<tr>
<td>HOLDS-IN</td>
<td>HOLD-SIT, HOLD-INFON</td>
</tr>
</tbody>
</table>

Table 2.2: More primitive types

collections we can say little without going into the structure of a particular scheme. In Section 1, we said that a scheme of individuation provided a collection of infons. In this section we will describe the structure of that collection as a construction from the other collections we are taking as primitive.

The new primitive types classify those collections of objects particular to our scheme of individuation:

- Any assignment for IND, SIT, INFON, or PROPERTY is appropriate.
- An assignment for IND is of type IND if and only if it assigns an individual to IND-INST.
- An assignment for SIT is of type SIT if and only if it assigns an situation to SIT-INST.

2.4.2 Complex Infons

There are two kinds of infon, basic and complex. Basic infons have already been discussed in Section 2.2. Both negative and positive basic infons are formed from a property \( r \) and an appropriate assignment \( a \) for \( r \). The term 'property' is used no matter how many argument roles \( r \) has: a relation is simply a property with more than one argument role. Assignments for properties are constrained by exactly analogous conditions those for assignments to types (Section 2.3.1).

Complex infons are either \( \{V, A\} \)-combinations of infons, or else quantified infons\(^{11}\).

\(^{11}\)Although some theorists would omit these because of the non-persistence of existentially quantified infons – see Section 2.4.4.
Negation of infons can be defined by cases. Given an infon \( \sigma \), \( \neg\sigma \) is often called the dual of \( \sigma \).

- The dual of a positive basic infon is the corresponding negative one. The dual of a negative basic infon is the corresponding positive one.

- The dual of a \( \{ V, A \} \)-combination of infons is defined in the de Morgan rules:

\[
\neg(\sigma \land \tau) = \neg\sigma \lor \neg\tau \\
\neg(\sigma \lor \tau) = \neg\sigma \land \neg\tau
\]

- The dual of a quantified infon is defined in the usual way as:

\[
\neg(\forall x)\sigma = (\exists x)(\neg\sigma) \\
\neg(\exists x)\sigma = (\forall x)(\neg\sigma)
\]

An assignment for INFON is of type INFON if it assigns an infon to INFON-INST.

In certain versions of the theory (e.g. Barwise and Perry (1983), Devlin (1988)) the infons defined above have been termed saturated infons since their assignments fill, or 'saturate', all the argument roles of the property. On their account, a infon can be unsaturated, i.e. have an assignment which leaves some of the argument roles unfilled. Unsaturated infons are to be distinguished from both the corresponding parametric and existentially quantified infons. The argument for this is usually motivated by linguistic examples (although one could see such distinctions arising in computer science as well).

Suppose the property \( e \) has two argument roles \( A \) and \( B \). We will write the infon formed by assigning the objects \( a \) and \( b \) to \( A \) and \( B \) respectively (with polarity \( i \)) as \( \langle e, A:a, B:b; i \rangle \). Also suppose that in the following sentences, 'is eating' is referring to \( e \), 'Jim' to \( a \), and 'the whole pig' to \( b \). The long arrow indicates a possible interpretation of the sentences' informational contents.

- 'John is eating the whole pig' \( \rightarrow \langle e, A:a, B:b; i \rangle \) (saturated).
- 'John is eating it' \( \rightarrow \langle e, A:a, B:x; i \rangle \) (parametric).
- 'John is eating something' \( \rightarrow (\exists x)\langle e, A:a, B:x; i \rangle \) (existential).
CHAPTER 2. SITUATIONS

- "John is eating" — \{e, A:a;i\} (unsaturated).

Dependencies between these interpretations, such as the rule 'If Jim is eating then Jim is eating something', would be specified by 'constraints' between the corresponding types (see Chapter 3).

2.4.3 Properties

There are two\textsuperscript{12} kinds of properties: primitive properties and complex properties. They come with a number of argument roles\textsuperscript{13} to which assignments can be made. Whether an assignment for a primitive property is appropriate or not, is also determined by the specific scheme of individuation we are concerned with.

Complex properties are formed by abstraction over infons in an analogous way to the formation of complex types. We write

\[ [X]|\sigma| \]

for the complex property determined by the list of indeterminates \( X \) and the parametric infon \( \sigma \). \([X]|\sigma|\) has argument roles \( Arg_i \), one for each indeterminate \( x_i \) in \( X \).

- An assignment \( a \) for \([X]|\sigma|\) is appropriate if and only if the corresponding anchor \( a_{x_i} \) anchors \( \sigma \) to a infon.

- An assignment for PROPERTY is of type PROPERTY if it assigns a property to PROPERTY-INST.

Following the accounts given in Cooper (1987) and Devlin (1988), and assuming \( a \)-conversion, we can define boolean combinations of unary properties as complex types.

- \( r_1 \land r_2 =df [x|((r_1, a_{r_1}) \land (r_2, a_{r_2}))] \)

- \( r_1 \lor r_2 =df [x|((r_1, a_{r_1}) \lor (r_2, a_{r_2}))] \)

\textsuperscript{12}In some versions of the theory, types are also properties. We do not take this line here since we wish to keep the theorist's framework separate from its application to classifying the structure of a scheme of individuation.

\textsuperscript{13}Argument roles may be indexed by a situation \( s \) (see Barwise, 1989c). This would make the collection of infons relative to a situation index also. We would say an infon in the collection determined by the situation \( s \) is an \( s \)-infon. Situations can only support infons from the collection they index.
\[ \neg \tau = \{ x \in \{ (r, a_x) \} \mid \neg (r, a_x) \} \]

where \( a_x \) is the assignment of \( x \) to the (unique) argument role of \( r \). As with combinations of types, these definitions will not work for properties with multiple argument roles unless we restrict our conception of argument role.

### 2.4.4 When Situations Support Infons

The principal classification given by a scheme of individuation is in terms of basic infons. Here we show how this is extended to complex infons. First we need to define the domain of objects which occur in a situation so that there is something for quantified infons to quantify over.

The constituents of a basic infon \( ([r], a; i) \) are those objects assigned to the argument roles of \( r \) by \( a \). The domain of a situation is the collection of objects which are constituents of some basic infon supported by \( s \).

There are many ways of saying \( s \models \sigma \): \( s \) supports \( \sigma \), \( \sigma \) holds in \( s \), \( s \) makes \( \sigma \) factual, \( \sigma \) is made true by \( s \). We give an inductive definition for when this condition obtains:

- If \( \sigma \) is the basic (non-parametric) infon \( ([r], a; i) \) and \( r \) is a primitive property then whether or not \( s \models \sigma \) is determined directly by the scheme of individuation.
- If \( \sigma \) is a basic infon \( ([X]^r], a; i) \) and \( a_x(r) \) is a infon, then \( s \models \sigma \) just in case \( s \models a_x(r) \) if \( i = 1 \), or \( s \models \neg a_x(r) \) if \( i = 0 \).
- If \( \sigma \) is the complex infon \( \tau \land \nu \) then \( s \models \tau \) and \( s \models \nu \).
- If \( \sigma \) is the complex infon \( \neg \nu \) then \( s \models \sigma \iff s \nmid \nu \).
- If \( \sigma \) is the complex infon \( \forall x \sigma \) then \( s \models \sigma \) iff for every object \( b \) in the domain of \( s \), \( s \models a(x) \), where \( a \) is the anchor which only anchors \( x \) to \( b \).
- If \( \sigma \) is the complex infon \( \exists x \) then \( s \models \sigma \) iff there is some object \( b \) in the domain of \( s \) such that \( s \models a(x) \), where \( a \) is the anchor which only anchors \( x \) to \( b \).

An assignment \( a \) for HOLDS-IN is appropriate if and only if it assigns a situation \( s \) to HOLD-SIT and a infon \( \sigma \) to HOLD-INFON. It is of type HOLDS-IN if and only if \( s \models \sigma \).
Propositions of the form \((a:\text{HOLDS-IN})\), called Austinian propositions (see Barwise and Etchemendy 1987), are often written as \((a \models \sigma)\). If the actual world \(W\) is a situation, Austinian propositions of the form \((W \models \sigma)\) are called Russellian propositions. If, in addition, all infons are persistent, \(W\) will support an infon iff it is a fact. The content of the Russellian proposition \(W \models \sigma\) can then be identified with \(\sigma\).

We have come to the end of our lightning tour through the taxonomy of ST entities. Much has been missed out on the way, but we hope to have conveyed to the reader an impression of the structure of a single scheme of individuation. The most important omission is that of a discussion of 'constraints'. This will be rectified in the next chapter.
Chapter 3

Constraints

In Chapters 1 and 2 we were only concerned with the ability of organisms to classify their environments. Chapter 1 studied the general properties of classifications, whilst Chapter 2 focused on the classification of situations. Situation Theory provides a very rich classification of entities of the various ontological categories induced by a scheme of individuation: objects, properties, relations and infons.

Nevertheless, there is clearly more to cognition that simply dividing the world into classes. In order for there to be regularity in the world, there must be certain dependencies between the types within a classification. If the types of a classification were entirely independent then the knowledge that something was of type \( t \) would not and could not have any bearing on our response to this knowledge. Of course, there may be classifications which do not have dependencies between the types, but these can play no part in a reasonable model of a world which is for the most part subject to laws and regularities of one kind or another. Our understanding is shaped by these regularities and our behaviour is dependent on this understanding.

Before proceeding to discuss the way in which we can model the dependencies between types of a classification, it may be useful to review some examples from Chapter 1. Although our definition of a \( D \)-classifier had nothing to say about the structure of the types or there interactions with other types, each of our examples focused on a particular kind of classificatory system with a characteristic structure on its collection of types.

The subtype relation is a dependency between the types of a hierarchy. It constrains the token domain, \( D \), so that if \( d \in D \) and \( d : t \) and \( t < t' \) then \( d : t' \). Speaking in
CHAPTER 3. CONSTRAINTS

epistemic terms, if we know that something is of type \(t\) and \(t\) is a subtype of \(t'\) then we also know (or can infer) that it is of type \(t'\). The incompatibility relation between features in a feature system has a similar role. If something has the feature \(f\) and \(f \perp f'\) then it cannot also have feature \(f'\). In both cases partial knowledge about how a token is classified has consequences for the complete classification of the token in the system: the subtype relation forces types into \(D\), whilst the incompatibility relation forces them out.

Both subtype and incompatibility relations are restricted in that they only constrain the classification of individual tokens. They do not express dependencies between the classifications of different tokens. An example of this kind of dependency is the notion of linguistic dependency of Section A.3. Here the tokens are words (or rather occurrences of words) and the classification of one word constrains the classification of its neighbours according to the grammatical constraints between their lexical types. Using a metaphor from communication theory, which we shall repeatedly exploit in this chapter, we say that information flows from one word to another.

In Section 3.1 we discuss Dretske's theory of the flow of information which he derives from certain principles of communication theory. We discuss some of the basic concepts of communication theory before examining Dretske's definition of the information carried by a signal. Of great importance to what follows is the relativity of his definition to a communication channel. We end by presenting the distinction between analog and digital modes of carrying information.

In Section 3.2 our exposition of Situation Theory (ST) from Chapter 2 is extended to cover constraints. We discuss various accounts of the logical properties of constraints in ST, their problems and some solutions. In doing this we discuss extensions of the basic idea of a constraint to general constraints and conditional constraints. We then use constraints to give a ST definition of the information carried by a situation and we compare this with Dretske's. The definitions are found to differ only in the implicit acknowledgement of relativity to the communication channel in Dretske's account.

In Section 3.3 we ask how an organism can use the information flow in its environment to further its own ends. Barwise and Perry's notion of attunement to a constraint appears too weak to be useful but, through an extended example, we show how attunement
works for organisms that are in the right place at the right time.

Finally, in Section 3.4 we follow up a suggestion of Devlin’s that the process of individuation can be characterized (in part) as one of digitalization. The same claim is made by Dretske about the cognitive processes as opposed to perceptual ones. We find both ideas lacking, but acknowledge the possibility of a ‘constraints before individuals’ approach. However, we find that the ST framework is too rigid since it fails to capture the essential relativity of Dretske’s approach.

3.1 The Flow of Information

In his book “Knowledge and the Flow of Information” (1981) Dretske gives a philosophical reconstruction of various traditional epistemological ideas in terms of the mathematical theory of information. This is a branch of communication science which was developed by Shannon and Weaver (1949) in order to understand and correct for noisy signals in large electronic communication networks. The basic idea is that our semantic idea of what it is for a message to convey certain information is an idealization of the communication theoretic idea of information flow. Shannon and Weaver are concerned with noisy signals and the quantity of information that a signal reveals about its source. Dretske is concerned with quality rather than quantity, but he finds a way of expressing the semantic notion of information (‘information that’) in the terms of communication theory.

Our treatment of Dretske’s theory is brief. We merely attempt to extract the main points of relevance to this thesis. Firstly, the communication theoretic view of information is expounded. Discussions of information based on probability can have an undesirably detached quality if they are conducted without some regard for the probabilistic models in question. Secondly, Dretske’s definition of information content is presented together with a discussion of his actual and potential responses to an objection. Lastly, the ‘analog/digital’ distinction is defined in preparation for later discussion.

3.1.1 Communication Theory

To appreciate Dretske’s position, it is useful to have some idea about the various quantities used in communication theory. A communication engineer is interested in modelling
CHAPTER 3. CONSTRAINTS

... the behaviour of communication systems in a way that provides an accurate measure of their efficiency in transmitting information, how much noise there is and so on. Because the behaviour of such systems is in general very complex and unpredictable, the engineer must make use of probabilistic models.

The fundamental quantity studied in communication theory is the amount of information 'generated' by a device, but before we can state what this is we need a few other definitions. Firstly, a 'device' is taken to be anything which is in one of a mutually exclusive set $S$ of states at any one time. Associated with each state, $s \in S$, is the probability $P(s)$ of the device being in state $s$ at an arbitrary time. In general, the probabilities associated with each state will be different so that the device is in some states far less frequently than it is in others. Being in an unlikely state has 'more information' associated with it than being in a very probable state. The quantity of information associated with the device being in state $s$ is called the surprisal of $s$, written $I(s)$ and defined by

$$I(s) = -\log P(s)$$

i.e. the amount one is 'surprised' by the device being in state $s$ is inversely proportional to the logarithm of the probability of it being in that state. For example, an 8 sided balanced dice (an octahedron) will have a probability of $1/8$ of landing on any one side. The surprisal associated with it landing on side 4 (or any other side) is given by $I(\text{side 4}) = -\log 1/8 = \log 8 = 3$. Intuitively the answer is 3 because it takes 3 bits of information (a 3 digit binary code) to make the number of distinctions necessary (i.e. 8) to specify the state of the dice. With unfair dice the surprisal associated with each side will differ according to the probability function.

The information generated by the device is the quantity of information that we might 'expect' to be produced by the device at an arbitrary time. 'Expectation' is to be taken in its strict probability theory sense: the expectation of some random variable is the average value of that variable weighted by the likeliness of each of its values, so that unlikely values contribute less to the expectation than likely values. It this case we are interested in the expected value of the surprisal, so the information generated, $I$, is defined by

$$I = \sum_{s \in S} P(s) \cdot I(s)$$

\(^1\text{All logarithms will be taken to the base 2.}\)
CHAPTER 3. CONSTRAINTS

In the example of the octahedral balance dice, the surprisal is constant (=3), so \( I = 1/8.3 + 1/8.3 + 1/8.3 + 1/8.3 + 1/8.3 + 1/8.3 + 1/8.3 + 1/8.3 + 1/8.3 = 3 \): we, unsurprisingly, expect the surprisal to be 3! For an unfair dice, however, the result would be different.

We now move to a two-device setting in which the big question is: how much information is (on average) transmitted from one device to the other? A two-device communication system is pictured in Figure 3.1. We suppose that a signal is sent from one of the devices (the source) to the other (the receiver). Each device is capable of being in one of a number of states. The states of the source are \( S_s = \{1', 2', 3', 4', 5', 6'\} \) and those of the receiver are \( R_r = \{a', b', c', d'\} \). At any one time, each device is in one and only one of its states.

The answer is provided by considering what can get in the way of successful communication. The information generated by the receiver \( (I_R, \text{say}) \) must be in some way dependent on the behaviour of the source. But how much of \( I_R \) is really influenced by the source and how much is just produced by noise in the system. It may be that the receiver behaves in a partly random way, either due to interfering processes at the source or in the communication channel between the source and the receiver, or even at the receiver itself. Noise is a phenomena with which we are all familiar, but in communication theory it has a specific meaning. The noise in the system is calculated by considering the noise-contribution \( N(s) \) of each of the states \( s \) of the source \( (s \in S_s) \).

Figure 3.1: A Two-Device Communication System.
CHAPTER 3. CONSTRAINTS

If the source actually is in state $s$ then to work out its noise-contribution we want to see how much of the information generated by the receiver is generated independently of this fact.

In a perfect communication system the state of the receiver would be completely determined by the state of the source, so if we 'held down' the source in state $s$, the state of the receiver would be constant. In this case the surprisal associated with that state would be 0 (we cannot be surprised since it could not be different) and as the receiver never changes state (whilst the source is 'held' in state $s$) the information generated (= average surprisal) would be 0 also.

In a noisy system, however, it may be that the receiver continues to change state even while the source is held in state $s$. Any fluctuation of this sort can only be due to something interfering with the communication, i.e. noise. A measure of this noise is given by the information generated by the receiver in this situation (i.e. when the source is held down in state $s$). This is calculated in the same way as before except using conditional probabilities, i.e. the surprisal of receiver state $r \in R_{st}$ given that the source is in state $s$ is $-\log P(r|s)$ and so the average information generated by the receiver (given $s$) - the noise-contribution of $s$ - is

$$N(s) = -\sum_{r \in R_{st}} P(r|s) \log P(r|s)$$

The noise in the system is determined as a weighted average of the noise-contributions of each $s \in S_{st}$, i.e.

$$N = \sum_{s \in S_{st}} P(s) N(s)$$

Now the noise in the system has been quantified, it is easy to calculate the average quantity of information transmitted. We know the amount of information generated by the receiver is $I_R$ and we know that some of this (the amount $N$) is down to noise. So the average amount of transmitted information, $T$, is given by

$$T = I_R - N$$

In studying information flow, we are not only interested in the amount of information transmitted, but how much got left behind. If $I_S$ is the amount of information generated by the source and $T$ is the amount transmitted to the receiver, then $T - I_S$ is the
amount of information that fails to get through. Technically, this quantity is called the equivocation, $E$, of the system. It is defined in a similar way to noise (see Dretske 1981 ch 1), as a weighted average of the equivocation-contributions $E(r)$ associated with each state $r$ of the receiver, i.e.

$$E = \Sigma_{r \in R_t} P(r).E(r)$$

where $E(r) = -\Sigma_{s \in S_t} P(s|r).\log P(s|r)$. A simple calculation shows that

$$T = I_S - E$$

A useful diagram from Dretske's book which summarizes the dependencies between these quantities is shown in Figure 3.2.

The quantities of information used above cannot be used as they stand to say much about the kinds of informational dependencies that Dretske is concerned with. They are measures of the average amount of information involved in the communication.

To say something about the information contained in a specific signal (a state of the receiver) they must be adapted. But the adaptation is fairly straightforward. We know that the information generated by the source's being in a particular state $s \in S_t$ is the surprisal of $s$, i.e.

$$I_S(s) = -\log P(s)$$

and the equivocation-contribution of the receiver's being in a particular state, $r \in R_t$ is $E(r)$. So if the source is in state $s$ and the receiver is in state $r$ then the information transmitted about the source to the receiver is given by

$$T(r, s) = I_S(s) - E(r)$$
CHAPTER 3. CONSTRAINTS

This equation provides us with a measure of the quantity of information transmitted by a specific signal about the source. However, the use of the equation to Dretske is not in determining amounts of information at all. Rather, it provides a condition on the possibility of the signal containing the information that the source is in a certain state: to do so the information transmitted must be as much as the information generated by the source's being in that state. In other words, \( T(r, s) = I_g(s) \) and so \( E(r) = 0 \).

Zero signal equivocation is therefore the criterion for all the information about the source (i.e. its state) to flow to the receiver. It is important to distinguish this condition from the condition that the system has zero equivocation \( (E = 0) \). There can be total information flow from source to receiver in a system which lacks perfect communication: all that is required is that the equivocation of the particular signal be zero.

The condition is also more general than might appear. Whilst zero equivocation is necessary for a signal to contain total information about the source (information that it is in state \( s \)) this does not mean that no information is transferred. Consider, for example, the information, \( \phi \), that the source is either in state '1' or in state '2'. If the source is in fact in state '1' and the signal equivocation is non zero, this does not imply that \( \phi \) is not transmitted, only that the information that the source is in state '1' is not transmitted. Whether or not \( \phi \) is transmitted is determined by a re-specification of the state space. The classification of the source into states \( S_{esr} = \{ '1' \) or '2' \}' '3', '4', '5', '6' \) is just as good as the original classification: the states are mutually exclusive and exhaustive. Also the probability measure \( P \) determines the probability of the source being in state '1' or '2' as \( P(1 \) or '2') = \( P(1) + P(2) \) (since the two events are mutually exclusive). Now in the new classification, the information generated by the source decreases since \( -\log(P(1) + P(2)) \leq -\ln P(1) \) (the function \( x \mapsto -\log x \) is decreasing for \( x \leq 1 \)), but the information transmitted must remain the same (see Dretske), so the equivocation must also decrease. If it decreases to zero then 'message' \( \phi \) is transmitted. The important point of this conclusion is that whether or not a certain 'message' is transmitted is crucially dependent on the classification of the source.
3.1.2 Information Carried and Information Content

Before supplying his criterion for a signal to carry semantic information, Dretske argues for three principles that any account of information content should obey.

Dretske's Principles

If a signal carries the information that \( s \) is \( F \) then

1. (A) the signal carries as much information about \( s \) as would be generated by \( s \)'s being \( F \)

2. (B) \( s \) is \( F \)

3. (X) if, in addition, \( s \)'s being \( F \) carries the information that \( s' \) is \( G \) then the signal also carries the information that \( s' \) is \( G \).

Condition (A) is really a restatement of the zero equivocation condition. Condition (B) demands that information is veridical. Condition (X), the Xerox condition, has independent motivation which will be discussed below. Finally, he gives the definition.

Definition 3.1 (Information Content.) A signal \( r \) carries the information that \( s \) is \( F \) if the conditional probability of \( s \)'s being \( F \), given \( r \) (and \( k \)), is 1 (but given \( k \) alone is less than one),

where \( k \) is a parameter standing for the prior knowledge about \( s \) of the observer to whom information is carried.

Rather than analyse Definition 3.1 here, we will discuss an objection to it. The definition certainly obeys each of the principles, but it seems very strong. Does the conditional probability really have to be as high as 1? And if so, is this condition ever met in practice: is there any information flow at all?

Dretske's answers to the first objections is contained in his defence of the Xerox Principle. The main idea, as the name implies, is one does not lose information by duplicating something that carries information as long as the duplicates carry information themselves about the thing they were copied from.

Suppose a message is propagated along a chain of speakers of English, each one communicating privately with the next in the manner of 'Chinese Whispers'. If the transmission is less than perfect at each link then an easy calculation shows that the conditional
probability of the message being \( m \) given the report at a certain stage in the chain will decrease along the chain. In fact, given a long enough chain the probability of the message being \( m \) given the report of the last member of the chain will be insignificantly different from the unconditional probability of its being \( m \) (which in an infinite language is likely to be zero). Dretske's point is that our semantic notion of information does not decay in this way. If the report of any member of the chain means that \( m \) then either the next member understands him and reports that \( m \), or there is some mistake and he reports that \( m' \). If there is such a break in the chain, then no semantic information about the original message is transferred. Whereas if there is no such break, the message will survive intact to the end of the chain, however long it is.

The second objection, that the condition always fails in practice has a revealing response. Firstly, the definition is relative to classifications of the source and the receiver. All that is required is that the classifications are strict state systems: the source and receiver must be in one and only state at each time. There is no requirement that the classification has to be in terms of some absolute notion of physical state, or even reducible to such terms. Dretske trades on the same notion of the 'objective relativity' of classification that we have discussed in previous chapters.

Secondly, as the example of disjunctive messages at the end of the last section showed, a classification which can be reduced to a state system is just as good. This means that any of the classification systems in Chapter 1 which could be represented as state systems would be suitable for the application of Definition 3.1.

Thirdly, classifications aside, the definition is relative to a probability measure. That the probability measure is 'objective' is important to Dretske's claim that he is providing account of information as an "objective commodity". But there is an implicit relativity in any measure of probability. To substantiate this claim we will use another of Dretske's examples.

Consider a simple electrical circuit in which the leads of a voltmeter are connected to measure the potential difference across a resistor. There is clearly an informational link between the reading of the voltmeter and the potential difference across the resistor. The conditional probability of the p.d.'s being 6 volts given that the needle on the voltmeter's dial is pointing to the figure '6' is 1: this is what allows the state of the voltmeter's
CHAPTER 3. CONSTRAINTS

needle to have informational content. However, this entire description assumes that the voltmeter is working normally (as well as many other less obvious conditions, for example, we assume that there is no large magnet in the vicinity). Dretske acknowledges these assumptions in noting that his definition of information content is relative to a communication channel and hence to the conditions under which the channel works.

The notion of a communication channel will be developed further in Chapter 6. so we will content ourselves here by pointing out an important misconception. A communication channel is a connection between specific situations on the basis of which information can be said to flow from one to the other. The resistor is connected to the voltmeter by the electrical circuit, so that information can be obtained in the voltmeter situation about the resistor situation. In the example it is tempting to think of the communication channel as constituted by the wires which connect the resistor to the voltmeter. But this is quite wrong: a communication channel is an informational connection between situations, not a physical connection between objects. Whereas the wire links one object to another, the channel links a range of situations (the different situations at the resistor) to a range of other situations (the corresponding situations at the voltmeter). The range of situations it connects cannot be defined in terms of the range of situations in which the resistor is connected to the voltmeter by the wires of the circuit since it is possible to break the informational connection (e.g. by damaging the voltmeter) without breaking the wires.

The relativity implicit in the communication channel is not to be confused with the parameter $k$ occurring in Definition 3.1. This provides a means of relativizing the notion of information flow to an observer's prior knowledge of the source, so that no information is carried that is already known and more information may be carried on the basis of this knowledge. One can think of conditional probabilities as ways of 'shifting' the underlying probability measure. Given a measure $P$ and a measurable event $e$ we can define a new probability measure $P_e$ by $P_e(x) = P(x|e)$. This is how $k$ works. If there are conditions on the effective working of a channel which are expressible as measurable events, then these conditions can be incorporated into the probability measure. But this internalization of conditions does not remove the fundamental dependence of the probabilistic model of the system on conditions which are not expressible as measurable
CHAPTER 3. CONSTRAINTS

events in the system.

The various kinds of relativity which are needed to support Dretske's definition will be important in the discussion in Section 3.4 and to a certain extent motivate the consideration of 'perspectives' in Chapter 4. It is unclear to what extent Dretske accepts the idea that probability measures have this degree of relativity since he sometimes writes as if everything would ultimately be reduced to some notion of absolute probability.

3.1.3 The Analog/Digital Distinction

Although Definition 3.1 purports to be a definition of information content, this is only so on a very broad sense of 'content'. Information flow is all pervasive. Whenever there is some causal connection between events, information is sure to flow. Moreover, a particular signal will in general carry many different 'messages'. But this does not accord with our intuitions about semantic information. The information content of a symbol is usually taken to be unitary. Dretske draws a distinction between two different ways in which a signal can carry a particular piece of information: an analog way and a digital way. He then claims that this distinction forms the basis of an explanation of the difference between the perceptual and cognitive abilities of an organism. We will assess this claim in Section 3.4, but first we must present the distinction itself.

To introduce Dretske's treatment of the analog/digital distinction we must first explain how information can be nested. The information that t is G is nested in s's being F if and only if s's being F carries the information that t is G. For example, the information that s is a rectangle is nested in s's being a square. Similarly, the information that it will be light outside in twelve hours time is nested in its being dark now. These examples do not make explicit that nestedness is also constraint relative. In the second case, for example, the relevant constraint is that night follows day on a 24 hour cycle (and there is more day than night). This works in Edinburgh now, but not in the winter.

Using nestedness Dretske makes a distinction between two ways in which information can be carried. A signal carries the information that s is F in digital form if and only if all the information it carries about s is nested in s's being F. Otherwise the signal carries the information that s is F in analog form.

A clear illustration of this distinction is the difference between a picture and a state-
CHAPTER 3. CONSTRAINTS

ment. For example, compare a photograph of the shops on Broadwater Parade with the veridical utterance "Steel's Electricals is between the florist and the Golf Shop". The photograph carries the information that Steel's electricals is between the other two shops in analog form since it carries additional information about the shops which is not nested in Steel's being between the florist and the Golf Shop. It carries the information that Steel's has many different lights in the window, that the sign for the Golf shop is green and that the Golf shop is to the right of Steel's whilst the florist is immediately to the left. The information carried by the utterance is in digital form since all the information carried about the shops, such as the fact that Steel's is closer to the florist than the Golf Shop is to the florist, is nested in Steel's being between the other two.

Although, we naturally think of the distinction in terms of the properties of the signal (photos give analog information, statements give digital information) it is important to note that the distinction is not absolute, but rather relative to the piece of information concerned. The photograph carries the information that the photographed scene was ... (substitute a very long description of every aspect of the photographed scene) in digital form. In fact every signal carries some information in digital form: For any object, s, the conjunction of all the information that a given signal carries about s is carried in digital form.

It is also worth noting that Dretske uses the terms 'analog' and 'digital' in a slightly non-standard way. These terms are usually used to distinguish continuous and discrete representations. The representation of the time by the continuous motion of the hand of a clock is said to be analog, whereas the numbers on the LCD of a wrist-watch is said to be a digital representation of the same thing. This clearly relates neither to the Dretske's analog distinction, nor to Devlin's usage of it. However, confusion on this point is encouraged by the fact that electronic analog/digital conversion from a continuous waveform into a discrete bit pattern typically involves a loss of information (that is, the waveform can never be perfectly recovered from the bit pattern). Notice that, unlike Dretske's distinction, the continuous/discrete distinction is a distinction among representations.
3.2 Constraints in Situation Theory

The development of ST presented in Chapter 2 was incomplete. We dealt with the aspects of the theory concerned with the classification of situations into types which a given scheme of individuation induces. But this is just a classification: there is no notion of information flow apart from that induced by the logical structure of infons and propositions. A scheme of individuation determines, for each situation \( s \), the type \([s] \models \sigma\) of infons supported by \( s \). The collection of infons of this type are those that are locally determined to be factual by the situation \( s \). From the point of view of the organism whose scheme of individuation we are modelling, this is the collection of information immediately available at \( s \). We can think of these collections of infons as constituting the organism's primary understanding of the world it lives in.

We say 'primary understanding' because situation theory models another way in which an organism can extract information from its environment. It is supposed that, because of the way the world is, there are various law-like dependencies between situation types. Whenever a situation is smoky there is sure to be a firey situation nearby. In ST these dependencies are captured by a new class of propositions called constraints. The secondary means an organism has of understanding the world is to be attuned to certain constraints. If a situation is classified by the organism as being of a smoky type and it is attuned to the constraint that smoke means fire then it is able to extract information about a nearby situation, that it is firey. Following Dretske, information arrived at in this way is said to be carried by the situation in virtue of the constraint. We write \( s \models \sigma \) for the proposition that \( s \) carries the information \( \sigma \) in virtue of a constraint in \( C \) (some collection of constraints).

The distinction between \( \models \) and \( \vDash \) will be further illustrated by means of an example.

Mandy ate two chocolate eclairs and three brownies when we were at Millie's Tea House last week. When she got to school the next day her face was all spotty. The other girls guessed that Mandy had been on another chocolate binge.

The information \( \sigma \) that Mandy ate lots of chocolate is made factual by the tea-time situation \( s_1 \) at Millie's last week (i.e. \( s_1 \models \sigma \)). The later situation \( s_2 \) of Mandy's
arriving at school, supports the information \( \tau \) that she is unusually spotty \((s_2 \models \tau)\). This information is available to Mandy’s school friends who are present when she walks into the classroom. But the additional information \( \sigma \) concerning Mandy’s chocolate binge is also conveyed to the girls. Although \( \sigma \) is not supported by the situation \( s_2 \) of Mandy’s arriving at school, it flows there in virtue of the constraint \( C \) that having a spotty face is a sure sign of having eaten too much chocolate \((s_2 \models \neg C \sigma)\). Information only flows if the connection between \( \tau \) and \( \sigma \) really is law-like. If the common wisdom that eating chocolate gives you spots is just an old wives tale then the information \( \sigma \) would not be carried by \( s_2 \), even if Mandy did eat all those chocolates at Millie’s Tea House. Nevertheless, the connection only has to be locally law-like. It can have all kinds of exceptions. If Mandy had measles, for example, the constraint between spots and chocolate would not apply: having spots would no longer be a sure sign of having eaten too much chocolate, only of Mandy’s measles.

Barwise and Perry (1983) provide a classification of what they take to be the most important kinds of constraint. These are:

1. **Necessary Constraints** reflect the dependencies which result purely from the way a scheme of individuation is organized. Given the nature of our classifications of the world, these are the dependencies which hold no matter what the world is really like. ST is a highly intensional theory, properties and relations with equal extensions are rarely identified. Any necessary connections between properties and relations must therefore be captured by constraints. Included in the assortment of connections that count as ‘necessary’ are taxonomic relations between properties like ‘moles are mammals’, argument role constraints between relations and properties like ‘sipping involves sipping something’, and incompatibility restrictions between properties like ‘red things aren’t green’, or even between individuals like ‘Kriesel is not Keisler’. Mathematical constraints might be thought of as necessary constraints too.

2. **Conventional Constraints** reflect dependencies which exist in systems of classification used by communities. Linguistic dependencies, for example, should be

---

2 And necessarily so. Plotkin (1988) has shown that for the fragment of ST dealing with relations, individuals and infons, even \( \beta \)-conversion between infons results in inconsistency.
captured by conventional constraints, as should the rules of backgammon or the rule of driving on the right side on the road.

3. *Nomic Constraints* include practically ever other kind of dependency which is determined by the way the world is. These are the 'natural laws' of science: not just the things that scientist's study, but also their applications to everyday situations. Nomic constraints capture the regularity that unsupported wine glasses fall to the floor, or that the radio works if you hit it, as well as the more esoteric regularities about chemical bonds or electromagnetism. The notorious 'smoke means fire' is a nomic constraint.

4. *Metatheoretic Constraints* are constraints which reflect the structure not of the world, but of ST itself. The inclusion of these kind of constraints reflects the hope that ST should be able to express things about its own structure. For example, there may be a constraint that if a situation supports $\sigma_1$ and $\sigma_2$ then it also supports $\sigma_1 \land \sigma_2$. This aspect of the theory has not been developed very far, so we will ignore metatheoretical constraints in what follows.

In ST there are several approaches to formalizing constraints. The most simple takes constraints to be propositions with two situation types as components. A situation type is a type with a single argument role which forms a proposition when that argument role is assigned a situation. For example, the complex types

$$[z](x \rightarrow \sigma)$$

formed by abstraction over Austinian propositions are situation types. The logical behaviour of constraints depends on what kind they are, but little has been said on this important and difficult subject. What follows from there being a constraint between $T_1$ and $T_2$ is far from clear. However, two kinds of constraint which have been discussed are the 'involves' and 'precludes' constraints. The classification of constraints into these kinds cross cuts the classification discussed above. Whether or not a constraint is of one of these kinds depends on its logical behaviour.

3The 'precludes' constraint has not been discussed very much. We are basing our account on some remarks in Barwise (1989c) and on a similar relation between 'events' in Barwise and Perry (1983)
'Involves' and 'precludes' constraints are specified using the binary connectives $\Rightarrow$ and $\perp$. The constraint that $T_1$ involves $T_2$ is written $T_1 \Rightarrow T_2$ and the constraint that $T_1$ precludes $T_2$ is written $T_1 \perp T_2$. Constraints of this form constrain by satisfying the following conditions:

- If $s$ is a situation of type $T_1$ and $T_1 \Rightarrow T_2$ (is true) then there is some situation of type $T_2$.
- If $s$ is an situation of type $T_2$ and $T_1 \perp T_2$ (is true) then there is no situation of type $T_2$.

It is important to realize that these statements only provide necessary conditions on $T_1 \Rightarrow T_2$ and $T_1 \perp T_2$. It is quite consistent for $T_1 \Rightarrow T_2$ to be true even if for all situations of type $T_1$ there is a situation of type $T_2$. This latter fact may just be an accident of the way the world is not a genuine law-like dependency.\footnote{For this reason $\Rightarrow$ and $\perp$ are not really connectives at all. Connectives must have introduction and elimination rules, whereas $\Rightarrow$ and $\perp$ only have elimination rules.}

As necessary conditions, the above statements are fairly reasonable for versions of ST which require that there is a greatest part of the world or no maximal parts. But if there is a maximal part (world) which is not the greatest part then the condition is not strong enough. If $s$ is an actual situation of type $T_1$ and $T_1 \Rightarrow T_2$ then there should be an actual situation of type $T_2$. This can be rectified by replacing 'situation' by 'actual situation' in the above. But even if there are non-actual situations the restriction to actual situations is still somewhat limiting. If there are other possible worlds, then surely constraints bite on them too. In this case the conditions can be improved by replacing 'situation' in the above by 'part of $w$' where the conditions now range over all possible worlds $w$. This still fails to satisfy our intuitions if there are situations which are not part of any world, but our intuitions about such situations (if they exist) are probably fairly incoherent anyway.

Example. The infamous 'smoke means fire' is the constraint $T_{\text{smoky}} \Rightarrow T_{\text{firey}}$ where $T_{\text{smoky}}$ is the type of smoky situations and $T_{\text{firey}}$ is the type of firey situations. Applying the modified condition, we get that, if $T_{\text{smoky}} \Rightarrow T_{\text{firey}}$ is true (i.e. smoke does mean fire) then, for all worlds $w$, if $s \leq w$ and $s$ is of $T_{\text{smoky}}$ (i.e. $s$ is a smoky part of $w$) then there is some $s' \leq w$ of type $T_{\text{firey}}$ (i.e. there is some part of $w$ that is firey).
CHAPTER 3. CONSTRAINTS

Although constraints should apply to other possible worlds (if they exist), they should not apply equally. Can we not conceive of constraints which hold in some possible worlds and not in others? Although it could be argued that necessary constraints apply equally to all possible worlds, and perhaps an argument could be extended for nomic constraints, it is surely no so for conventional constraints. Different constraints have different notions of possibility attached to them and sometimes they interact in strange ways. This much at least has been established by research into Conditional Logic.

A potential solution to this problem (one that is adopted by Barwise) is to take constraints to be infons rather than propositions. In this way sense could be made of constraints holding in some worlds and not in others. On this approach, the symbols \( \Rightarrow \) and \( \perp \) are taken to denote relations between situation types. The condition would then be

- If \( s \not\subseteq w \) and \( s \) is of type \( T_1 \) and \( w \models \left( \Rightarrow, T_1, T_2 \right) \) then there is some \( s \subseteq w \) of type \( T_2 \).
- If \( s \not\subseteq w \) and \( s \) is of type \( T_1 \) and \( w \models \left( \perp, T_1, T_2 \right) \) then there is no \( s \subseteq w \) of type \( T_2 \).

The problem here is that this would mean that the worlds in which the constraint held would have to be situations. However, the approach is valuable and will reappear in our discussion of constraints in Chapter 4. For the remaining point we wish to make we will stay with the constraints as propositions approach.

Infons or propositions aside, there are various problems with the account presented so far. Firstly there is a double role being played by situations. On the one hand they are validating the inference from \( T_1 \) to \( T_2 \) in virtue of the above conditions. But on the other, they are acting as the objects of predication of the types. We can only really express dependencies between situation types on this account. Types like \( T_{\text{smoky}} \) and \( T_{\text{firey}} \) work because the properties 'smoky' and 'firey' can be thought of as applying to situations (although they are probably better thought of as applying to locations). We cannot express a dependency like 'kissing means touching' in this way. The best we can do is to express the dependency as a collection of more specific constraints like the constraints...
for each appropriate pair of individuals $a$ and $b$. A more sophisticated approach is to consider a new class of propositions called \textit{general constraints}. The new approach treats $\Rightarrow$ and $\perp$ more like quantifiers than connectives as they require a set $X$ of parameters and parametric situations types $T_1$ and $T_2$ whose parameters are contained in $X$ in order to make the propositions

$$T_1 \xrightarrow{X} T_2 \text{ and } T_1 \perp T_2$$

These propositions are not parametric since each parameter in $X$ is taken to be bound. The conditions which ensure that general constraints constrain are given by:

- If $T_1 \xrightarrow{X} T_2$ (is true) and $\alpha$ is an anchor for the indeterminates in $X$ then $\alpha(T_1) \Rightarrow \alpha(T_2)$ (is also true).
- If $T_1 \perp T_2$ (is true) and $\alpha$ is an anchor for the indeterminates in $X$ then $\alpha(T_1) \perp \alpha(T_2)$ (is also true).

Example ‘Kissing means touching’ can now be expressed by the general constraint

$$[x | x \models \langle \text{kissing}, y, z \rangle] \Rightarrow [x | x \models \langle \text{touching}, y, z \rangle]$$

The collection of specific constraints mentioned above can be recaptured since for any appropriate assignment $(a, b)$ for kissing the anchor taking $y$ to $a$ and $z$ to $b$ will transform the general constraint into the specific constraint

$$[x | x \models \langle \text{kissing}, a, b \rangle] \Rightarrow [x | x \models \langle \text{touching}, a, b \rangle]$$

The second problem for this kind of account of constraints cannot be so easily rectified. According to the above, perhaps all necessary constraints and some nomic constraints are true, but most nomic constraints and all conditional constraints are not. Of the sort of constraints which are relevant to cognition, the constraint ‘all rules have exceptions’ is the only exceptionless one. ‘Smoke means fire’ only if there are no artificial smoke generators nearby, ‘unsupported wine glasses fall’ near the surface of the Earth by not in
interstellar space, and utterances of 'cookie!' by a teasing sister rather than by mummy do not mean that there is a cookie on the way.

The ST solution to this problem is to generalize the notion of a constraint again to that of a conditional constraint. The idea is that some constraints come with 'background conditions' attached to them which specify when the constraint holds and when it does not. An organism attuned to a conditional constraint need not be aware of the background conditions at all. As long as the organism keeps to situations in which the background conditions are met all is well. If it strays into situations in which they are not met then there is no guarantee that the inferences it draws will be correct.

Conditional constraints are propositions of the form

\[ T_1 \Rightarrow T_2 | B \quad \text{and} \quad T_1 \perp T_2 | B \]

where \( T_1, T_2 \) and \( B \) are situation types. The type \( B \) is called the background type and, intuitively, it constraints which situations the constraint work in. The big problem for ST is to explain how conditional constraints constrain. The obvious approach is to say that (modulo the modifications concerning worlds)

1. If \( s \) is a situation of type \( B \) then if \( s \) is also of type \( T_1 \) and \( T_1 \Rightarrow T_2 \) (is true) then there is some situation of type \( T_2 \).
2. If \( s \) is an situation of type \( B \) then if \( s \) is also of type \( T_1 \) and \( T_1 \perp T_2 \) (is true) then there is no situation of type \( T_2 \).

But on this approach, if the conditional constraint \( T_1 \Rightarrow T_2 | B \) is true then there is no explanation why the unconditional constraint \( T_1 \land B \Rightarrow T_2 \) is not also true. Although there is no requirement that the truth of \( T_1 \land B \Rightarrow T_2 \) follows from the truth of \( T_1 \Rightarrow T_2 | B \), it is difficult to see how we can argue against it. One possible modification is to distinguish between the situation which satisfies the antecedent type \( T_1 \) and the 'background situation' which satisfies the background types. This would give us:

1. If \( s \) is a situation of type \( T_1 \) and \( T_1 \Rightarrow T_2 | B \) and the background situation \( b \) is of type \( B \) then there is some situation of type \( T_2 \).
2. If \( s \) is a situation of type \( T_1 \) and \( T_1 \perp T_2 \) and the background situation \( b \) is of type \( B \) then there is no situation of type \( T_2 \).
However, the relationship between the situation and the background situation is mysterious. We will return to this point later in this section and in Chapter 4, where a solution in terms of 'perspectives' will be offered and compared to a solution of Barwise.

At the beginning of this section we claimed that there were two ways in which a situation \( s \) could be classified by an infon \( \sigma \), which are given by the propositions \( s \models \sigma \) and \( s \models \lnot \sigma \). where \( C \) is a set of constraints. The information carried by a situation, \( s \models C \sigma \), is hard to define for an arbitrary set of constraints, \( C \). We can stipulate that if either of the constraints

\[
T \models [x](x \models \tau) \quad \text{or} \quad T \not\models [x](x \models \lnot \tau))
\]

are in \( C \) and \( s : T \) then \( s \models C \tau \) (this assumes that either \( \tau \) or its dual is supported by some situation), but there may be other constraints in \( C \) which contribute to the information carried by \( s \) which are not of either of these forms. This is the best we can do without further principles governing the interaction of constraints.

It is instructive to briefly compare the ST account of 'carrying information' with Dretske's. We can view the specification of a communication systems as providing all the necessary ingredients for a ST account of the same informational relationships.

For each device, we are interested in the situation of the device's being in a certain state at a certain time. These situations are classified by a state system in the obvious way. Infons can be represented by an attribute-value system whose induced state system is isomorphic to this state system, so that the attribute-value classification relation determines which infons are supported. The only restriction is that the issue as to whether or not a particular infon is supported by a situation is resolved entirely by the state of the device in that situation. Now the probability measure on the communication system determines the information carried by a situation in accordance with Dretske's definition, i.e. the situation of device \( d \) being in state \( s \) at time \( t \) carries the information \( \sigma \) iff the conditional probability of \( \sigma \) given that \( d \) is in state \( s \) is 1. Since the infon is 'reducible' to a condition on the states of the devices in the system, this probability can be calculated.

The reverse construction is harder to see. How does the information carried by an arbitrary set of ST constraints determine a Dretskean analysis of the information flow
between situations? Barwise views the ST analysis of information flow as more general than Dretske's analysis because he does not see the need to reduce arbitrary ST constraints to talk about probabilities. This is true: there is certainly less information in a specification of the ST constraints in a communication system than in a probability measure. But Dretske does not use this information in his definition. He only refers to \textit{conditional probabilities equal to 1}. A specification of the system purely in terms of conditional-probabilities-equal-to-1 would do just as well for Dretske as a fully probabilistic specification. This specification would be just as 'objective', but more importantly it would be just as \textit{relativistic} as the probabilistic specification.

So what is the difference between a ST specification in terms of constraints and a revised Dretskean specification in terms of conditional-probabilities-equal-to-1? As far as the flow of information is concerned there is no difference. But there is an important difference in the status of the $\mathcal{E}_C$ relation. On the ST account this relation is determined absolutely. Given a classification induced by a scheme of individuation and a set $C$ of constraints, we can say what information is carried. But for Dretske, the information carried is crucially relative to the communication channel. This relativity is captured by the notion of a probability measure, since all the probabilities, including the conditional-probabilities-equal-to-1, have an implicit reliance on the communication channel working properly.

Again, we defer a detailed discussion of communication channels to Chapter 6., but note here a possible misconception. It might be thought that a communication channel is nothing more than a (possibly conditional) constraint. In the voltmeter example, the relevant constraint would be between the type of situations in which there is a potential difference of 6 volts across the resistor and the type of situation in which the voltmeter reads '6'. This constraint may also be an instance of a general constraint in which the voltage is parameterized. Although the communication channel underlies such a constraint, it is more specific since it links particular situations rather than situation types. Information flows from a particular situation at the resistor to a \textit{corresponding} situation at the voltmeter (where 'corresponding' means 'synchronous' in this example). Of course, we can approximate the communication channel by parameterizing the dimension of correspondence (in this case, time) in the situation types of the constraint.
CHAPTER 3. CONSTRAINTS

For example, we might approximate the channel by the general constraint between the parametric type of situations in which there is a potential difference of $v$ volts across the resistor at time $t$ and the parametric type of situations in which the voltmeter reads $v$ volts at time $t$. But this strategy will only work when the specific instances of the general constraint link individual situations: i.e. when each situation type is sufficiently precise that there is only one situation of that type.

In our view, this observation is the key to a proper treatment of conditional constraints in ST. The trick of including a 'background type' does not work because it ignores the real cause of exceptions to constraints: the breakdown of the communication channel. We can see an analogous device being employed by Dretske to account for the relativity of information flow to the knowledge of an observer. The parameter $k$ works just like background types. $k$ shifts the probability measure. Background types restrict the antecedent. Both have the same effect on the flow of information. But as we have discussed in the previous section, the parameter $k$ is no substitute for the channel conditions since, in general, these cannot be expressed as a measurable event in the probability space. Likewise, the background type is no substitute for the conditionality of conditional constraints.

3.3 Attunement

In order for the concept of a flow of information to have any use in explaining cognition there must be a way of understanding how an organism can utilize the flow of information to its own ends. Barwise uses the term 'situated inference' (in his 1986) to describe the process by which an organism, a computer, or even a simple control device, like a thermostat, gains information from the flow. The important difference from usual conceptions of inference is that many of the factors which may be necessary to guarantee the validity of the inference can be omitted simply because the organism is situated in the world. The fact that an organism is in a particular situation may be enough to ensure that certain constraints hold even though they do not hold in general.

A good example (from Barwise 1986a), is the constraint “if Claire rubs her eyes then she is sleepy”, used by the author as a way of telling when to put his daughter to bed. In general the constraint does not hold: when there is a high pollen count, for example,
CHAPTER 3. CONSTRAINTS

Claire will rub her eyes even though she is not in the least bit sleepy. However, in situations when the pollen count (and any other potentially misleading factor) does not get in the way, it enables the flow of information and validates the inference.

In Barwise and Perry's (1981) terminology, situated inference is a matter of being attuned to constraints which hold in the situation. Following a constraint is not enough to ensure that one's inferences are correct. It may be that some other (unrepresented) factor, like a high pollen count, gets in the way. But, if we just happen to be in an area of low pollen count then we will make our inferences successfully without ever having to know about the effects of pollen on Claire's eyes. To a certain extent attunement comes down to being in the right place at the right time.

Aside from the effects of conditional constraints, there is another potential difficulty with the notion of attunement given the weak logical structure of constraints. Modulo the modifications discussed in the previous section, the logical properties of an 'involves' constraints can be stated simply. If a situation is of type T1 and T1 involves T2 then there is some situation of type T2. So if an organism in situation s1 classifies it as being of type T1 and is attuned to the constraint T1 \(\rightarrow\) T2 then it can rely on their being a situation of type T2. But to use this information successfully it must surely know where s2 is! That there is a T2 situation out there somewhere will not be of much use.

For humans inference it seems that it does not really matter which situation is of type T2 as long as there is one. But this is an illusion due to use of a highly expressive classification system: language. As long as we know that a certain statement is true it does not matter which situation makes it true. It is this view of language that Situation Semantics is in complete opposition too. Nevertheless, the relative context independence of human language as compared with simpler systems of classification only serves to confuse the issue when it comes to talking about attunement. Consequently we will use an example of a much more primitive information processor to explore the idea of information flowing and of an organism being attuned to that flow.

Our example is constructed around the feeding behaviour of a fictitious unicellular animal called a Locipod. Locipods have thousands of tiny threads, or cilia, attached to their cell walls which beat rhythmically to propel them through fluid in which they live. They feed by simply absorbing useful proteins through their cell walls. The cell
CHAPTER 3. CONSTRAINTS

wall is a delicate mechanism which requires a fine chemical balance across it if it is to work properly: allowing food to pass through whilst preventing toxins from entering the cell body. Chemical $X$, which is present in the environmental fluid in varying concentrations, disturbs this balance. To adjust for a change in concentration of $X$ in the immediate surroundings, a locipod can initiate chemical changes in the fluid in its own cell body to restore the necessary balance across the wall.

We suppose that the internal changes are controlled by production of a chemical, $Y$, in such a way that the necessary balance is achieved when the concentration of $Y$ inside the cell wall is in a certain proportion to the concentration of $X$ outside. If the organism is functioning correctly in a normal environment, the balance across its cell wall will be maintained and feeding will continue safely. If, however, the balance is upset then at best feeding will be inefficient and at worst the animal will die from toxins crossing its cell wall.

The problem for a locipod is therefore to ensure that the correct adjustment of internal fluid composition occurs to match the changing external concentration of chemical $X$. This is made difficult because of the time delay between initiating the appropriate internal changes and the balance being restored. In order to avoid disaster the locipod must anticipate changes in $X$ concentration. We assume that the animal has two kinds of corrective action: either increase or decrease production of $Y$. If external $X$ concentration is going to increase then $Y$ production must be increased to match it and if $X$ concentration is going to decrease, $Y$ production must also decrease.

The space in which our locipod swims at a constant speed (of $d$ units per unit time step) is taken to be two-dimensional for simplicity of exposition. The concentration of $X$ is determined by microscopic plants in the environment. We suppose that the plants are randomly distributed throughout the plane but are separated from each other by a distance significantly larger than $d$. The concentration of $X$ varies from plant to plant and decreases with the distance from the nearest plant.

To give a simple model of the locipod's environment we divide the plane in which it moves into a grid of squares, length $d$. The animal's movements are approximated to lateral and vertical movements in the grid so that in any one time step it moves exactly one square North, South, East or West. The concentration of $X$ is monitored
CHAPTER 3. CONSTRAINTS

Figure 3.3: $9 \times 9$ region of the locipod’s environment.

at each edge that it crosses. To survive the locipod must be fairly good at predicting the concentration of $X$ one time step ahead, i.e. at the next edge that it crosses. If, for example, it is crossing an edge where the $X$ concentration is 3.5, it may predict that the concentration will increase to 4.5 over the next square. Internal $Y$ concentration should then be increased to balance this predicted level.

There are several additional assumptions that we make about the locipod’s environment. We suppose

- that the change in $X$ concentration is reasonably smooth (in our model this is interpreted as requiring that the average $X$ concentration in each square changes by at most one unit), and
- that the change in $X$ concentration is reasonably frequent: there are no plateaus or rapid fluctuations (in our model this will be interpreted as requiring that the average $X$ concentration in each square changes by at least one unit).

An example model of a $9 \times 9$ region of the locipod’s environment is depicted in Table 3.3. The figures in each square represent the average $X$ concentration. In line with our assumptions, the change in the average concentration of $X$ from one square to a neighbouring square is exactly one unit. If, for example, the concentration of $X$ in the current square is 4 units then the concentration of $X$ in neighbouring squares is either
3 or 5 units. It is important to note that these are not the figures monitored by the locipod. Monitoring takes place as the animal crosses an edge. The $X$ concentration at an edge will be taken to be the average of the concentrations in each of the bordering squares. For example, proceeding down the first column, the locipod would monitor the $X$ concentration as

![Concentration Values]

We will review the dilemma facing a locipod at a particular point in its journey. As it enters a square of the grid, it has various choices to make. Firstly, it must decide which direction to go in: straight on, left, right, or back the way it came. Secondly, it must decide what to do with its production of $Y$. Should it increase $Y$ production in anticipation of higher $X$ concentration, or decrease it, predicting a fall?

There is little information immediately available which will help. If the last monitor result was, say, 3.5 then the average concentration of $X$ in the current square is either 3 or 4. So all one can say about the next edge it crosses is that the $X$ concentration is between 2.5 and 4.5. At least the concentration in the previous square is confined to being 3 or 4, so perhaps the best strategy would be to move forwards and backwards across that edge. We discount this as a long term strategy on the assumption that the food in the area would soon be used up and our locipod would be forced to move on.

Clearly some use must be made of past experience. If the locipod has spent a long time wandering around the region and has a capacity for storing the results of its monitoring, it is possible that it could make some use of its knowledge of the environment. Perhaps it has a ‘cognitive map’ showing the chemical landscape of the area. Such a map would not, of course, be complete, but it would contain a large amount of information which would constrain the possible $X$ concentrations in uncharted squares. Certain inference procedures could infer from the map which was the safest way to go and what concentration of $X$ to expect.

Figure 3.4 depicts the results of a survey of the whole region in graphical form. The boxes in the diagram represent the edges between the squares. The higher the concentration of chemical $X$, the darker the boxes are shaded. The lines between the boxes represent
Figure 3.4: A Locipod's Monitoring of the Environment.
CHAPTER 3. CONSTRAINTS

the paths from one edge to another, so that the squares of the region are represented by
the points where the lines cross. We can think of this diagram as a representation of the
total knowledge the locipod could possibly have about its environment. It is made up of
slopes, peaks and troughs in a landscape of chemical gradients. A locipod in possession
of such a map would only have to keep track of its current position and adjust \( Y \)
production according to whether the next edge has a higher or lower concentration of
\( X \) than the last edge crossed.

There is an obvious problem with the hypothesis that the locipod has such a map.
Firstly, it would have to learn the map by conducting an exhaustive survey of \( X \) con-
centration in the region. To do this, and survive, it would have to have another, albeit
temporary, strategy for maintaining a balance across its cell wall whilst it was conduct-
ing the survey. If the temporary strategy was good enough then it would not need a
cognitive map at all. If not then it would probably die before completing the survey.

We can meet this objection by supposing that, at any time, the locipod has a partial map
of the region recording the \( X \) concentration at edges that it has visited. A representation
of its acquired knowledge might therefore consist of a continuous portion of Figure 3.4
together with a 'marker' for its current position. If it moves into a part of the region
that it has visited before it can rely on the map. If it moves across a new square then
there are certain regularities in the way \( X \) concentration changes which it may be able
to use to calculate the expected concentration at the next edge. Moreover, it would
be able to use the partial map to decide which is the safest direction to go in without
doubling back on its tracks. As it monitors the \( X \) concentration at each edge, it can
incorporate the newly acquired information into an extension of its partial map so that,
eventually, it will build up a complete map of the region.

But now suppose that the \( X \) concentration in the region is changing. It need not be
changing very fast, say, at most one unit every ten time steps. Now the locipod must
do even more if it is to have a correct representation of its environment. Information
acquired more than ten time steps ago is no longer guaranteed to be correct, so it
cannot simply add incoming information to its partial maps; it must also delete or
correct information that is older than ten time steps. This requires some representation
of the age of parts of its map. If the animal is to create maps larger than ten squares,
it must adopt an even more sophisticated approach, somehow monitoring the rate of change of concentration in different locations, that is assuming that the concentration is changing at a constant rate, ... 

The point of the story is simple. If the locipod is to simultaneously solve the problem of which direction to go in and what to expect when it gets there, it needs a fairly sophisticated representation of its environment. The handling of such a representation is likely to be computationally complex. It is also clear that even slight changes in the environmental conditions require further representational and computational complexity to correct for. It cannot ignore these changes because its strategy depends on the assumption that its representation is correct. Even slight changes can accumulate to invalidate its representation.

However, we need not abandon the locipod to extinction from fictional evolutionary forces. There is a very simple strategy it can adopt providing it ignores the choice of where to go. We will suppose that the locipod moves in a more or less constant direction across the plane. In our model, this will be interpreted as movement in a vertical or horizontal direction, but in practice a wide arc would do. What is important is that the direction is constant locally.

The computational solution that the locipod can adopt is represented by the flow chart in Figure 3.5. The concentration of X is monitored each time step as it crosses an edge (although the solution generalizes to continuous monitoring). The value of the result of each sample is passed (as value c) to a unit that can test equalities. The same value is passed to the same unit after a delay of one time step (as value p). The test unit checks to see if the current sample value, c, is equal to the previous value, p. If they are not equal then nothing more happens. But if they are equal the value of an internal
CHAPTER 3. CONSTRAINTS

register is flipped from 0 to 1 or vice versa. Production of $Y$ is controlled by the value of the internal register. If it is 0 the $Y$ production will be increased. If it is 1 then $Y$ production will be decreased.

In this way, the locipod can be seen as a two-state device. In one of its states (corresponding to the internal register having the value 0) it constantly increases $Y$ production at a rate of one unit per time step. In its other state (corresponding to the internal register having the value 1) it constantly decreases $Y$ production at the same rate. By monitoring $X$ concentration in the way suggested, the locipod only changes state when the level of $X$ concentration is the same at its current position as it is at a distance $d$ in the direction that it came from. Now, supposing that the locipod is traveling in a (locally) straight line, this can only occur if it has just passed over a peak or a trough in the chemical landscape. If it has just crossed a peak then decreasing $Y$ production is the right thing to do for the proper functioning of its cell wall. Moreover, it continues to be the right thing to do until the locipod reaches a trough. After the trough, it should increase $Y$ production, which it does, until the next peak. Assuming that each locipod has a 50 per cent chance of starting off in the right state, this strategy should ensure that about 50 per cent of locipods survive.

Not only is the strategy very simple, it is also less effected by small changes in the environment. It does not matter to this breed of locipod whether the monitoring it did ten time steps ago is no longer accurate since its strategy only depends on its current and previous sample.

The morals concerning attunement to the flow of information are, firstly, that it really can be about being in the right place at the right time. Secondly, that increasing your knowledge of the regularity by including more specific information in the constraint is not necessarily going to help and computationally is likely to be much worse. And, thirdly, that this applies just as much to the problem of finding the ‘right situation’ as it does to the problem of satisfying the conditions on a conditional constraint. These are different problems, but they have the same solution.

\footnote{The reader should check that this is in fact so by trying a few traverses of Figure 3.4.}
3.4 Digitalization, Individuation and Cognition

In this section we consider a suggestion of Devlin's (1988) that the individuation of facts by a cognitive agent can be explained, or at least elucidated by, Dretske's analog/digital distinction. This is of great interest to us since it purports to shed some light on the connection between schemes of individuation and constraints. Although we shall reject Devlin's suggestion, the discussion reveals the importance of the relativistic aspects of Definition 3.1. We claim that this relativism needs to be incorporated into the way constraints work in ST if the theory is to make any sense of the idea that objects, properties and relations arise as uniformities across situations.

3.4.1 From Kapok to Bits

In Chapters 1 and 2 of his book 'Logic and Information', Devlin relates the situation theoretic account of constraints to Dretske's notion of information flow. He regards Dretske's theory as giving a picture of the world as having a complex, continuous, even 'fractal', information structure from which cognitive agents 'extract' certain facts, individuate objects and properties and so on.

"In general, the situation facing the agent will be host to a whole spectrum of information – information that the agent may or may not be capable of picking up. In order to function in an appropriate fashion, the agent has to be able to extract from this ‘continuum’ of information those basic items (facts) it requires for whatever purposes it has. Any further information resident in the situation, information the agent is not equipped to pick up (i.e. cannot form a digital representation of) may, as far as the agent’s cognitive activity is concerned, be ignored – it might as well not be there. This means that in developing our theory of the cognitive activity of the agent, what counts is the collection of facts that may be picked up from a situation. ... Thus we are concentrating on those aspects of the situations that are eligible to play an active role in the agent’s cognitive activity with

6The work presented here is a minimally revised version of joint work carried out with Nick Chater and reported in Seligman and Chater (1989). Chater’s subsequent thesis (Chater 1989) contains a far more detailed and extensive treatment of these and other related issues.
CHAPTER 3. CONSTRAINTS

respect to a situation, adopting very much an agent relative, rather than an absolute view of reality." (Devlin 1988, p.21)

Note that the information 'resident' in a situation which is to be excluded from the theory, is excluded on the grounds that 'the agent is not equipped to pick (it) up'. Now surely if an agent is capable of individuating a situation \( s \), a relation \( r \), and (appropriate) individuals \( a_1, \ldots, a_n \), then that agent is certainly 'equipped to pick up' the information that \( a_1, \ldots, a_n \) stand in the relation \( r \) in \( s \). So Devlin is not excluding any information constructed from the primitives of ST. The additional information 'resident' in the situation must be information about other individuals not individuated by the agent, or else some other kind of information altogether. So the infons extracted from the kapok are just those that are provided by a scheme of individuation.

The interesting point is that Devlin offers us an alternative way of characterizing the information that ST addresses. It is that information that the agent can 'form a digital representation of'. The claim seems to be that the information constructed as infons from the primitives individuated by an agent can be equally well characterized as the information that the agent can form digital representations of.

This is not to say that for a particular situation the agent has digital representations of all the infons constructed from the primitives that the agent individuates from the situation, only that those infons are exactly those pieces of information that the agent could represent digitally: other pieces of information about the situation are distinguished from those infons because the agent is not capable of representing them digitally.

The key question is whether this distinction is able to tell us about the individuation of objects and properties. Intuitively, the ability to individuate the property (one place relation) of being a chair seems to be closely related to the ability to recognize chairs. For Devlin, this crucially involves digitalization: in recognizing an object \( a \) as a chair, some state of the agent carries the information that \( a \) is a chair in digital form. On this view, the individuation of a property, \( p \), is a matter of recognizing (in general) that an object has \( p \), and that is a matter of digitally encoding the information that the object has \( p \). So, the analog/digital distinction promises to help explain what it is to be situation, relation or object (as individuated by an agent).
Typically, digitalization involves a huge loss of information—indeed, only one item of information is digitalized. Further, given that any item of information has indefinitely many further items nested within it, Devlin’s description of the structure of information as ‘fractal’ is intuitively appealing. For example, the information that something is a rectangle is nested within the information that it is a square; the information that something is a plane figure is nested within the information that it is a rectangle. It seems that we can iterate at will from plane figure to geometric figure, to geometrical object, to mathematical object and so on. The notion of digitalization allows us to abstract away from the fractal structure and focus on a single piece of information.

Devlin’s suggestion, encapsulated in the similarity between Figure 3.6 (a) and (b), is that the process of individuation is to be explained as a process of digitalization of an analog reality. An agent can individuate, or ‘extract’ just those items of information that he can digitalize. To individuate the property ‘red’ an agent must be able to detect red things, which, at the very least, involves carrying the information that a certain object is red in digital form.

We will examine the extent to which the detection of such informational uniformities can be captured using Dretske’s theory in the next section. But if such an account can be given, we are lead to some exciting conclusions. It would mean that the building
CHAPTER 3. CONSTRAINTS

blocks of ST – individuals, relations and infons – can characterized in purely (Dretskean) informational terms. Dretske’s theory of information would provide the ontological underpinning of ST. Moreover, Devlin claims that Dretske’s primary notion of information flow can be characterized by the Situation Theoretic notion of a constraint. Rather than providing an explanation of information flow in probabilistic terms, ST treats informational constraints as given. So, prima facie, Dretske’s analog/digital distinction can be framed in ST as a condition on constraints. If this is really true then an explanation of the individuation of the ‘primitives’ of ST can be given within the theory. Such a projects hangs dangerously on the edge of circularity. If constraints are constructed as relations between situation types (themselves constructed out of infons), we are doomed. Take the regularity between smoke and fire. Because ‘smoke means fire’ is indeed a constraint, a smoky situation carries the information concerning the proximity of fire. But this constraint can only be identified as such a relation between ‘smoke’ and ‘fire’ if we have a prior understanding of these categories. An account of the individuation of ‘smoke’ and ‘fire’ in terms of what an agent can digitalize must therefore omit any reference to this constraint. But every constraint identified in this way must assume prior understanding of some such categories. So if constraints are identified as relations between situation types, a complete theory internal explanation of individuation is impossible. If the status of individuals and relations is to be explained within the theory then something else must bear the ontological burden that they have traditionally assumed. Constraints must be truly primitive if ST is to be a mathematical theory of information which is able to characterize individuation.

3.4.2 From Digitalization to Cognition

We have assumed that the notion of digitalization may be able to explicate the ability of agents to individuate the world into objects and properties – that is, to arrive at an appropriate categorisation for cognitive activity. Indeed, this ability to conceptualize

---

Devlin, however, notes that "it is with information in digital form that we shall be primarily concerned" and views infons as idealizations of 'digital information' alone. If this is the case then the analog/digital distinction could not be expressed for the trivial reason that there is no 'analog information'. But information can only be 'digital' with respect to certain constraints. Quite what is meant by a restriction of the subject matter of the theory to 'digital information' alone is unclear.
CHAPTER 3. CONSTRAINTS

the world could be seen as distinctive of cognitive agency.

"... the extraction of information from the environment by agent can be regarded as taking place in two separate stages, corresponding to the analog/digital distinction. The first stage is perception, where the information in the environment becomes directly accessible to the agent by way of some form of sensor ... At this stage the information flow is an analog one (relative to whatever information we are concerned with). The second stage (if there is one!) involves the extraction of a specific item (or items) of information from that perceived 'continuum'; that is to say, it involves the conversion from analog to digital information. This stage is cognition." (Devlin 1988 p.14)

Many examples seem to suggest that the analog/digital distinction and the perception/cognition distinction are closely related. When looking at the armchair in my office I may suspect that the information that the armchair is red may perhaps be digitally encoded in some state of my head, whereas this information is encoded in analog form on my retina. Yet if this account relies upon the intuition that retinas are only capable of carrying information in analog form then it is surely mistaken. The retina does carry some information in digital form; trivially, for example, a characterization of the state of the retina itself. This is because any information carried by the retina's being in a certain state is, by definition, nested within the information that the retina is in that state. The point is that analog/digital conversion is only relative to a particular piece of information. In the previous example, the relevant piece of information is that the armchair is red.

Although Devlin acknowledges this relativity, he defines what it is to be a cognitive agent in absolute terms: 'A cognitive agent is an agent that has the capacity of cognition in this sense; i.e. the ability to make the analog to digital conversion.' (p.14)

Given that analog/digital conversion is relative to a piece of information rather than being an absolute distinction, there is no sense in which the above can be a definition

---

On Dretske's definition the granularity, or otherwise, of the retina is not relevant to whether or not the information is coded in analog form. Rather it is only relevant to whether the information is represented discretely or continuously.
CHAPTER 3. CONSTRAINTS

of cognitive agency. As the above example shows, every physical object would be a
cognitive agent relative to some pieces of information, and a non-cognitive agent relative
to others, on such a definition. Devlin admits that this definition is not enough to
fully explicate the perception/cognition distinction. Perhaps further conditions must
be added.

That a bucket contains water (rather than ice or steam) carries the information that
the temperature is between 0 and 100 degrees Centigrade in digital form. Under such
a characterization the bucket is a device with three states: contains water; contains
ice; contains steam. That the bucket is in one of these states digitally encodes some
information about the temperature, that a more sensitive instrument, such as a ther-
nometer, would encode in analog form. A thermometer placed in the water reading 22
degrees centigrade would carry the information that the temperature is between 0 and
100 degrees in analog form. So under this specification, the bucket performs analog to
digital conversion. But we do not say that the bucket of water exhibits cognition. So
although we have specified a particular piece of information (that the temperature lies
between 0 and 100 degrees Centigrade), we still find our definition to be inadequate.

How could it be extended to rule out the bucket as a cognitive agent? Perhaps the
bucket, although digitally encoding some information does not digitally encode the right
kind of information for cognition. In what way does the kind of information that we, as
cognitive agents, process differ from the kind of information in these trivial examples?
One sort of act that we regard as cognitive is the recognition of everyday objects: the
information that something is a chair would seem to be of the right kind. Surely human
beings (and perhaps domestic pets) are the only agents to interact with chairs qua
chairs. Many organism and even inanimate objects will causally interact with chairs
qua physical objects. A rolling ball may stop when it contacts a chair, but not because
it is a chair; it responds to the chair only as a heavy rigid object. In contrast, a person
might stop by the same object with the intention of sitting on it. In this case the person
is responding to the property of being a chair. Since people are our paradigm example
of cognitive agents, and only people are able to respond to such everyday properties as
the property of being a chair, it seems that at least one example of the right kind of
information is that something is a chair.
Prima facie we have at least made some progress towards characterizing cognition. All we now have to do is to specify the kinds of information which cognitive agents process. Information about everyday categories like chair-hood seem to be good candidates. We might say that only cognitive agents can 'extract' the information that an object is a chair from the environment. There seem to be no plausible non-cognitive chair recognizers. Indeed such tasks as chair recognition are notoriously difficult for Artificial Intelligence programs. To make this story precise, it would be necessary to explain exactly what distinguishes the right kind of information from the wrong kind, other than that only the right kind is encoded by cognitive agents. Even if such an explanation could be formulated (an extremely difficult task in itself) there are further problems. We will argue that the analog/digital distinction is relative to more than the piece of information under consideration, be it of the right or the wrong kind. Because of this, the corresponding definition of cognition is sufficiently pliable to allow such patently non-cognitive objects as mirrors and photographs to count as chair recognizers.

The analog/digital distinction relies on a notion of aboutness. The only information carried by a signal relevant to determining whether the information that s is F is coded in digital or analog form is information about s. Devlin's illustration of the analog/digital distinction is the difference between the utterance "Jon is taller than Fred" and a photograph of Jon and Fred standing next to each other. The photo carries the information that Jon is taller than Fred in analog form, whilst the utterance carries the same information in digital form.

The photograph of Jon and Fred carries the information that Jon is taller than Fred in analog form because it carries additional information about Jon. It also carries additional information about the library in which they are standing. Perhaps the photograph shows the library's vaulted windows. But in the example this information is not relevant to the analog/digital distinction because the photograph is tacitly understood to convey information about Jon and not about the library. Similarly, the utterance 'Jon is taller than Fred' may be said to carry the information in digital form only because it is tacitly understood that none of the additional information carried by the utterance (for example, its pitch and volume) is information about Jon.
CHAPTER 3. CONSTRAINTS

The presence of information carried by a signal which is irrelevant to making the analog/digital distinction is by no means exceptional. This information is noise in the system. Recall that noise is information at the receiver that is not generated by the source, i.e. it is information that is not about the source. For information of the type $s$ is $F$ we seem to have an obvious criterion for distinguishing signal from noise: information that is about $s$ is part of the signal, information that is not about $s$ is noise.

In the example of the utterance 'Jon is taller than Fred' it seems clear what information is carried about Jon. Unfortunately, this is rather more than one might first expect. That Jon is being talked about, Jon is being talked about loudly, and that Robin is talking about Jon, are all pieces of information about Jon, none of which are nested in Jon's being taller than Fred. Once these pieces of information are admitted as bona fide information about Jon, we have the counter-intuitive conclusion that the utterance no longer carries the information that Jon is taller than Fred in digital form.

If we accept such additional information as being about Jon, the analog/digital distinction has failed to capture our intuitions as to the different ways in which an utterance and a photograph carry information. To block such counter-examples we need some characterization of what properties of Jon are properties about Jon, in the requisite sense. We need a more restrictive notion of aboutness, to exclude the above examples as not being genuinely about Jon. Let us suppose that we have such a restriction. This would pick out some subset, $S$, of the properties of Jon which are really about Jon. In particular it would include the property of being taller than Fred, but reject the property of being talked about. Only properties in this subset would be relevant to determining whether or not a given piece of information is carried in digital or analog form. That is, in making analog/digital judgements about a piece of information of the form $s$ is $F$ carried by a signal, we must ignore all information of the form $s$ is $G$, where $G$ is not in $S$. Yet by excluding all such $G$ from consideration, we ensure that the information that $s$ is $G$ can never be carried in digital form. On this view, the utterance 'Jon is being talked about' cannot, in principle, digitally encode the information that Jon is...

---

*It is important to note that there is nothing special about simple acoustical properties of the utterance, like pitch and volume. We can construe any property of the utterance (e.g. that it is $G$) to be (indirectly) a property of Jon (Jon is being talked about On). For example, $G$ might be the property of having been made 40 miles from Glasgow.
CHAPTER 3. CONSTRAINTS

being talked about (and neither can anything else). If having the concept $G$ requires the ability to carry the information that $s$ is $G$ (for some $s$) in digital form, then we cannot have the concept $G$, if $G$ is not in $S$. In particular, we cannot even have the concept of being talked about, let alone believe that Jon is being talked about. Thus any subset $S$ is too strong a restriction to be seriously considered.

The problems are not only with properties which fall outside $S$. Allowing that we have made such severe restrictions, let us consider what properties might be left inside $S$. Since we want the utterance 'Jon is over 6 feet tall' to carry the information that Jon is over 6 feet tall in digital form, we must at least allow the property of being over 6 feet tall to fall within $S$. If there is to be any principled criterion for membership of $S$, then surely it must at least include such non-relational, physical properties. Nevertheless this is still too weak to maintain our intuitions about what is coded digitally.

Imagine a black box lying on a table in Robin's office. Robin says 'The black box is a tape recorder'. Intuitively, this utterance carries the information that the black box is a tape recorder in digital form, in the same way that 'Jon is over 6 feet tall' carries the information that Jon is over 6 feet tall in digital form. But suppose that the tape recorder was recording Robin's utterance. The state of the magnetic tape can be specified in purely physical terms. Suppose that the tape is in state $T$ having recorded Robin's utterance. Then the black box has a non-relational physical property $P$, of its tape being in state $T$. Moreover, Robin's utterance carries the information that the black box is $P$. By our minimal assumption, the property $P$ is in the subset $S$. We must conclude that the utterance carries the information that the black box is a tape recorder in analog form. In fact, whether or not the information is analog or digital depends on whether or not the tape recorder is on.

We have established that no general way of distinguishing between properties which are genuinely about an object from those which are not will comply with what is needed to make utterances encode their contents in digital form. The underlying problem in the tape recorder example is that utterances have the power to causally interact with the events that they are describing. The same argument will go through for any other information bearing object with such causal powers (e.g. mental states). Hence, the attempt to use the analog/digital distinction as a basis for distinguishing cognitive from
non-cognitive agents has so far proved unsuccessful.

The only remaining option is to say that which properties are members of $S$ is decided for each specific informational occasion. That is, in order to get a satisfactory notion of aboutness, we must specify which properties of the object $s$ are relevant to determining whether the information that $s$ is $F$ is carried by a signal in digital or analog form. In Dretske's account, this specification is built into the very definition of information. The source in any informational exchange is defined as a set of mutually exclusive and exhaustive states. The properties of being in each of these states are precisely those required. According to a specification which only gives that the black box is a tape recorder, the information that the black box is a tape recorder is carried by Robin's utterance in digital form. But according to a specification in which the audio signal on the tape in the black box is given, the information is carried in analog form.

So far we have seen that for the purposes of drawing a distinction between perception and cognition, cognitive and non-cognitive activity, it is crucially important to recognize that the analog/digital distinction is inherently relative. It is not only relative to the piece of information we are considering, but it is also relative to the specification of the source. Even this is not enough. The analog/digital distinction is also crucially relative to a specification of the signal.

Devlin introduces a useful example for illustrating this point in contrasting the thermometer (non-cognitive) and the thermo-stat (cognitive) on the basis of the analog/digital distinction:

"A thermometer simply registers the temperature (analog coding of information) and hence is a perceiving agent but not a cognitive agent...: the thermo-stat classifies the temperature into two classes ('warm' and 'cold'), and thus exhibits a form of cognition."

Saying that the thermo-stat codes the information that it is cold in digital form amounts to saying that the thermo-stat's being in the 'on' position carries this information in digital form. It is the fact that a receiver is in a particular state (a signal) which carries information about the source, not the receiver itself. There are other facts about the thermo-stat which carry more information about the temperature (e.g. the stress on the
bimetallic strip, and even the temperature of the thermo-stat itself). These facts carry the information that it is cold in analog form. Thus any appeal to the information coded in an object makes implicit appeal to a specification of the relevant facts about that object. Similarly, saying that 'a thermometer simply registers the temperature (analog coding of information)', is to say that the fact that the thermometer is in a particular state carries specific information about the temperature in analog form. For example, the fact that the mercury is at the 5 degrees Centigrade mark carries the information that it is cold in analog form. However, the fact that the mercury is below the 10 degrees Centigrade mark carries the information that it is cold in digital form. Whether or not the information that it is cold is carried in analog or digital form by a thermometer is dependent on what we take the states of the thermometer to be. This amounts to a specification of the signal. As before, this specification is built into Dretske's definition of information content. Hence, if we are to maintain Devlin's intuition that there is some important way in which the information processing capacity of the thermo-stat is more akin to cognitive activity than that of the thermometer, we must recognize the relativity of the analog/digital distinction to a specification of the signal.

Carrying this over to Devlin's characterization of cognition, we find that an act is cognitive only relative to

- the information carried
- a specification of the source (what the information is about)
- a specification of the signal (in this case the state of the cognizer)

To return to problems of chair recognition, it may seem that a mirror is at best a perceiving object. But if this is so, then when a chair is reflected in the mirror, the information that the reflected object is a chair is carried by the state of the mirror in analog form. However, suppose that the mirror is specified as a receiver with two states: \( A \), reflecting a chair; or \( B \), not reflecting a chair. On this specification the fact that the mirror is in state \( A \) carries the information that the reflected object is a chair in digital form. Hence, the mirror demonstrates hitherto unsuspected cognitive powers.
CHAPTER 3. CONSTRAINTS

3.4.3 Constraints before Individuals?

Clearly, that an agent is in some state which carries the information that \( s \) is \( F \) is not sufficient for that agent to know or believe that \( s \) is \( F \), or have the concept \( F \). A first move to bridge this gulf is the introduction of the analog/digital distinction. 10 Both Devlin and Dretske acknowledge that there is more to cognition than digitalization and there seems to be some agreement as to what must be added. Devlin cites a personal communication with Dretske:

"... provided the agent has the means of manipulating and utilizing the information it obtains, then the digitalization of perceived information is the essence of cognitive activity." (Devlin 1988 p.15)

As we have seen, for any information carried by a receiver there is always some specification of that receiver as being in a certain state (i.e. some signal), under which the information is carried in digital form. Let us call a property of being in such a state a digitalizing property.

If the state of thermo-stat is specified by its being in the 'on' position, then it carries the information that it is cold in digital form. But if the state of thermo-stat is specified by its bimetallic strip's having a certain stress (which is monotonically related to the temperature) then it carries the information that the temperature is, say, 5 degrees Centigrade in digital form. The properties used to specify the state of the thermo-stat (being in the 'on' position; having such and such a stress) are both digitalizing properties. For the purposes of making the required step towards an elucidation of cognition, we want some way of distinguishing between good and bad digitalizing properties.

The difference in the above example seems to be that the heating system is responsive to the former property but not the latter. More generally, the right kind of digitalizing properties seems to be those which are suitable for understanding the larger informational system, in which the thermo-stat is embedded. These matters are discussed more fully in Chater's 1989.

10Israel and Perry (1987) have another approach. They attempt to characterize what it is for an agent to have information rather than to merely carry it, in terms of purposeful activity towards a goal.
CHAPTER 3. CONSTRAINTS

Given an optimistic attitude towards the prospect of characterizing the information processing capacity of organisms in situation theoretic terms, Devlin's incorporation of Dretske's ideas seems the right way to proceed. But more attention must be given to the assumptions on which Dretske's theory of information is based. We have seen that the project of using the analog/digital distinction to characterize cognition relies crucially on the relativity of Dretske's notion of information content. The specification of a source and a signal and certain assumptions about the communication channel are needed to determine whether or not a piece of information is carried in digital or analog form. We have suggested that the choice of specification should be made according to the information processing role of the receiver within a larger informational system.

Another aspect of the project which must be clarified is exactly how ST captures the relativity to specifications and the communication channel. We have already criticized the 'background type' approach as failing to capture the implicit relativity to the communication channel. In the next chapter we will suggest a solution.
Chapter 4

Perspectives

A perspective\(^1\) is a way of seeing some part of the world, or, more correctly, a part of the world seen from one point of view. If I look across the room at the black vase, I see it from a certain point of view. From my point of view the vase is in a certain perspective. Only someone looking at the vase from precisely the same point as me will see exactly the same thing. The perspective of the vase I see is not essentially subjective, but because of the normal geometry of space no one else can adopt this perspective at the same time as me.\(^2\) Whilst being relative to a point of view, the perspective has an objective quality.

The vase's perspective is not really a property of the vase alone, or even a relation between my point of view and the vase. My point of view determines the perspective of everything I am seeing. The vase is only seen by me in the way it is seen because it emerges as a visual object from its visual context. Visual perspectives are made up of edges and slopes in the visual scene, but these only emerge from interaction between different parts of the scene.

One of the original motivations behind the development of situation theory was to model visual scenes (e.g. Barwise 1981). A scene is a situation which supports information about what is seen. To incorporate a notion of perspective, we suppose that a scene is made up of many situations. Each situation supports visual information about their

---

\(^1\)Apart from a few revisions and omissions, this chapter together with Chapter 5 make up the paper *Perspectives in Situation Theory* - Seligman 1989.

\(^2\)Tricks with mirrors can allow more than one person to see something in the same perspective.
geometry, colour and whatever else is required for, or is implicit in, the informational organisation of our visual perception. The perceived objects and the relations they stand in are uniformities across these smaller situations and, most probably, many other situations which are not part of the scene.

Central to a model of perspectives, is a classification of situations into visual types. But more is involved in seeing than simply classifying things together. There are dependencies between the visual types which allow similarities to be noticed, expectations to be made, and explanations given as we construct an understanding of the scene as one in which some objects stand in various relationships with each other. The perspective in which I see the room is a classification of the various parts of the room, together with dependencies between the types used to classify.

We abstract the notion of perspective from visual classifications of the world to arbitrary classifications of situations, taking with us the qualities of objectivity and relativity which are more clearly manifested in the visual case. The dependencies between types will be modelled by binary relations. Each relation captures a particular kind of dependency with associated conditions which place restrictions on the classificatory role of the dependent types. Following ST, we say that a pair of types linked by a dependency relation is a constraint in the perspective. A perspective can therefore be thought of as a classification of situations together with a collection of constraints grouped into dependency relations.

On this broad definition, many of the classification systems from Chapter 1 would count as perspectives. However, in this chapter two fundamental kinds of dependency will be considered; the positive \( \Rightarrow \), or 'involves' relation, and the negative \( \Lambda \), or 'precludes' relation. We will find that these two kinds of dependency are widely applicable and give rise to a rich theory. Section 4.1 concerns a theory of perspectives based on these kinds of dependency. In Section 4.2 we compare constraints in a perspective with Barwise and Perry's 'conditional constraints' in which the applicability of a constraint is determined by a background type. Section 4.3 uses perspectives to model Dretske's (1981) theory of information flow, with the purpose of contrasting the perspectives used with those needed to model a logical conception of information content in Section 4.4.
4.1 A Theory of Perspectives

The conditions we associate with 'involves' and 'precludes' constraints are derived in an albeit modified form from Dretske, Barwise and Perry.

Facticity: the involves relation $\Rightarrow$ is fact preserving. If a situation $s$ is of type $t$ ($s : t$) and $t \Rightarrow t'$ then there must be some situation $s'$ of type $t'$ ($s' : t'$).

Xerox Principle: the involves relation $\Rightarrow$ is transitive. If $t \Rightarrow t'$ and $t' \Rightarrow t''$ then $t \Rightarrow t''$.

Local Preclusion: the precludes relation $\bot$ represents incompatibility. If a situation $s$ is of type $t$, $s : t$ and $t \bot t'$ then $s \not t'$.

Mutual Preclusion: the precludes relation $\bot$ is symmetric. If $t \bot t'$ then $t' \bot t$.

The Facticity condition is just the constraint satisfaction condition from ST, discussed in Section 3.2. Alternatively, in Dretske’s terms, the facticity condition expresses the veridicality of information. The Xerox Principle, discussed in Section 3.1.2, is Dretske’s. The two conditions for preclusion express the idea that $\bot$ is an incompatibility relation in the sense of Section A.1.3. We note two rival conditions for $\Rightarrow$ and $\bot$ which we have chosen not to adopt.

- if $s : t$ and $t \Rightarrow t'$ then $s : t'$ (Strong Facticity)
- if $s : t$ and $t \bot t'$ then there is no $s'$ of type $t'$ (Global Preclusion)

Both of these conditions are stronger than the ones we have chosen. Strong Facticity implies Facticity and Global Preclusion implies Local Preclusion. We will find examples of perspectives which satisfy one or other of these conditions, but they are not satisfied in the general case. The main distinction we will be concerned with in Sections 4.3 and 4.4 is that between ‘informational’ perspectives, in which the classification of one situation from a perspective carries information about other situations, and ‘logical’ perspectives, in which the dependencies provide closure and coherency conditions on the classification. The satisfaction of Strong Facticity or Global Preclusion will be found to distinguish
between these two kinds of perspective. Our four conditions are incorporated into the following definition of a perspective:

**Definition 4.1** A perspective is a structure \((S, T, :, \perp)\) where \((S, T, :)\) is a classification, ‘:’, of a collection \(S\) (of situations) by a collection \(T\) (of types), and \(\perp\) (involves) and \(\perp\) (precludes) are binary relations on \(T\), such that for all \(s \in S, t \in T,\):

1. if \(s : t\) and \(t \perp t'\) then there is an \(s' \in S\) such that \(s' : t'\) (facticity).
2. if \(t \perp t'\) and \(t' \perp t''\) then \(t \perp t''\) (xerox).
3. if \(s : t\) and \(t \perp t'\) then \(s \not: t'\) (local preclusion).
4. if \(t \perp t'\) then \(t' \perp t\) (mutual preclusion).

Given a particular perspective \(P = (S, T, :, \perp)\) we will refer to the situation domain \(S\) of \(P\) as \(S^P\) and the type set \(T\) of \(P\) as \(T^P\) in contexts where they have no other name, but we will rely on textual context to discriminate in which perspective \(s : t, t \perp t',\) or \(t \perp t'\) The symbols \(s, s', s''\) and \(t, t', t''\) will be used schematically over situations and types respectively.

**4.1.1 Information Content**

One sense in which information is localised at a particular situation \(s\) within a perspective \(P\) is given directly by the : relation of \(P\). We will say that information \(t \in T^P\) is located at, or grounded by, \(s\) if \(s : t\). The set of types grounded by \(s\) will be called the type set of \(s\), written \(\mathcal{T}\). We still retain the understanding of : as a classification of situations, and so we also speak of \(\mathcal{T}\) as the set of types which classify \(s\). Consequently, when \(\mathcal{T}\) is non-empty, \(s\) will be said to be classified and unclassified otherwise.

**Definition 4.2** For \(s \in S^P, \mathcal{T} = \{t \in T^P \mid s : t\}\)

Given that the same information \(t \in T^P\) can be located in different places, we will also make use of the set of situations \(s \in S^P\) at which \(t\) is located. This will be called the situation set of \(t\), written \(\mathcal{S}\). When \(\mathcal{S}\) is non-empty, \(t\) is said to be grounded and ungrounded otherwise. By the dual conception of : this set will also be referred to as the set of situations classified by \(t\).
A 'typical' perspective with some situation sets and type sets marked is illustrated in Figure 4.1. We extend the previous two definitions to sets of situations and sets of types. For $X \subseteq \mathcal{T}$, $X = \bigcup_{x \in X} x$, and for $R \subseteq \mathcal{S}$, $\bar{R} = \bigcup_{r \in R} r$. Thus $X$ is the collection of situations of some type in $X$ and $\bar{R}$ is the collection of those types which classify some situation in $R$.

**Proposition 4.4** $\mathcal{T}^P$ is the set of all classified situations in $P$. $\mathcal{S}^P$ is the set of all grounded types in $P$.

A second sense in which information is localised is given by the law-like relations on types, $\Rightarrow$ and $\perp$. The information $t$ is said to be carried by a situation $s$ within a perspective $P$ iff there is a type $t' \in \mathcal{T}$ such that $s : t'$ and $t' \Rightarrow t$. Similarly, the information $t'$ is said to be precluded at $s$ iff there is a type $t' \in \mathcal{T}$ such that $s : t'$ and $t' \perp t$. The sets of types carried or precluded by a situation are defined in terms of two operations on sets of types.

**Definition 4.5** For $X \subseteq \mathcal{T}$,

- $(X)^+ = \{ t \in \mathcal{T} \mid \exists x \in X x \Rightarrow t \}$
- $(X)^- = \{ t \in \mathcal{T} \mid \exists x \in X x \perp t \}$
For single types \( t \), write \((t)^+\) for \( (\{t\})^+\) and \((t)^-\) for \( (\{t\})^-\).

The sets \((X)^+\) and \((X)^-\) are referred to as the positive and negative completions of \( X \) in \( P \). The set of types carried (precluded) by \( s \) in \( P \) is therefore just the positive (negative) completion of its type set in \( P \). Sets of types \( X \) for which \((X)^+ \subseteq X ((X)^- \subseteq X)\) are said to be positive (negative) complete.

Proposition 4.6 \( T^P, \emptyset \), and \( \overline{3}^P \) are all positive complete. If \( t \in T^P \) is grounded then so is every type in \((t)^+\).

4.1.2 Inside Perspectives

Perspectives may be classified in various ways according to their internal structure. In particular, perspectives which satisfy additional requirements for their constraints will be studied as special cases. We will consider three ways in which perspectives may be given additional structure.

Restrictions on the Constraints

A plausible restriction to make is that \( \Rightarrow \) should be reflexive. This is a property of all logical consequence relations and is clearly justified by the extensional behaviour of \( \Rightarrow \) (the presence or absence of \( t \Rightarrow t \) places no restriction on the \( t \) relation). In situation theoretic terms, an insistence on reflexivity for \( \Rightarrow \) amounts to the claim that whatever information is made factual by a situation is also carried by that situation. The motivation for reflexivity is high from the logical point of view: if we know \( p \) about the world then surely we can infer \( p \). But this observation is quite out of place in an informational setting. If the receiver is in state \( r \) then this signal does not carry the information that the receiver is in state \( r \), since this is not information: it is part of the specification of the system.

Another sensible requirement is a version of xerox constraining the interaction of \( \Rightarrow \) and \( \perp \): if \( t \Rightarrow t' \) and \( t' \perp r' \) then \( t \perp r' \). However, the intuitive appeal of this requirement relies on the strong (global) sense of preclusion. Although not equivalent to the global preclusion condition it shares some of its extensional properties. Suppose, for example, that we have a perspective over a series of coin tossing experiments, suitably large in number so that the probability of all the tosses landing heads is \( \approx 0 \). Within our
perspective we will have heads $\Rightarrow$ tails, since up to our level of idealisation it is a law that if a particular toss lands heads some other toss must land tails. But we will also have that tails $\perp$ heads, since a given toss landing tails precludes that toss from landing heads. The interaction principle would then force heads $\perp$ heads, ensuring that heads was ungrounded. Yet we do not want to say that a toss of heads is impossible; in fact quite the reverse. The problem stems from our desire to have weak laws in such an application: for a positive constraint this reflects a global condition ("there is a toss landing heads"), whereas for a negative constraint the condition is local ("this toss didn't land heads").

Having made these arguments against the reflexivity of $\Rightarrow$ and the interactive form of transitivity, it should be noted that all the concrete applications of Sections 4.3 and 4.4 satisfy both principles.

Extensional Relations Between Types

The defining clauses of facticity and local preclusion express the relationship between the law-like relations $\Rightarrow$ and $\perp$ and relations between types which are defined extensionally (i.e., in terms of the : relation). For example, facticity states that the relation $\Rightarrow$ is included in the relation

$$\{(t,t') : t, t' \in TP, \forall s : t \text{ then } \exists s' : t' \}$$

We will examine the strengthenings of facticity and local preclusion in terms of various relations defined in this way.

**Definition 4.7** Extensional relations between types. For all $t, t' \in TP$,

- $t \gg t'$ iff $\forall s : t \text{ then } \exists s' : t'$ (constant conjunction)
- $t \parallel t'$ iff $\forall s : t \text{ then } s / t'$ (local inconsistency)
- $t \gg t'$ iff $\forall s : t \text{ then } s \gg t'$ (co-groundedness)
- $t \gg t'$ iff $\forall s : t \text{ then } \neg s \gg t'$ (global inconsistency)

(where the quantifiers range over $SP$)

Each of these relations can be characterised by a simple condition on the situation sets of $t$ and $t'$.
CHAPTER 4. PERSPECTIVES

Proposition 4.8 For all $t, t' \in \mathcal{T}^p$,

- $t \succ t'$ iff if $I = \emptyset$ then $I' = \emptyset$
- $t \iff t'$ iff $I \cap I' = \emptyset$
- $t \Rightarrow t'$ iff $I \subseteq I'$
- $t \Rightarrow t'$ iff either $I = \emptyset$ or $I' = \emptyset$

We can express facticity and local preclusion in terms of the above properties as: for all $t, t' \in \mathcal{T}^p$

- if $t \Rightarrow t'$ then $t \succ t'$ (facticity)
- if $t \iff t'$ then $t \iff t'$ (local preclusion)

and we have various means of classifying perspectives according to whether or not they satisfy the additional properties of

- if $t \Rightarrow t'$ then $t \succ t'$ (strong facticity)
- if $t \iff t'$ then $t \iff t'$ (global preclusion)
- if $t \Rightarrow t'$ then $t \Rightarrow t'$ (Hume's property (weak))
- if $t \succ t'$ then $t \Rightarrow t'$ (Hume's property (strong))

Strong facticity and global preclusion have been discussed in previous sections. Hume's properties are best understood as an abandonment of the distinction between accidental correlations and law-like dependencies. If whenever a situation is classified by $t$, there is another situation of type $t'$; the two pieces of information are constantly conjoined within the perspective. In general this is not enough to ensure that there is a law-like dependency between $t$ and $t'$. The strong version of Hume's property says that it is enough. Unlike Hume's condition for constant conjunction, there is no requirement that the second situation is contiguous with the first, or even that the observation is repeatable. The attribution 'Hume's property' (in the strong sense) follows Barwise and Perry (1983). The weak version says that $t \Rightarrow t'$ if any situation of type $t$ is also of type $t'$.

3) In other words, the strong version is the converse of facticity and the weak version is the converse of strong facticity.
A (weakly / strongly) humean perspective is one which has Hume's property (weak / strong). By default we will use 'humean' and 'Hume's property' to mean the weak version.

Special Elements of Perspectives

The third way of classifying perspectives is to isolate some elements with special properties. The significance of these properties will be justified in Sections 4.3 and 4.4.

Definition 4.9 A situation s is full iff its type set is positively complete, i.e., \((s)^+ \subseteq \Theta\).

Definition 4.10 A perspective is static iff all its situations are full.

A full situation is a place from which no information can flow. All information it carries is already there. Consequently, a static perspective is one in which every context is informationally independent from every other context.

Proposition 4.11 A perspective is static iff it satisfies strong facticity.

This trivial proposition links the strong facticity of a perspective with information saturation and the lack of information flow. We will see these two aspects of the static perspective brought out in the examples concerned with logical calculi and information theory respectively.

Definition 4.12 A situation is maximal iff every grounded type is in its type set.
A type is top iff it is in the positive completion of every grounded type.
A type is bottom iff its positive completion contains every grounded type.

Proposition 4.13 A full situation grounding a bottom type is maximal.

Proposition 4.14 Every ungrounded type in a humean perspective is bottom.

The reader might try to prove these last two propositions as an exercise in understanding our terminology!
4.1.3 Perspectives from the Outside

So far we have only studied the internal structure of perspectives: the relationships between situations, types and constraints within a perspective. In order to make the notion of a perspective a useful one, we have to be able to compare them. The whole point of introducing perspectives is that there is more than one.

The story will not be completed until Section 5.5 where we provide the conditions under which a situation supports a single unit of information, or 'infon'. In this section we develop the basic notions that will allow us to continue. Firstly, we give the criteria for one perspective to be part of another (a subperspective). After that, we examine possible conditions for equivalence between perspectives. We end up with the notion of a shift as our preferred way of comparing perspectives.

Subperspectives

Definition 4.15 The structure \((S,T,\vdash,\bot)\) is a substructure of the perspective \(P\) iff \(S \subseteq S^P\), \(T \subseteq T^P\), and \(\vdash\) and \(\bot\) are the restrictions of \(\vdash^P\) and \(\bot^P\) to \(S\) and \(T\). A substructure \(P'\) of \(P\) is a subperspective of \(P\), written \(P' \leq P\), iff \(P'\) is also a perspective.

The basic intuition of how to form a subperspective from a perspective \(P\) is to form a substructure \(P'\) of \(P\), i.e. restrict the situation domain \(S^P\) and type set \(T^P\) of \(P\) to \(S^{P'}\) and \(T^{P'}\) respectively, but leave things otherwise unchanged. By "otherwise unchanged" we mean that the internal structure of \(P'\) (its classification relation and constraints) is just the restriction of the internal structure of \(P\) to a smaller situation domain and type set. However, we need to strengthen this condition slightly to be sure that the resulting substructure is indeed a perspective. This is because an arbitrary restriction of the situation domain of a perspective may violate facticity. For example, if \(s\) is a situation in the restricted domain of type \(t\) and \(t \Rightarrow t'\) then facticity (for \(P\)) would demand that there is a situation in the restricted domain of type \(t'\). This cannot be guaranteed, since facticity (for \(P\)) only ensures that there is some situation \(s'\) in \(S^P\) of type \(t'\), but we cannot be sure that \(s'\) is in the restricted domain \(S^{P'}\).

Proposition 4.16 Given a substructure \(P'\) of a perspective \(P\), \(P' \leq P\) iff \(S^{P'}\) is positively complete in \(P'\).
CHAPTER 4. PERSPECTIVES

In other words, we can form a subperspective of $P$ from any subset $S$ of its situation domain from which no information flows to other situations of $P$ outside $S$.

Perspective Shifts

There is a difference in the theoretical status of situations and types in our theory. Situations are independent parts of the world which may be shared by many different perspectives. In contrast, situation types are only significant within a perspective. They play no role beyond their contribution to the internal structure of a perspective and the information flow that this structure licenses. Nevertheless, in comparing perspectives we need to be able to relate types from different perspectives together. In particular, we want to know when two types from different perspectives are playing the same role.

Our method of comparing perspectives is to consider mappings between their type sets which map types to types with similar roles. There are two aspects of a perspective influencing whether or not types have a 'similar role':

- the types may classify the same situations, and
- the types may be dependent on other types in a similar way.

The factors determining whether a mapping between type sets is a good comparison are therefore whether or not it maps types to types classifying the same situations and whether or not it preserves (or anti-preserves) informational relationships between types. Mappings satisfying both these conditions give us the standard notions of homomorphism and isomorphism.

Definition 4.17 A mapping $\rho : TP \rightarrow TP'$ is a weak homomorphism from perspective $P$ to perspective $P'$ (called the source and target of $\rho$) iff $SP' \subseteq SP''$ and for every $t, t' \in TP$ and $s \in SP'$,
- if $s : t$ then $s : \rho t$,
- if $t \Rightarrow t'$ then $\rho t \Rightarrow \rho t'$, and
- if $t \perp t'$ then $\rho t \perp \rho t'$.

If the converses of these conditions also hold then $\rho$ is a strong homomorphism. If $\rho$ has an inverse which is a homomorphism from $P'$ to $P$ then $\rho$ is an isomorphism and $P$ and $P'$ are isomorphic perspectives.
We regard the test of a good comparison between perspectives to be whether or not we end up with a new way of looking at the old situations. If we look at the situations of a perspective through the image of that perspective under a mapping, is what we see coherent? To make this test precise we need to define what we been by the image of a perspective under a mapping.

**Definition 4.18** If $\rho : TP \rightarrow T'P'$ then the image of $P$ under $\rho$ is the structure $P\rho = \langle S^\rho, \text{ran}(\rho), ;, \Rightarrow, \bot \rangle$ where $\Rightarrow$ and $\bot$ are the restrictions of the corresponding relations in $P'$ to $\text{ran}(\rho)$, and for $s \in S^\rho$ and $t' \in \text{ran}(\rho)$

$$s : t' \text{ in } P\rho \text{ iff there is some } t \in T^P \text{ such that } s : t \text{ in } P \text{ and } \rho t = t'.$$

Since this definition is somewhat complex on paper, but really fairly straightforward an idea, we have illustrated the image of $P$ under $\rho$ in Figure 4.2.

By asking that the image under comparison map be 'coherent', we are requiring that it is a perspective, i.e. that it satisfies the axioms of Definition 4.1. The notion of weak homomorphism is not strong enough to guarantee this since there may be dependencies in the image which are not respected by the original situation domain. However, strong homomorphisms do ensure that the images they create are perspectives. In fact, if $\rho : TP \rightarrow T'P'$ is a strong homomorphism then the image $P\rho$ is a subperspective of $P'$. 

---

*Figure 4.2: The image $P\rho$ of $P$ in $\rho : TP \rightarrow T'P'$*
The image under an isomorphism \( \rho : T^P \rightarrow T^{P'} \) is the perspective \( P' \) itself, or rather it is almost \( P' \) since \( P' \) may have some 'hidden' unclassified situations which are not in the domain of \( P \). If we are happy with the above definition of isomorphism, then this result tells us that unclassified situations are essentially uninteresting as far as the theory is concerned. If we simply add a number of unclassified situations to a perspective, we just get an isomorphic copy of the same perspective. We might as well have done without unclassified situations from the start. In fact, in Section 5.1 we ban them for technical reasons.

Despite the fact that strong homomorphisms are well behaved in that they create images that are perspectives, they are too inflexible for our purposes. As we noted above the image of a perspective under a strong homomorphism is a subperspective of the target perspective of the map. We have not really found a new way of looking at the old situations at all. In particular, when finding a perspective to compare our original perspective to, we are confined to perspectives over the old situation domain. We want a more flexible notion of comparison which allows us to compare perspectives over remotely separated situations in virtue of their informational structure. The minimal conditions are given by the notion of a perspective shift, defined below.

**Definition 4.19** A map \( \rho : T^P \rightarrow T^{P'} \) is a shift between perspectives \( P \) and \( P' \), written \( \rho : P \leftrightarrow P' \), iif for all \( t, t' \in T^P \),

- if \( \rho t = \rho t' \) then \( t \succ t' \) (shifted facticity)
- if \( \rho t \perp \rho t' \) then \( t \parallel t' \) (shifted local preclusion)

The image \( P\rho \) of \( P \) under a shift \( \rho \) is called the shifted perspective. This title is justified by the following proposition:

**Proposition 4.20** If \( \rho : P \leftrightarrow P' \) then \( P\rho \) is a perspective.

**Strong and Weak Shifts**

The problem with shifts is that, in general, they do not compose to give shifts.

\[
P \trianglelefteq P' \trianglelefteq P''
\]
CHAPTER 4. PERSPECTIVES

If \( \rho_2 \rho_1 x \Rightarrow \rho_2 \rho_1 y \) in \( P'' \) then \( \rho_1 x > \rho_1 y \) in \( P' \), but this is not enough to ensure that \( x > y \) in \( P \). This stems from the mixture of constraints (\( \Rightarrow \) and \( \bot \)) and extensional relations (\( > \) and \( \| \)) in the definition.

There are two minimal ways of strengthening the definition of a shift to get composition: by insisting that either the constraints or the extensional relations are anti-preserved. These two approaches will give us the definitions of strong and weak shifts respectively.

Definition 4.21 A map \( \rho : T^P \rightarrow T^{P'} \) is a strong shift between perspectives \( P \) and \( P' \) iff for all \( t, t' \in T^P \),
- if \( \rho t \Rightarrow \rho t' \) then \( t \Rightarrow t' \)
- if \( \rho t \bot \rho t' \) then \( t \bot t' \)

Definition 4.22 A map \( \rho : T^P \rightarrow T^{P'} \) is a weak shift between perspectives \( P \) and \( P' \) iff for all \( t, t' \in T^P \),
- if \( \rho t > \rho t' \) then \( t > t' \)
- if \( \rho t \| \rho t' \) then \( t \| t' \)

Proposition 4.23 Strong and weak shifts have the following properties:

1. both strong and weak shifts are shifts.
2. the identity map is a strong and weak shift.
3. the inclusion shift from a subperspective to a perspective is strong and weak.
4. a strong homomorphism is a strong shift.

Let \( \rho_1 : P \hookrightarrow P' \) and \( \rho_2 : P' \hookrightarrow P'' \) be shifts.
5. if \( \rho_1 \) and \( \rho_2 \) are both strong then so is \( \rho_2 \rho_1 \).
6. if \( \rho_1 \) and \( \rho_2 \) are both weak then so is \( \rho_2 \rho_1 \).
7. if \( \rho_2 \) is strong then \( \rho_2 \rho_1 \) is a shift.
8. if \( \rho_2 \) is weak then \( \rho_2 \rho_1 \) is a shift.
The last proposition shows that both strong and weak shifts impose stronger conditions on the mapping than shifts do. The terms "strong" and "weak" refer to the extent of similarity between the constraints of the source and target perspectives which is revealed by the shift. A strong shift indicates a large degree of similarity since the constraints ⇒ and ⊥ are anti-preserved, whereas a weak shift can only reveal a small degree of similarity since the induced comparison is possible because of extensional similarities between the perspectives alone.

A strong shift from a perspective can be seen as a generalisation of the informational structure in the perspective. The shift may 'forget' about some dependencies between types in the move, but any dependency in the shifted perspective will have its correlate in the original perspective. Looking at it the other way round, if there is a constraint between t₁ and t₂ in P then any strong shift into P which has t₁ and t₂ in its range will 'copy' the constraint to types corresponding to t₁ and t₂ in the source perspective. Many such 'copies' can be made in different perspectives. Following this metaphor, we think of the relation of general to specific as being that of a constraint to a copy.

Finally, we mention that shifts, strong shifts and weak shifts all determine notions of equivalence between perspectives in the same way that homomorphisms determine the notion of isomorphism. For example, a shift ρ is a shift equivalence if it has an inverse which is also a shift. Strong shift equivalence is very close to isomorphism except regarding the situation domains of the source and target perspectives. If ρ is a strong shift equivalence which preserves situation sets (i.e. t = ρt) then it is an isomorphism.

### 4.2 Conditional Constraints

We ended Chapter 3 with the promise that we would suggest a solution to the problems concerning conditional constraints discussed in Section 3.2. What counts as a solution is not very well determined owing to the weak logical structure of constraints in ST. Despite this, some additional principles have been revealed through the applications of ST. Barwise (1989a) uses conditional constraints to give an analysis of conditional statements in which he lists five axioms describing those logical properties of constraints on which his account is based. They are:
1. if \( t_1 \Rightarrow t_2 | b \) and both \( t_1 \) and \( b \) are grounded then so is \( t_2 \).

2. if \( t_1 \Rightarrow t_2 | b \) and \( t_2 \Rightarrow t_3 | b \) then \( t_1 \Rightarrow t_3 | b \).

3. if \( t_1 \Rightarrow t_2 | b \) is parametric and \( f \) is an appropriate anchor then \( t_1(f) \Rightarrow t_2(f)b(f) \).

4. if \( b' \) is stricter than \( b \) and \( t_1 \Rightarrow t_2 | b \) then \( t_1 \Rightarrow t_2 | b' \).

5. if \( t_1 \Rightarrow t_2 | b \) then \( t_1 \) is compatible with \( b \).

Our suggestion is to model the background condition of a conditional constraint by a perspective, \( P_b \). Constraints within a perspective are automatically conditional since they apply specifically to the way of looking at the world from that perspective. There can be no reduction to an unconditional constraint because there can be no reference made to a perspective from within itself. A perspective is analogous to a probability measure in Dretske’s account of information flow. No event which is measurable within the system can capture the implicit relativity to the communication channel.

We can see that this interpretation for the role of the background condition satisfies each of Barwise’s axioms.

1. is given by the facticity condition on perspectives.

2. is given by the xerox principle.

3. is justified by appeal to the notion of a strong shift. Given a perspective \( P \) in which \( t_1 \Rightarrow t_2 \) and a strong shift \( \rho : P_b \leftrightarrow P \), we have for all \( t, t' \in T^{P_b} \) such that \( \rho(t) = t_1 \) and \( \rho(t') = t_2 \), \( t \Rightarrow t' \). A strong shift can be seen as a way of generalising laws inside perspectives without an identification of types between perspectives.

4. is given by interpreting ‘stricter’ as subperspective. The axiom follows from the previous argument since, by Proposition 3, the inclusion shift from a subperspective to a perspective is strong.

5. is given by \( t_1 \in P^{T_P} \).

One problem with this approach is that the ‘range’ of a conditional constraint is restricted to the same collection of situations as the domain: if \( s \in T^{P_b} \) and \( t_1 \Rightarrow t_2 | P_b \)
CHAPTER 4. PERSPECTIVES

then there is an $s' \in T^B$ such that $s' : t_2$. This restricts inference to reasoning about situations where the background conditions are the same as they are in the situation of the antecedent type.

This problem is resolved by considering more than one perspective. Suppose that $s' \in T^B$ and $\rho(t_1) \Rightarrow t_3 | P_c$, where $\rho : P_b \leftrightarrow P_c$ is a shift. Now we do not have the general rule (HS):

$$\text{if } s : t_1 \text{ and } t_1 \Rightarrow t_2 | P_b \text{ and } t_2 \Rightarrow t_3 | P_c \text{ then } \exists s'' : t_3$$

which is just as well since the point of axiom 2 is to permit Hypothetical Syllogism under constant background conditions, but disallow such inferences under changing background conditions (see Barwise 1989a). However, for an agent moving through situations $s, s', s''$ the shift in perspective $\rho$ is sufficient for it to make the right prediction about $s''$. This is an illustration of how the transitivity of inference can work through a shift given that a particular path of inquiry is followed. Of course, we are not attempting to formally justify the inference (which is not justified), but merely remarking on how a situated agent can exploit changing conditions by shifting perspective.

Barwise (1989c) suggests a new approach to the "somewhat mysterious" background conditions of a conditional constraint. He suggests a strengthened form of axiom 1 for both positive and negative constraints which amounts to the following:

1a. If $b = \langle \Rightarrow, t_1, t_2 \rangle$ and $s \not\subset b$ and $s : t_1$ then there is a situation $s' \not\subset w$ such that $s' : t_2$.

1b. If $b = \langle \Leftarrow, t_1, t_2 \rangle$ and $s \not\not\subset b$ and $s : t_1$ then there is no situation $s' \not\subset w$ such that $s' : t_2$.

Here the background condition, $b$ is interpreted as a situation rather than a type, and its extension is the domain of situations which are 'part of' (\not\subset) $b$. The conditional constraints $t_1 \Rightarrow t_2 | b$ are just the insons $\langle \Rightarrow, t_1, t_2 \rangle$ supported by $b$. The situation $w$ is the (actual) world.\footnote{Barwise's definitions are in terms of insons, not types, but our translation of them is hopefully a faithful one.}\footnote{i.e., a distinguished maximal part of reality and possibly a situation—see branch points 1 and 3 in Barwise (1989c).}
An interesting observation from our point of view is that when the background situation \( b \) is the world \( w \) (or, equivalently, when the background type is universal), axioms 1a, 1b and 2 define a perspective! The remainder of the section pursues this point.

**Proposition 4.24** Given axioms 1a, 1b and 2, if \( b = w \), and \( S = \{ s \mid s \leq b \} \) and \( T \) is the collection of situation types and \( \Rightarrow, \perp \) are relations on \( T \) such that for \( t, t' \in T \)

\[
t \Rightarrow t' \text{ iff } b \models \langle \Rightarrow, t, t' \rangle
\]

\[
t \perp t' \text{ iff } b \models \langle \perp, t, t' \rangle \tag{6}
\]

and \( \perp \) is the usual situation theoretic 'of type' relation, then \( P_3 = (S, T, \Rightarrow, \perp) \) is a perspective with global preclusion.\(^7\)

**Proof:**

1. **Facticity** is given by 1a and \( b = w \)

2. **Xerox Principle** is given by axiom 2.

3. **Local Preclusion** is given by **Global Preclusion** which follows from 1b.

4. **Mutual Preclusion** follows from 1b.

If \( b \neq w \) then it is prima facie possible that \( P_3 \) is not a perspective since it is consistent with 1a that the 'range' of a conditional constraint \( t_1 \Rightarrow t_2 \mid b \) may fall outside \( b \). In other words there may be a situation \( s \leq b \) of type \( t_1 \), but no type \( t_2 \) situation part of \( b \). By 1a, there must be some situation of type \( t_2 \), but since \( b \neq w \), it may not be part of \( b \). Suppose that this is indeed the case.

Also let us suppose that there are no type \( t_3 \) situations at all. It would seem that the constraint \( t_2 \Rightarrow t_3 \mid b \) should be consistent with this, since it is local to \( b \) and so entirely

\(^6\)This somewhat contorted definition is only necessary because Barwise does not insist that \( \perp \) is symmetric.

\(^7\)We are not suggesting this construction as a definition of 'background perspective', since in general the domain of a perspective may not be characterisable as the parts of a unique situation (i.e. \( S^\phi \) may not have a least upper bound in the \( \leq \) ordering). The construction is intended to show how Barwise's suggestion, axiomatised by 1a. and 1b., can be related to our analysis of background conditions using perspectives.
vacuous as no situations part of \( b \) are of type \( t_2 \). However, by 2, we have that \( t_1 \Rightarrow t_2 \mid b \) and so by 1a, there must be a situation of type \( t_3 \). To resolve this conflict we must decide between:

the consistency of the constraint \( t_2 \Rightarrow t_3 \mid b \) when there are no situations of type \( t_2 \) which are part of \( b \), and

the possibility of there being no situations of types \( t_2 \) which are part of \( b \) when the constraint \( t_1 \Rightarrow t_2 \mid b \) holds.

If we choose to reject the latter then \( P_b \) is a perspective for every \( b \).

4.3 Informational Perspectives

In this section we apply the idea of a perspective to bridging the gap between Dretske's (1981) theory of information and situation theory. We build models of Dretske's illustrative example of a two unit communication system in 4.3.1 and generalise it in 4.3.2. In each case we show that the model is a perspective. We also identify various properties of these models which we think of as characteristic of 'informational' perspectives. They will be used to compare the perspectives of this section with the 'logical' perspectives of Section 4.4.

Our model of communication is initially restricted to a single source, \( S \), and receiver, \( R \). The behaviour of \( S \) and \( R \) is taken to be entirely specified by the states they can adopt, together with the probabilities of their being in each state. \( S_{st} \) and \( R_{st} \) are distinct, exhaustive and mutually exclusive sets of states for \( S \) and \( R \) respectively.

At any given moment, the state of the system as a whole is represented by a pair, \( \langle s_1, s_2 \rangle \), where \( s_1 \in S_{st} \) is the state of the source and \( s_2 \in R_{st} \) is the state of the receiver. The behaviour of the system is therefore a trajectory in \( S_{st} \times R_{st} \).

Although we have included a temporal dimension in our model, we will ignore the possibility of information flow over time. We are only interested in the information conveyed by the receiver about the present state of the source, not about future or past states of the system. For all it matters, the behaviour of the system may be random over time: the states that \( S \) and \( R \) adopt now could be entirely independent of the states they adopted in the past.
CHAPTER 4. PERSPECTIVES

The state space of the system has a probability measure \( \mu \) defined over it. The space of 'possible events', \( E \), consists of those subsets of \( S_t \times R_t \) with non-zero probability. For events \( e_1 \) and \( e_2 \in E \), the conditional probability of an \( e_1 \) given \( e_2 \) is defined as:

\[
\mu(e_1|e_2) = \frac{\mu(e_1 \cap e_2)}{\mu(e_2)}
\]

We are particularly interested in the events

'S is in state \( \sigma_1 \)', given by \( t_{\sigma_1} = \{ (\sigma_1, x) | x \in R_t \} \), and

'R is in state \( \sigma_2 \)', given by \( t_{\sigma_2} = \{ (x, \sigma_2) | x \in S_t \} \).

for each \( \sigma_1 \in S_t \) and \( \sigma_2 \in R_t \), since these classify the states of \( S \) and \( R \) independently.

We insist that each \( t_{\sigma} \) has a non-zero probability and so is in \( E \). Note that the events \( t_{\sigma_1} \) and \( t_{\sigma_2} \) are disjoint (i.e., \( t_{\sigma_1} \cap t_{\sigma_2} = \emptyset \)) when \( \sigma_1 \neq \sigma_2 \) and \( \sigma_1, \sigma_2 \in S_t \) or \( \sigma_1, \sigma_2 \in R_t \).

Also \( \bigcup \{ \{ \sigma \in S_t \times R_t \} = S_t \times R_t \), so one and only one event in each of the sets \( \{ t_\sigma | \sigma \in S_t \} \) and \( \{ t_\sigma | \sigma \in R_t \} \) occurs.

4.3.1 The Source/Receiver Perspective

The perspective we construct is a perspective on the system at a particular (but arbitrary) time, \( \tau \). The system at \( \tau \) is divided into two epistemic contexts: the situation at the source and the situation at the receiver. These are modelled by two situations, \( s \) and \( r \). They are classified independently, according to the respective states of the source and receiver at time \( \tau \). To capture this we model the classification by the relation 't' between the situations, \( s \), \( r \), and the events \( t_\sigma \), for \( \sigma \in S_t \cup R_t \), defined by

\( s : t_\sigma \iff S \text{ is in state } \sigma \in S_t \)

\( r : t_\sigma \iff R \text{ is in state } \sigma \in R_t \)

The set of types in the classification is just the set \( T = \{ t_\sigma | \sigma \in S_t \cup R_t \} \). Since the events \( t_\sigma \), \( t_\sigma' \) are mutually exclusive when \( \sigma \) and \( \sigma' \) are both states of the same situation, the type sets \( s \) and \( r \) are just the singletons \( \{ t_\sigma \} \) and \( \{ t_\sigma \} \), where \( (\sigma_s, \sigma_R) \) is the state of the system at time \( \tau \). Because we are modelling an actual state of the system, we insist that

\[ \mu(\{ (\sigma_s, \sigma_R) \}) > 0 \] (Actuality)
CHAPTER 4. PERSPECTIVES

The 'actuality' property is an important constraint on the model and the proof of facticity crucially depends on it.

In this simple example, the only constraints we are interested in are positive and negative dependencies between types. These are defined by

\[ t_{t_1} \Rightarrow t_{t_2} \text{ iff } \mu(t_{t_1} | t_{t_2}) = 1 \text{ and } \mu(t_{t_2}) < 1 \]

\[ t_{t_1} \perp t_{t_2} \text{ iff } \mu(t_{t_1} | t_{t_2}) = 0 \text{ and } \mu(t_{t_2}) > 0 \]

The \( \Rightarrow \) constraint is a direct translation of Dretske's definition, whereas the \( \perp \) constraint is our addition. Both definitions, however, can be simplified in all but the most trivial case:

Proposition 4.25 If the source and receiver are both non-constant (i.e., have more than one possible state) then for any states \( \sigma_1, \sigma_2 \)

\[ t_{\sigma_1} \Rightarrow t_{\sigma_2} \text{ iff } \mu(t_{\sigma_2} \cap t_{\sigma_1}) = \mu(t_{\sigma_1}) \]

\[ t_{\sigma_1} \perp t_{\sigma_2} \text{ iff } \mu(t_{\sigma_2} \cap t_{\sigma_1}) = 0 \]

Proof: Each \( t_\sigma \) is in \( E \), so \( \mu(t_\sigma) > 0 \). Suppose \( \mu(t_\sigma) = 1 \) for some \( \sigma \in S_m \). Then, assuming the source is non-constant, there is a \( \sigma' \in S_m \) such that \( \sigma \neq \sigma' \). But then \( t_\sigma \cap t_{\sigma'} = 0 \), so

\[ 1 \geq \mu(t_\sigma \cup t_{\sigma'}) = \mu(t_\sigma) + \mu(t_{\sigma'}) = 1 + \mu(t_{\sigma'}) \]

so \( \mu(t_{\sigma'}) = 0 \), contradicting the fact that \( t_{\sigma'} \in E \). So for all \( \sigma \in S_m \), \( 0 < \mu(t_\sigma) < 1 \).

The argument for \( \sigma \in R_m \) is identical and the rest follows directly from the definition of conditional probability.

Taken together the above structure \( \langle \{x, y\}, T, \Rightarrow, \perp \rangle \), which we will call \( D_2 \), constitutes a model of the flow of information of the form "U is in state \( \sigma \)" (where \( U = S \) or \( R \)) between a single source and a single receiver, according to Dretske's definition. We will assume that both source and receiver are non-constant.

Proposition 4.26 \( D_2 \) is a perspective.

Proof: We must verify the axioms for a perspective on \( D_2 \).
1. Facticity is reasonably straightforward since \( t_e : t_e \) and \( t_e \) are the only grounded types. So the only cases to consider are when \( t_e \) or \( t_e \) for some arbitrary state \( \sigma \). The argument is the same for both cases, so we will only consider the first. If \( t_e \) then \( \mu(t_e \cap t_e) = \mu(t_e) \), and we must show that there is a situation of type \( t_e \). This amounts to showing that either \( \sigma = t_e \) or \( \sigma = \sigma_e \).

If \( \sigma \in E_R \) then either \( \sigma = \sigma_e \), as required, or \( t_e \cap t_e = \emptyset \), so that \( \mu(t_e) = \mu(t_e \cap t_e) = 0 \), contradicting the fact that \( t_e \in E \).

If \( \sigma \in S_M \) then \( t_e \cap t_e = \{\sigma, \sigma_e\} \) so that
\[
\mu(t_e) = \mu(\{\sigma, \sigma_e\})
\]
Now either \( \sigma = \sigma_e \), as required, or \( \{\sigma, \sigma_e\} \) and \( \{\sigma_e, \sigma_e\} \) are mutually exclusive events, so that
\[
\mu(\{\sigma, \sigma_e\}) + \mu(\{\sigma_e, \sigma_e\}) = \mu(\{\sigma, \sigma_e\}, \{\sigma_e, \sigma_e\})
\]
Since \( \{\sigma, \sigma_e\}, \{\sigma_e, \sigma_e\} \subseteq t_e \), we have \( \mu(\{\sigma, \sigma_e\}, \{\sigma_e, \sigma_e\}) \leq \mu(t_e) \). Hence,
\[
\mu(\{\sigma, \sigma_e\}) + \mu(\{\sigma_e, \sigma_e\}) \leq \mu(\{\sigma, \sigma_e\})
\]
and so \( \mu(\{\{\sigma_e, \sigma_e\}\}) = 0 \), contradicting the actuality property. This exhausts the possibilities for \( \sigma \) and so completes the proof.

2. Xerox Principle is straightforward in this case, but we omit the proof (see Proposition 4.32 for the general case).

3. Local Preclusion. For each situation \( x \), if \( x : t_e \) and \( t_e \perp t_e \) then \( \mu(t_e \cap t_e) = 0 \).

But \( t_e \in E \), so \( \mu(t_e) \neq 0 \), so \( t_e \neq t_e \cap t_e \), and so \( \sigma \neq \sigma' \). But then \( x \not\perp x' \), since \( x \) is a singleton.

4. Mutual Preclusion. By Proposition 4.25, \( \perp \) is clearly symmetric.

Although we have made a notional distinction between 'source' and 'receiver', the model does not have this asymmetry. In generalising to a system with many sources and receivers, we will refer to them simply as 'units'. The proofs can easily be extended to the \( n \)-unit case, \( D_n \).
CHAPTER 4. PERSPECTIVES

The restriction of types to classifications of the form 'x is in state σ' is only a small simplification of the general case in which types classify situations in an arbitrary way depending on their states. In the general case each type t would be such that

\[ z : t \iff P_t(\sigma_X) \]

where (for each t) \( P_t(\sigma_X) \) is a proposition, the truth of which depends only on \( \sigma_X \), the actual state of unit X. Before we construct a general model, we note a few properties of \( D_2 \) for comparison with the 'logical' perspectives later.

**Proposition 4.27** \( D_2 \) is static iff there is no information flow from source to receiver or vice versa.

**Proof:** If r is full and \( r : t_0 \) and \( t_0 \Rightarrow t_2 \) then \( r : t_2 \), so \( s \not\in t_2 \), so there is no flow of information from S to R. Similarly, there is no flow from R to S if s is full.

**Proposition 4.28** \( D_2 \) is non-humean.

**Proof:** Trivial. For example, for all \( \sigma \in S_M, \{ s \} = t_{x_2} \), so by Proposition 4.8 \( t_0 \not\Rightarrow t_{x_2} \). But if \( \sigma \not\in s_2, t_0 \cap t_{x_2} = \emptyset \), so \( \mu(t_0 \cap t_{x_2}) = 0 \neq \mu(t_{x_2}) \) since \( t_{x_2} \in E \). Therefore \( t_0 \not\Rightarrow t_{x_2} \) contradicting Hume's property.

**Proposition 4.29** There are bottom and top types in \( D_2 \) iff there is information flow from source to receiver or vice versa.

**Proof:** Information flows from source to receiver iff the signal, \( t_{x_2} \), carries information \( t_{x_2} \) about the source's state, \( \sigma_S \), i.e., \( t_{x_2} \Rightarrow t_{x_2} \). But since \( t_{x_2} \) and \( t_{x_2} \) are the only grounded types, \( t_{x_2} \Rightarrow t_{x_2} \) iff \( t_{x_2} \) is bottom and \( t_{x_2} \) is top.

4.3.2 Informational Perspectives in General

In this section we put Dretske's ideas about the flow of information in the general setting of an arbitrary dynamic system.

We start with a state space \( S \) which is evolving over time. We characterise its behaviour by a probability space \( P = (S, M, \mu) \), where \( \mu \) is a probability measure defined over a
collection of measurable events, $M$.

Again we distinguish the set $E$ of 'possible events' as those with non-zero probability.

Rather than explicitly defining the state space as a product of the state spaces of a set $\{U_0, \ldots, U_n\}$ of units, we will remain uncommitted to any such factorisation of the space. Instead we will associate an arbitrary set of possible events $T_i(r)$ with each $U_i$ at time $r$. A particular $T_i(r)$ constitutes the specification of unit $U_i$'s possible behaviour at time $r$, in the same way that the state space of a unit would; the extra degree of freedom arises from the flexibility of the connection between this specification and the state of the system as a whole.

Since, again, we are only interested in the description of the system at a particular (but arbitrary) time, $r$, we will drop the parameter $r$ from the $T_i(r)$'s. As before, the units at time $r$ can be perceived as dividing the system into situations, $s_0, \ldots, s_n$, which form the situation domain, $S$, of our model. The domain of types $T$ is just $\bigcup \{T_i \mid i \in n\}$ and the map $\pi : S \to T$ defined by $\pi(s_i) = T_i$ recovers the $T_i$'s from $T$. At time $r$, the system will be in a particular state $\sigma \in S$. Now the singleton $\{\sigma\}$ may not be a measurable event, but we can approximate it by measurable events, and so define

$$[\sigma] = \bigcap \{e \mid e \in e \in M\}$$

Since $\sigma$ is an actual state of the system, we insist on the actuality property which here amounts to saying that $[\sigma] \in E$. The classification relation $\vdash$ between situations and types is defined by

$$s_i : t \iff \sigma \in e \in T_i$$

The constraints between $t, t' \in T$ are defined, as before, by

$$t \Rightarrow t' \iff \mu(t') = 1 \text{ and } \mu(t') < 1$$

$$t \perp t' \iff \mu(t') = 0 \text{ and } \mu(t') > 0$$

Since $t' \in T \subseteq E$ we can be sure that $\mu(t') > 0$, but it need not be the case that $\mu(t') < 1$, so the simplification in Proposition 4.25 holds for $\perp$ but not for $\Rightarrow$.

\textit{i.e., $M$ is a collection of subsets of $S$ containing $\emptyset$ and closed under complements and countable union.}
The structure \((S,T,\Rightarrow,\lambda,\mathfrak{I})\), as defined above, will be referred to as \(P_{\mathfrak{I}}\). This is a generalisation of \(D_2\) since different situations can ground the same types and different types can classify the same situation. The classification in \(D_2\) is just a special case of the above, since in \(D_2\),

\[
s : t_r \iff (s,\sigma) \in t_r \in S_{\mathfrak{I}}
\]

\[
r : t_r \iff (s,\sigma) \in t_r \in R_{\mathfrak{I}}
\]

Before showing that \(P_{\mathfrak{I}}\) is a perspective, we will prove a simple Lemma about conditional probabilities:

**Proposition 4.30** For events \(a \in E, b \in M\),

\[
\mu(b|a) = 1 \iff \mu(b^*|a) = 0,
\]

where \(b^*\) is the complement of \(b\).

**Proof:** \(a \cap b\) and \(a \cap b^*\) are disjoint, and \(a = (a \cap b) \cup (a \cap b^*)\), so

\[
\mu(a) = \mu(a \cap b) + \mu(a \cap b^*)
\]

and \(\mu(a) > 0\), since \(a \in E\), so we can divide through to get \(\mu(b|a) + \mu(b^*|a) = 1\) from which the result follows.

**Proposition 4.31** For types \(t_1, t_2, t_3 \in T\),

\[
t_1 \Rightarrow t_2 \iff t_1 \perp t
\]

**Proof:** Noting that \(t_2 \in T\) implies \(\mu(t_2) > 0\) and \(\mu(t_2) < 1\), \(t_1 \Rightarrow t_2\) is an immediate corollary of Proposition 4.30.

**Proposition 4.32** \(P_{\mathfrak{I}}\) is a perspective.

**Proof:**

1. **Facticity** Suppose \(s : t_1\) and \(t_1 \Rightarrow t_2\), for \(s \in S\), \(t_1, t_2 \in T\), so \(\sigma \in t_1 \in \pi s\) and \(\mu(t_2|t_1) = 1\). By Proposition 4.30, \(\mu(t_2|t_1) = 0\), and so \(\mu(t_1 \cap t_2) = 0\). Now, either \(\sigma \in t_1 \cap t_2\) or \(\sigma \in t_1 \cap t_2\). But if \(\sigma \in t_1 \cap t_2\) then \([\sigma] \subseteq t_1 \cap t_2\), and so

\[
\mu([\sigma]) \leq \mu(t_1 \cap t_2) = 0
\]

contradicting the actuality property. So \(\sigma \in t_1 \cap t_2\) and so

\[
\sigma \in t_2 \in T = \text{ran}(\pi)
\]

so there is some \(s' \in S\) such that \(\sigma \in t_2 \in \pi s',\) i.e., \(s' : t_2\)
2. Xerox Principle Suppose \( a \Rightarrow b \) and \( b \Rightarrow c \), where \( a, b, c \in T \), then \( \mu(b|a) = \mu(c|b) = 1 \), so by Proposition 4.30,
\[
\mu(a \cap b^*) = \mu(b \cap c^*) = 0
\]
Now \( c^* \cap a \cap b \) and \( c^* \cap a \cap b^* \) are disjoint, and \((a \cap c^*) = (c^* \cap a \cap b) \cup (c^* \cap a \cap b^*)\), so
\[
\mu(a \cap c^*) = \mu(c^* \cap a \cap b) + \mu(c^* \cap a \cap b^*)
\]
but \( c^* \cap a \cap b \subseteq b \cap c^* \) and \( c^* \cap a \cap b^* \subseteq a \cap b^* \) so \( \mu(c^* \cap a \cap b) \leq \mu(b \cap c^*) = 0 \)
and \( \mu(c^* \cap a \cap b^*) \leq \mu(a \cap b^*) = 0 \), so \( \mu(a \cap c^*) = 0 \), and so by Proposition 4.30,
\( \mu(c|a) = 1 \). Also \( \mu(c) < 1 \) since \( b \Rightarrow c \). Hence \( a \Rightarrow c \).

3. Local Preclusion If \( s : t_1 \perp t_2 \) then \( \sigma \in t_1 \in \pi s \) and \( \mu(t_1 | t_1) = 1 \), so by
Proposition 4.30, \( \mu(t_1 \cap t_2) = 0 \). But if \( \sigma \in t_2 \) then \( [\sigma] \subseteq t_1 \cap t_2 \), so \( \mu([\sigma]) \leq \mu(t_1 \cap t_2) = 0 \), which contradicts actuality. Then for all \( s' \in S \), \( \sigma \not\in t_4 \in \tau s' \). This
is the stronger property of Global Preclusion.

4. Mutual Preclusion The symmetry of \( \perp \) is clear from the definition.

4.4 Logical Perspectives

In this section we look at the elementary model theory of various logical systems as
defining perspectives when we regard models as situations. These 'logical' perspectives
are contrasted with the 'informational' perspectives of Section 4.3. In particular, we
relate the notion of consequence in the logical perspectives to the notion of information
flow in informational perspectives.

4.4.1 The Classical Logic Perspective

The perspective from the armchair is indeed a perspective. A formal language supplied
with a truth definition is a system for classifying models: this is the subject matter
of Model Theory. Together with a definition of logical consequence, a formal language
induces a flow of information. If a model \( M \) is classified as making the sentence \( \phi \) true,
the sentence \( \phi \lor \psi \) is also made true, as is every other logical consequence.
Definition 4.33 Let $L$ be a set of sentences of a language with a class $M$ of interpretations, or models. Define a classification $(M, L, :)$ by $m : \phi \iff \phi$ is true in $m$. Supposing some logical calculus $C$ on $L$, write $\Delta \vdash \phi$ just in case the $L$-sentence $\phi$ is provable from the set $\Delta$ of $L$-sentences. Define $\phi \Rightarrow \psi \iff (\phi) \vdash \psi$, and $\phi \perp \psi \iff (\phi, \psi) \vdash \phi \land \neg \psi$, for some propositional variable, $p$. Then let $LCM$ be the structure $(M, L, :, \perp, \Rightarrow, \perp, \perp, \land)$.

The definition of a 'logical' perspective is of necessity vague without an encompassing definition of a 'logic'. Nevertheless, the above covers many well known examples. The following three propositions characterise classical propositional logic as a perspective. We include discursive 'proofs' of these somewhat trivial propositions in order to facilitate the reader's understanding of the terminology introduced in earlier sections.

Proposition 4.34 Classical Propositional Logic is a perspective. In other words, if $P$ is a propositional language, $V$ is the set of valuations on the atomic sentences of $P$, and $K$ is a calculus for classical propositional logic, then $PKV$ is a perspective.

Proof: Firstly, if $m : \phi$ and $\phi \vdash K \psi$ then $m : \phi$, by the Soundness Theorem of classical propositional logic. But this property is strong facticity and so $PKV$ satisfies facticity a fortiori. Secondly, if $\phi \vdash K \psi$ and $\psi \vdash K \theta$ then, by the Cut Rule, $\phi \vdash K \theta$, so $\Rightarrow$ is transitive, and so $PKV$ satisfies the zero principle. Thirdly, $\phi \perp \psi \iff (\phi, \psi) \vdash K \phi \land \neg \psi \iff \psi \perp \phi$, so $PKV$ satisfies mutual preclusion. Finally, if $m : \phi$ and $\phi \perp \psi$ then $(\phi, \psi) \vdash K \phi \land \neg \psi$, so $(\phi) \vdash K \neg \psi$, so $\phi \Rightarrow \neg \psi$, and, by strong facticity, $m : \neg \psi$. Hence $m \vdash \psi$ and so $PKV$ satisfies local preclusion. Global preclusion is clearly not satisfied since although $p \perp \neg \psi$ and there are valuations which make $p$ true, there are other valuations which make $p$ false and so $\neg p$ true.

Proposition 4.35 $PKV$ is static and (weakly) humean.

Proof: It is well known that $\vdash K$ can be given an exact semantic characterisation by

$\phi \models K \iff \forall m \in V \text{ if } m : \phi \text{ then } m : \psi$

This is exactly equivalent to the definition of the $\models$ (constant conjunction) relation in $PKV$. The Soundness and Completeness Theorems state that $\vdash \models$ and $\models \vdash$ respectively. So it is clear that $\phi \Rightarrow \psi \iff \models \Rightarrow \psi$, i.e., the perspective $PKV$ is humean and satisfies strong facticity, so is static.
Proposition 4.36 The theorems and contradictions of classical propositional logic correspond exactly to the top and bottom types of PKV.

Proof: The theorems of a logical calculus are the sentences which can be proved from no assumptions. By monotonicity of $\vdash_K$, this is equivalent to being provable from any assumptions. In informational terms, then, they are in the positive completion of every type, so they are top types. Conversely, every top type is in the positive completion of every grounded type. Since PKV is humane, Proposition 4.14 ensures that every grounded type is in the positive completion of every ungrounded type. So, since top types are grounded in PKV, they are theorems.

It should be noted that, in general, the notion of a top type is weaker than that of a theorem. All that is required of a top type is that it is universally available information within the domain of situations of the perspective. Top types flow to every situation. A theorem must satisfy the stronger requirement of being a consequence of every sentence, even those that are nowhere true.

The contradictions are those sentences which are provably inconsistent with themselves, i.e., those $\phi$ such that $\{\phi}\vdash_K \neg p \land \neg p$, or in informational terms, $\models \phi$. By local preclusion, $\vdash \phi$, so $\phi$ is not grounded by any situation. Again by Proposition 4.14, any sentence is in the positive completion of a contradiction (ex falso quodlibet), so they are bottom types. Conversely, if $\phi$ is bottom then any propositional variable, $p$, and its negation, $\neg p$, are both in the positive completion of $\phi$ since they are both grounded (i.e., have a model). Hence $\phi \rightarrow p$ and $\phi \rightarrow \neg p$, so, $\vdash_K \neg p \land \neg p$, and so $\phi \neq \phi$.

In PKV $\phi \models \psi$ iff $\phi \land \psi \models p \land \neg p$. As long as we have the conjunction and negation of arbitrary types in a perspective, it is possible to use this fact to define $\models$ in terms of $\models$. In the following, we will only consider the positive aspects of perspectives, taking $\models$ to be defined in this way.

4.4.2 Information Flow and Modality

The classical logic perspective, PKV, is somewhat limited as a model of information flow. Since it is static, all its situations are full. In informational terms this means that the classification of a situation as being of certain types carries no information about the way in which other situations are classified. Information flow is entirely local to a
situation. Also, because PKV is humean, any constant conjunction of phenomena over the situation domain is reflected as a constraint between types. There is no distinction between law-like and contingent dependencies between types.

One way of modifying PKV to overcome these objections is to restrict the domain of situations. $V$ includes every valuation for $P$, so that the propositional variables are without significance: that $p$ is true in one model tells us nothing about what is true in any other model. In the sense given by probability theory, the variables are independent, which implies that the domain is essentially unstructured. For any observable structure to exist in a physical space the measurable variables must demonstrate some degree of dependence. The same is true in this abstract setting. Information flow relies crucially on the structure of the domain: if there are no dependencies to be characterised then no posited correlations between types will characterise them!

There are two ways of restricting the situation domain: externally or internally. The internal approach isolates a class of models with certain properties in common which can be distinguished by the language. For example, a set of sentences $\Delta$ defines a class of models, $V_\Delta$, namely those which make every sentence in $\Delta$ true. The perspective, $PKV_\Delta$, obtained by the restriction of $V$ to $V_\Delta$ is still static but in general it will lack Hume's property.

In general, a domain of situations $D$ is said to be definable in a perspective if there is a set of types $X$ such that for each situation $s$, $X \subseteq \exists s : s \in D$. As long as $\overline{D}$ is positively complete, the restriction to $D$ will always be a subperspective. In static perspectives every situation is full, so this condition is always met.

Logically the failure of Hume's property amounts to incompleteness. If we were to re-define the semantic consequence relation, $\models$, relative to $V_\Delta$, the Completeness Theorem would fail. The sentences of $\Delta$, for example, would be true in every model in $V_\Delta$ but not provable as theorems.

In informational terms, when we restrict the domain we are effectively adding structure which may not be reflected in the constraints on the type set. Sentences $\phi$ and $\psi$ which are logically independent (i.e., neither one is provable from the other), may be true in the same models in the restricted domain, so that $\phi \models \psi$ even though $\phi \not\models \psi$. 
CHAPTER 4. PERSPECTIVES

Although in general Hume's property is informationally unattractive, since it means that there is no selectivity in the way the constraints reflect structure in the domain, in this case no advantage is to be had from its failure. The structure we would like to be reflected is that peculiar to the restricted domain $V_\Delta$, not the purely logical regularities inherited from $PKV$. Nevertheless, there is a simple way to include constraints over the domain of types which do reflect this extra structure. The sentences $\Delta$ which define the restricted domain can be added as axioms to the calculus $K$ to give $K_\Delta$, so that the resulting perspective, $PK_\Delta V_\Delta$, is again non-humean.

By restricting the domain in a way that is internally definable within the perspective, we have not really gained. The extra structure can be absorbed by using the sentences defining the restriction as axioms in the background. At each stage the situations remain full and so information fails to flow between them. In effect, this kind of restriction introduces further complexity to the structure of the classification of individual situations without allowing dependencies between situations.

The external approach to restricting the situation domain is to select a class of models, $M$, which may not be definable by any set of sentences of $P$. Of course, there may well be some condition which precisely determines which models are included in $M$, but it need not be expressible in $P$: the class $M$ may be externally if not internally definable. The restriction of $PKV$ to $M$ ($PKM$) is again static and non-humean. But $PKM$ may be non-humean in an interesting way since there is no obvious way of modifying $K$ to restore Hume's property. What kind of calculus can reflect the structure introduced by this restriction?

The answer is to be found in the semantics for modal logics. If we augment the propositional language $P$ with a modal operator, $\Box$, to get $P^+$, we can extend the truth definition over $M$ in accordance with the rules for the modal logic S5:

\[ \Diamond \phi \text{ is true in } m \iff \exists m' \in M, \phi \text{ is true in } m' \]

Since the truth definition for modal sentences is independent of $m$, we can usefully define a notion of global truth on $M$. A sentence $\phi$ of $P^+$ is globally true on $M$ iff $\forall m \in M, \phi$ is true in $m$. In effect, we are turning the domain $M$ into a Kripke model for S5.

It is easy to see that $\phi \models \psi$ if the sentence $\phi \rightarrow \Diamond \psi$ is globally true on $M$. Given a
CHAPTER 4. PERSPECTIVES

calculus K5+ for S5, we can define for \( \phi, \psi \in P \),

\[
\phi \vdash_{K5} \psi \iff \phi \vdash_{K5^+} \Box \psi
\]

We can be sure of facticity by the soundness of K5+ with respect to the modal truth definition, so the resulting structure PK5M is a perspective.

Unfortunately, the perspective PK5M is identical to PKM since (small result) \( \psi \) is provable from \( \phi \) in K iFF \( \Box \psi \) is provable from \( \phi \) in K5+! The situation is easily rectified by putting some modal facts about \( M \) in the 'background' as axioms. To fully restore Hume's property, we could add every globally true sentence of \( P^+ \) as axioms. A more selective alternative is to view \( M \) as one of a class of Kripke models for S5 which are axiomatised by some set, \( \Delta \), of \( P^+ \) sentences. By defining,

\[
\phi \vdash_{K5^+} \psi \iff \Delta \cup \{ \phi \} \vdash_{K5^+} \Box \psi
\]

we obtain a new perspective, PK5M.

The role of \( \Delta \) is to provide a distinction between law-like relations, \( \phi \Rightarrow \psi \), between types and contingent regularities, \( \phi \Rightarrow \psi \). In general, not all contingent regularities are law-like since it may be that \( \phi \Rightarrow \Box \psi \) is globally true on \( M \) even though \( \Box \psi \) is not provable from \( \Delta \cup \{ \phi \} \). In other words, there may be a model of \( \Delta \) and \( \phi \) which is not a model of \( \psi \), i.e. \( M \) need not be 'the only' model of \( \Delta \) and \( \phi \). Consequently PK5M is non-humean. \( \Delta \) could be understood as a 'theory' of the domain which underlies the law-like dependencies of the perspective. Nevertheless, \( \Delta \) is not itself expressible in the internal language of the perspective, and so need not be part of the classificatory apparatus of the organism whose perspective is being modeled.

As long as \( \Delta \) is sufficiently rich there will be law-like dependencies in PK5M which do not occur in PKM, making PK5M fail to satisfy strong facticity. In this case, by Proposition 4.11, PK5M is not static. So some situations in PK5M will not be full and there will be a genuine flow of information between situations. For example, suppose \( \Box \psi \) is provable from \( \phi \) in K5M, so that \( \phi \Rightarrow \psi \) in PK5M. If \( m : \phi \) then there is no guarantee that \( m : \psi \), only that there exists an \( m' \in M \) such that \( m' : \phi \). The information \( \psi \) flows from \( m' \) to \( m \).
4.4.3 Modal Orderings and the Direction of Information Flow

Although we have shown how to construct perspectives with non-trivial information flow, there is a sense in which they are too unconstrained. Given that \( \phi \Rightarrow \psi \) and \( m : \phi \), we know that \( \phi \) is true somewhere in the situation domain, but we have no idea where. There is nothing which determines the source of the information, or even the general direction from which it is flowing. Again there are internal and external approaches to this problem.

The internal approach is to include some mechanism for referring to situations within the perspective itself. If, for example, the types are indexed with their situation sets, then the new perspective with types \((t, f)\) would satisfy our requirement. Nevertheless, this answer is not in keeping with our attempt to model situated inference since it assumes the classificatory capability of a theorist observing the perspective rather than an organism adopting one. It would be like speaking in a language in which every sentences is labeled with its truth value: "true—the engine won’t start because the points are sticking", "false—Mangy Mister wins the 1990 Derby!"

A more modest version of the internal approach is to have as types every pair \((s, t)\) of situations \(s\) and types \(t\), and say \( s : (s', t) \) iff \( s = s' \) and \( s : t \). In such a perspective, a type is grounded by at most one situation and so the constraints explicitly encode the source and receiver of the information. For example, the constraint \( \phi \Rightarrow \psi \) in \( PKV \) would be transformed to the family of constraints \((m, \phi) \Rightarrow (m, \psi)\), one for each \( m \in V \).

Since all situations are full, there would be no constraints of the form \((m, \phi) \Rightarrow (m', \psi)\) when \( m \neq m' \). The problem with this approach is that the connection 't' between \((s, t)\) and \((s', t)\) is not apparent within the perspective, since these types are made true by at most one situation. The interesting possibility of a mixture of the purely contextual types \(t\) and the 'referential types' \((s, t)\) is left for future work.\(^9\)

The external approach is to find additional regularities in the distribution of types across situations. For example, suppose that the situations are small intervals of time, ordered by precedence and that their classification is at least partly time dependent. The ordering relation, \( R \), can be taken to underlie the modality \( \mathcal{O} \), so that the truth

\(^9\)We suspect that there are fruitful connections with the work of Blackburn on sorted intensional logics (see Blackburn 1990).
definition for \(P^+\) is changed to:

\[O\phi \text{ is true in } m \text{ iff } \exists m' \in M, mRm' \text{ and } \phi \text{ is true in } m'\]

Instead of adopting a modal calculus for \(S5\) (which axiomatises the universal relation), the \(\Rightarrow\) constraint within the perspective can be defined by provability in a modal calculus for the relation \(R\). Of course, the particular modal logic we choose will not describe \(R\) completely, only those properties of it that are common to the class of relations axiomatised by the logic. Nevertheless, the restriction to the relation \(R\) provides an extra level of understanding of the flow of information within the perspective: information flows backwards along \(R\). In our example of an ordering of time intervals, if \(\phi \Rightarrow \psi\) and \(m : \psi\) then the information \(\psi\) which flows into \(m\) is a prediction about the future. Of course, it may not be that we chose the 'right' relation \(R\). The regularities in the domain may occur in directions other than \(R\), so that the above construction would yield little increase in information flow.

A further development, which we will not pursue here, is to include several different relations on the domain corresponding to actions an organism can perform which change the situation. For example, thinking of the situation domain as a landscape with individual situations being small regions of the terrain, the actions 'move south', 'move north', 'move east' and 'move west' can each induce modalities in the way described above. The corresponding four \(\Rightarrow\) constraints would be aligned to the flow of information from the appropriate cardinal point.\(^{10}\)

---

\(^{10}\)Dynamic Logic is exploited in this development, in a similar way to the use of Modal logics in this section.
Chapter 5

Individuals

5.1 Individuation and Predication

In Chapter 1 we claimed that the ability of an organism to classify its environment was essential to its cognitive function. In ST this ability is modeled by a scheme of individuation. But the scheme of individuation of an organism tells us nothing about its ability to use this classification to make inferences. For this we must be able to account both for the flow of information in the environment and for the organism’s ability to be attuned to that flow. Chapter 3 concerned various ways of accounting for the flow of information in ST and in Dretske’s theory of information. The problem of conditional constraints lead to the perspective model in Chapter 4, which has an implicit relativity similar to the relativity of Dretske’s account to the communication channel. ¹

To a first approximation, a perspective can be viewed in ST terms as a scheme of individuation together with a collection of constraints. But there are two key differences. Firstly, because of the implicit relativity of a perspective, it can only have a small domain. Not many situations will be classified by a single perspective because within the perspective the dependencies must be satisfied without exception. Whereas a scheme of individuation is intended to model the whole of an organism’s interaction with the world, a perspective is only intended to model a small part. Perspectives are adopted

¹Apart from a few revisions and omissions, this chapter together with Chapter 4 make up the paper Perspectives in Situation Theory – Seligman 1989.
and discarded. Schemes stay.

The second difference is that a scheme of individuation presupposes more of the basic cognitive abilities an organism must have. It supposes an ability to individuate objects, relations and locations; an ability which is taken on a par with the ability to classify situations. The perspective account has had nothing to say so far about these other cognitive abilities. In this chapter we show how a primitive kind of individuation can be characterized using perspectives.

In Section 5.2 we use the idea of information flow within a perspective to say when the classification of one situation can be predicted from the classification of another. This leads to the definition of when a sequence of situations is predictively coherent, meaning that the classification of each situation in the sequence is predictable from the previous one.

In Section 5.3 we define an object of a perspective to be a sequence of situations that is predictively coherent. This gives substance to the ST idea of objects being 'uniformities' across situations. Here the 'uniformity' is measured in terms of predictability.

In Section 5.4 properties are characterized as perspective shifts and the possession of a property by an object amounts to the uniformity that is the object persisting under the shift in perspective.

In Section 5.5 we see how a collection of perspectives (together with shifts) over a domain of situations can characterize some of the aspects of a scheme of individuation. A notion of basic infon is defined and it is shown how the perspectival domain determines which situations support which infons. Finally, some unusual logical properties of these definitions are uncovered.

Technical Point For the definitions in this chapter to make sense they must be restricted to the classified situations of a perspective. Our analysis of individuation and predication is based on the idea of one situation predicting things about another. Because nothing is known about unclassified situations, they make no predictions about other situations and satisfy every prediction made about them. This vacuous satisfaction of predictions would lead to absurd results, so we confine our attention to classified situations. The restriction is not so bad, since, as we saw in Section 4.1.3, any pers-
CHAPTER 5. INDIVIDUALS

5.2 Predictive Coherence

The $\Rightarrow$ and $\perp$ relations in a perspective are law-like in that the properties of facticity and local preclusion hold; but they are in general not 'deterministic'. For any one situation, a perspective provides a set of types $\mathcal{I}$ located at that situation, from which information will flow concerning the existence of situations of certain other types. But the extent to which the laws determine the types of situations is no stronger than that; from a particular situation the world could be formed in many different ways, all of which would be consistent with the law-like structure of the perspective.

Because of the lack of determinism within a single perspective, we can make sense of predictions driven by the informational structure either succeeding or failing. What is meant by a prediction within a perspective? From the armchair there are no distinctions between different parts of the world: all situations have equal status, so the location of information is relatively unimportant. For finite beings, it is crucially important that the world is made up of parts. We constantly draw and redraw epistemological boundaries around aspects of the world that interest us: this is what we know; what can we say about that? The world is repeatedly divided into the known and the unknown. It is this that allows the possibility of prediction.

Given a perspective, a predictive task will be characterized by labeling one situation $s_a$ as "known" and one other $s_b$ as "unknown". The positive and negative information carried by $s_b$ will be taken to be predictions about the information located at $s_a$. The criterion for a successful prediction is that the type sets $\mathcal{I}_k$ and $\mathcal{I}_l$, are related by a predictive regularity. The following definition defines what we mean by a predictive regularity. For technical reasons we relativize the definition to a subset $W$ of the types.

Definition 5.1 A pair $(X,Y)$ of sets of types of a perspective $P$ form a predictive regularity relative to a set $W$ of types, written $X \leadsto Y|W$ if for all $x \in X$ and $t \in W$
CHAPTER 5. INDIVIDUALS

• if $x \Rightarrow t$ then $t \in Y$
• if $x \Rightarrow t$ then $t \notin Y$

In the case where $W = T^P$ we just write $X \leadsto Y$.

A pair of situations $(s_1, s_2)$ is predictively coherent within $P$, written $P[(s_1, s_2)]$ iff $s_1 \leadsto s_2$.

In other words, the prediction will be successful if the positive information carried by $s_1$ is located at $s_1$ and the negative information is not. If this is the case then the pair of situations $(s_1, s_2)$ is said to be predictively coherent within the perspective. The following results about predictive regularity and coherence can be taken as more exercises for the enthusiastic reader.

**Proposition 5.2** Given a perspective $P$ and $X, Y, W \subseteq T^P$,

1. If $X \leadsto Y | W$ and $W' \subseteq W$ then $X \leadsto Y | W'$
2. $X \leadsto Y | W$ iff $((X)^+ \cap W) \subseteq Y \subseteq (T^P - ((X)^- \cap W))$
3. $P[(s_1, s_2)]$ iff if $s_1$ carries $t$ then $s_2 : t$ and if $s_1$ precludes $t$ then $s_2 \not\vdash t$
4. $P[(s, s)]$ iff $s$ is a full situation in $P$.
5. if $\rho : P \rightarrow P'$, then in the shifted perspective $P\rho$, $P\rho[(s_1, s_2)]$ iff $\rho(s) \leadsto \rho(s_2)$ in $(P \rho)$.
6. if $\rho : P \rightarrow P'$ is a strong shift and $X \leadsto Y | W$ then $\rho(X) \leadsto \rho(Y) | \rho(W)$ in $P'$.
7. strong shifts preserve predictive coherence, i.e., for strong $\rho$, if $P[(s_1, s_2)]$ then $P\rho[(s_1, s_2)]$.

**5.3 Objects as Uniformities**

We will exploit the condition of predictive coherence to give a definition of objects as uniformities across situations. A sequence of situations within a perspective will be taken as a uniformity if it is locally predictively coherent; i.e., each pair of successive situations in the sequence is predictively coherent.
Definition 5.3 A sequence \( \langle s_0, \ldots, s_n \rangle \) of situations in a perspective \( P \) (i.e., \( s_i \in S^P \) for \( i \in n^* \)) \(^2\) is a uniformity in \( P \), written \( P[s] \), iff for all \( i, i^* \in n, s_i \sim s_{i^*} \).

The pair \( (P, \langle s_0, \ldots, s_n \rangle) \) is said to be an object iff \( \langle s_0, \ldots, s_n \rangle \) is a uniformity in \( P \). The object \( (P, \langle s_0, \ldots, s_n \rangle) \) is illustrated in Figure 5.1. If \( a \) is the object \( (P, \langle s_0, \ldots, s_n \rangle) \) then we write \( \sigma^a \) for \( \langle s_0, \ldots, s_n \rangle \) and \( \pi^a \) for \( P \). We are also inclined to say that \( a \) is an object in \( P \) when \( \pi^a = P \).

If \( a \) and \( a' \) are objects in the same perspective, i.e. \( \sigma^a = \pi^{a'} \), and \( \sigma^a \) is a subsequence of \( \sigma^{a'} \) then we say that \( a \) is a subobject of \( a' \) and write \( a \subseteq a' \).

The motivation for regarding objects as being locally predictively coherent sequences of situations is based on the intuition that an object is something with temporal (or conceptual) duration. If we compare a vase of flowers with a gust of wind we feel that the former is somehow more of an object than the latter. Our explanation for this is that the vase of flowers, as a sequence of situations, is more predictively coherent: from one moment to the next we can be more sure of what to expect the vase to do than the gust of wind. In fact, the gust of wind is only individuated as a gust if it is sufficiently distinguished from the general flow of air around us. It comes into being

\(^2\)By \( n^* \) we mean the successor of \( n \), i.e., \( n + 1 \)
and disappears in an instant. Within a perspective concerned with local changes in air velocity, no information flows out of the situation that is the gust: it is full, and so by Proposition 5.2.4 a uniformity and therefore an object!

A consequence of our definition of objects is that any subobject of an object is an object. This, amongst other things, follows directly from the following simple Proposition.

**Proposition 5.4** Given a perspective $P$ and $\sigma = (s_0, \ldots, s_n)$ where $s_i \in S^P$ for each $i \in n^+$,

1. $P[\sigma]$ if for all $i \in n$, $P[(s_i, s_{i+1})]$.
2. If $P[\sigma]$ and $\sigma'$ is a subsequence of $\sigma$ then $P[\sigma']$ so $(P, \sigma') \subseteq (P, \sigma)$.
3. the empty sequence, $()$, is a uniformity in any perspective $P$, so $(P, (l))$ is an object.
4. for any $s \in S^P$, $(P, (s))$ is an object.
5. $(P, \{s_1, \ldots, s\})$ is an object if $s$ is full in $P$.

The vase is an object with temporal duration, but so is the vase-segment that I am looking at now. I am not seeing a part of an object, but a fully determinate one. If I struck it with a hammer and it broke into a thousand pieces I would not have been cheated from seeing a 'whole' vase. It is just that vase-segments that might have been, now will not be.

Because we have motivated perspectives from ideas about perception, it may be thought that we are open to a Berkelean problem: what happens to an object when no one is looking at it? Our answer is straightforward. Perspectives are objective things, independent of any organism adopting them. If I go out of the room the vase continues to be a predictively coherent sequence of situations in the perspective which I have moved out of.

### 5.4 Properties as Shifts

In formalising predication within our framework we follow the intuition that the ascription of a property to an object is a matter of comparison. For example, Ben is identified as an object of interest by being a man; his behaviour, say, is uniform across situations
CHAPTER 5. INDIVIDUALS

with respect to certain regularities associated with being a man. If men are similar to wolves in some respect, then we say that Ben is a wolf if his behaviour seen through this comparison is uniform with respect to certain regularities associated with being a wolf. The property of "being a wolf" is given by the comparison between those regularities which identify a subject and those of wolfishness.

In our theory, a comparison between two perspectives is a shift. We associate properties with shifts, and say that a certain object in a perspective $P$ has the property given by the shift $\rho : P \rightarrow P'$ iff the object is still a uniformity when shifted through $\rho$. For this to make sense the object must be an object in the source perspective of the shift. When this is the case we say that the object is appropriate for the shift.

Definition 5.5 Given a shift $\rho : P \rightarrow P'$ an object $a$ is appropriate for $\rho$ iff $\rho(a) = P$. For an appropriate $a$, we say that $a$ has the property $\rho$, and write $\rho(a)$, iff $\rho(a) = (s_0, \ldots, s_n)$ and for each $i \in n$, $\rho(s_i) \sim \rho(s_i')\tau(\rho)$.

Proposition 5.6 If $a$ is an object which is appropriate for $\rho : P \rightarrow P'$ then

1. $\rho(a)$ iff $P\rho(a)$
2. if $\rho$ is strong then $P\rho(a)$ so $\rho(a)$
3. if $\rho(a)$ and $a' \subseteq a$ then $\rho(a')$

The results of Proposition 5.6 can be interpreted as follows. 5.6.1 confirms our statement of what predication means: an object $a$ has the property $\rho$ iff $a$ is an object in the shifted perspective $P\rho$.

5.6.2 says that any appropriate object $a$ has the property $\rho$ if $\rho$ is a strong shift. As was hinted at in Section 4.1.3, we regard strong shifts as generalisations, so this can be interpreted as saying that uniformities are preserved from the special case to the more general.

5.6.3 is important in understanding the limitations of our model of properties. It says that if an object has a property then all of its subobjects have that property too. Put in a positive light this means that any property of the whole object is a property of its parts: if we are mortal then we are mortal now, not just over the whole span of our lives.
CHAPTER 5. INDIVIDUALS

Put negatively it means that the only properties an object has are those that every part of it has: we cannot be mortal because this is not a property that can be determined from a small segment of our lifetime. We acknowledge this limitation and so we do not claim to be providing a model of all kinds of property. However, it could be that a different understanding of the ‘part-of’ relation goes some way towards overcoming these limitations.

5.5 Basic Infons Revisited

Situation theory is a theory in which information is localised to situations. The theory aims to tell us more about a piece of information than that it is a fact: to be a fact it must be supported by a particular situation. The way this tends to be explained is as follows. A scheme of individuation determines the objects and relations present in a situation. This raises various issues as to which objects stand in which relations. For example, given a binary relation \( r \) and appropriate objects \( a \) and \( b \) there is an issue as to whether or not \( a \) and \( b \) stand in the relation \( r \). The situation resolves these issues either positively or negatively or not at all. If situation \( s \) resolves the issue of whether or not \( a \) and \( b \) stand in the relation \( r \) positively then \( s \) is said to support the basic infon \( \langle (r; a, b); + \rangle \). If \( s \) resolves the issue negatively then it is said to support the basic infon \( \langle (r; a, b); - \rangle \). On the other hand, \( s \) may have nothing to say about the issue, in which case it supports neither infon.

So far our notion of predication has been a global one. We have defined what it is for an object (individuated as a uniformity across situations) to have a property (a comparison between perspectives). But the possession of a property is not dependent on any one situation, although being an object is dependent on a perspective and being a property is dependent on two perspectives (the source and target of the shift). To create basic infons we need a way in which the information that an object has a property can be located at a particular situation. Our approach is to associate with each situation the collection of situations which are ‘part of’ it. We consider the ‘part of’ relation to be a non-perspectival relation between situations in a sense that will soon be made clear (in Definition 5.7). The objects in a situation \( s \) will be those objects which are uniformities across situations that are part of \( s \). Which objects a situation has will be perspectival,
since objects are only defined relative to a perspective.

In order to make these ideas clear we will need to consider a collection of perspectives with a 'part of' relation defined on their situation domains. We call such structures \textit{perspectival domains}.

\textbf{Definition 5.7} A \textit{perspectival domain} is a structure $D = (S, \preceq, \mathcal{P})$, where $S$ is a set (of situations) partially ordered by $\preceq$ and $\mathcal{P}$ is a collection of perspectives each with situation domain contained in $S$.

We define the set of objects and properties in $D$ by

\begin{align*}
\text{Obj } D &= \{ \text{object } | \sigma^+ \in \mathcal{P} \text{ and } \sigma^+ \text{ is made up of situations in } S \}, \\
\text{Prop } D &= \{ p : P \leftarrow P' | P, P' \in \mathcal{P} \}.
\end{align*}

Now in order to decide which basic infons a situation $s$ supports, we first have to determine which objects occur in $s$. An object is a sequence $(s_0, \ldots, s_n)$ of situations individuated by some perspective: it occurs in $s$ if some of the $s_i$'s are part of $s$. For example, the left hand part of Figure 5.2 illustrates some objects in a perspectival domain with situation set $S$. They are shown as strings of situations (the black dots), labeled $a, b, c, d, e, f$ and $g$. We can think of a situation $s \in S$ as providing a 'window' onto $S$ through which only those situations which are part of $s$ can be seen. In the diagram this window is represented by a shaded circle. The objects occurring in $s$ are...
CHAPTER 5. INDIVIDUALS

those which can be seen through the window: in this case, b, c, d and e. Formally, the condition for being an object in a situation is given by the following definition.

Definition 5.8 Given a perspectival domain \( D \) and a situation \( s \) of \( D \) we define the \textit{window} \( s_2 \) of \( s \) to be the set of situations that are part of \( s \). We say that an object \( a \in \text{Obj}(D) \) occurs in \( s \) iff \( \sigma^a = (s_0, \ldots, s_n) \) and \( \{s_0, \ldots, s_n\} \cap s_2 \neq \emptyset \), i.e. for some \( i \in \pi^+ \), \( s_i \subseteq s \). We call the set of objects of \( D \) occurring in \( s \) the set \( \{s\} \) of \textit{components} of \( s \).

Note that we do not need to see the 'whole' of any object through the window in order for it to occur in \( s \). The vase in my room is an object with a lifetime of, let us say, several years, but it does not have to have spent all that time in my room in order to be an object in my room now. Because of this, it may be that an object occurs in a situation \textit{more than once}. If I take the vase out of the room and then bring it back sometime later then there are two 'maximal' subobjects of the vase in the room situation, one consisting of the vase before I took it out and the other consisting of the vase after I brought it back. An example of this phenomenon is illustrated in Figure 5.2 with the object labeled e. To keep track of multiple occurrences, we associate objects with the set of their occurrences in a situation. The occurrence of objects in a are illustrated in the right half of Figure 5.2.

Definition 5.9 Given a set of situations \( W \subseteq S \) and an object \( a \) of \( D \), the set of maximal subobjects of \( a \) lying entirely within \( W \) is called \( \mathfrak{a}(W) \). More precisely,

\[
\mathfrak{a}(W) = \{a' \subseteq a \mid \sigma^{a'} = (s_0, \ldots, s_n), s_0, \ldots, s_n \in W \text{ and } \forall s \in W \langle \pi^+, (s_0, \ldots, s_n, s) \rangle \nsubseteq a\}
\]

Given a situation \( s \in S \) we say that \( \mathfrak{a}(s) \) is the set of occurrences of \( a \) in \( s \). \(^3\)

(The following is a list of results about the interaction between our notions of object occurrence, subperspective and property. They are needed for the technical development, but may easily skipped by the general reader. Alternatively they can be proved as exercises.)

\(^3\)As expected, the object \( a \) occurs in \( s \) iff the set of its occurrences in \( s \) is non-empty.
Proposition 5.10 Given a property $p : P \rightarrow P'$, an object $a$ in $P$, and a subperspective $P_0 \leq P$, let $\rho_0$ be the restriction of $\rho$ to $P_0$, then

1. $\rho_0$ is a shift.
2. $P_0 \rho_0 \leq P \rho$
3. if $a_0 \in a|_{P_0}$ then $P_0[\sigma^{a_0}]$.
4. if $a_0 \in a|_{P_0}$ and $p[a]$ then $\rho_0[a_0]$.
5. if $\sigma^a$ is made up of situations in $S|_{P_0}$ then $(P_0, \sigma^a)$ is an object, and if $p[a]$ then $\rho_0[(P_0, \sigma^a)]$.

We are now ready to show how basic infons are represented in our theory. Firstly, it should be noted that our treatment is restricted to infons built from unary relations (properties). For a situation $s$, we have shown how to determine which objects occur in $s$. If $a$ is an object occurring in $s$ and $p$ is a property for which $a$ is appropriate then the issue is raised as to whether or not $a$ has property $p$ (in $s$). The resolution of this issue will clearly be dependent on whether or not the occurrences of $a$ in $s$ have property $p$. But we must be wary of the case where $a$ has more than one occurrence in $s$. Suppose that the vase is black before I take it out of the room, but has been painted white when I bring it back. Does the 'room situation' support the information that the vase is black or that the vase is not black or neither? Here we pursue the 'not black' option, although the other alternatives are also worth investigating. The general rule is that a situation supports the information that an object has a property whenever all the occurrences of the object in the situation have the property. Similarly, a situation supports the information that an object does not have a property whenever there is some occurrence of the object in the situation which does not have the property.

Definition 5.11 Basic Infons In a perspectival domain $D = (S, \mathcal{A}, \mathcal{P})$ an issue is raised in situation $s \in S$ iff there is an object $a \in |s|$ which is appropriate for a property $p \in \text{Prop } D$. Then

$s |= \langle p; a; + \rangle$ iff for all $a' \in a|_{S_a}, p[a]$.

$s |= \langle p; a; - \rangle$ iff for some $a' \in a|_{S_a}, \neg p[a]$.

$\langle p; a; + \rangle$ and $\langle p; a; - \rangle$ are called basic infons and $\models$ is read 'supports'.

CHAPTER 5. INDIVIDUALS
CHAPTER 5. INDIVIDUALS

The final proposition of this section tells us something about the kind of partiality we have modeled. The result is an illustration of the kind of results that we might expect from this approach. Much depends on what we choose to say about multiple occurrences of objects.

Proposition 5.12 Given an object $a$ which is appropriate for a property $\rho$ and situations $s$ and $s'$,

1. not both $s \models (\rho; a; +)$ and $s \models (\rho; a; -)$

2. if $s' \subseteq s$ then
   - if $s \models (\rho; a; +)$ and $a \in |s'|$ then $s' \models (\rho; a; +)$
   - if $s' \models (\rho; a; -)$ then $s \models (\rho; a; -)$

In other words, if a situation $s$ supports some positive information about an object $a$ then that information is also supported by any part of $s$ in which $a$ occurs: positive infons are downwardly persistent. Negative infons are upwardly persistent, without the caveat since an object occurring in a situation will also occur in any situation of which it is a part.

3. if $s' \subseteq a$ then
   - if $s \models (\rho; a; +)$ then $s \models (\rho; a'; +)$
   - if $s \models (\rho; a; -)$ then $s \models (\rho; a; -)$

This states a similar fact about the persistence of information with respect to the subobject ordering. Positive facts about an object apply to its subobjects. Negative facts about an object apply to any object of which it is a subobject.

4. if $a \in |s|$ then either $s \models (\rho; a; +)$ or $s \models (\rho; a; -)$.

   We see the limits of this form of partiality. Only if $a$ does not occur in $s$ at all will $s$ fail to support either that it has or that it does not have the property $\rho$.

5. if $s' \supseteq s$ and $a \in a|s'\_a$ then
   - $s' \models (\rho; a; +)$ iff $s \models (\rho; a; +)$
   - $s' \models (\rho; a; -)$ iff $s \models (\rho; a; -)$

If an object lies entirely within the window of a situation then any larger situation will support exactly the same information about the object, positive or negative.
These results give a hint at the kind of logical characteristics that are built in to this approach to individuation. What appears to be missing from the above account is a general treatment of properties. But in a way it is not surprising that generality is difficult to obtain. Both objects and properties are bounded by the perspectives we can have on them, so much so that it is unclear what criteria would decide on their identity across different perspectives. If two sequences of situations are individuated as objects by two different perspectives is there any fact of the matter as to whether they are parts of the same object or not?

The position that there is no fact of the matter is initially unattractive. But if we are accused of submitting defeat to the phenomenalists, we should quickly add that perspectives are not necessarily perceptual. Although our initial example of a perspective was a perceptual one, this is by no means the general case. If we are capable of thinking about a remote situation, then we are equally capable of adopting a perspective towards it. In fact, the indeterminacy of identity for situations (or sequences of situations) beyond our powers of thought does not seem so unreasonable.
Chapter 6

Worlds

In this chapter we consider a model of information flow which is different but related to both the ST model and Dretske's probabilistic models. Rosenschein (1987, 1989) is interested in designing robots by specifying constraints rather than by explicit internal representations. He therefore wants a model which is both very concrete (i.e. near to physics) and yet abstract enough to be able to specify complex conditions at a high level (i.e. near to semantics).

Rosenschein's (1989) model is made up of a set of (physical) locations which can be in various (physical) states at different times. The actual world is therefore a function assigning a state to every location at every time. This is the part that is close to physics. A specification of various environmental conditions is given in a (possibly modal) language which is interpreted in an algebra of world conditions (or propositions). This is the part that is close to semantics. The connection between the two parts is that the world conditions are taken to be sets of 'possible world-time' pairs, where a world is a total function from locations and times to states. A world is 'possible' if the distribution of states across the locations and times is physically possible.

The algebra on world conditions is just given by the normal set operations on the sets of possible world-time pairs. Implication is given by inclusion. The information content of a location's being in a certain state is given by the set of world-time pairs that are compatible with this.

In Section 6.1 we generalize Rosenschein's model to account for locations which are aggregates of other locations, 'locations' which are moving with respect to the frame
of reference (dynamic locations) and the behaviour of locations over time. At each stage we arrive at a new definition of information content. We use throughout the notion of a state distribution function which models the pattern of physical states over regions of space and time. These functions are compared by means of spatial and temporal 'invariants' which enable us to define complex behavioural 'states'.

In Section 6.2 the distribution functions are used to model situations. Rosenschein's world conditions are generalized to situation conditions. Intuitively, a situation condition holds between two situations if one situation satisfies some condition from the point of view of the other situation.

In Section 6.3 information flow between situations is characterized in terms of information channels. In this section and Section 6.4 the model of information flow along channels is developed and compared with alternatives based on Rosenschein's definition of information content and our generalizations of it.

In Section 6.5 we show how a perspective can be defined from a collection of information channels and a 'point of view'. Visual perspective is naturally thought of in this way.

6.1 World Systems

The world, for Rosenschein, is made up of locations, L, times, T, and states D. At any time each location is in a specific state determined by the world function \( w : L \times T \rightarrow D \).

The intuition for this metaphysics is presumably based on something like the following model of physical reality: locations can be thought of as points (or perhaps small quantized regions) of three dimensional physical space over which there is a distribution of energy (or matter) over time. The state of a location can be thought of as something like its energy level. No doubt a more sophisticated physicalist metaphysics could be given and Rosenschein's ideas could be extended to that setting. Here we will stay with the simple model, using 'physical state' and 'energy level' more or less interchangeably.

Rosenschein also makes use of a computational interpretation of L, T and D. The locations, L, are thought of as registers in a machine which hold values from D. T is thought of as computational time, as measured out by the machine's clock. Registers may hold different values at different times just as locations can be in different states. In
CHAPTER 6. WORLDS

fact, Rosenschein identifies the two interpretations, thinking of his robots as embedded in the physical world: their internal registers occupy physical locations and the values they store are determined by the physical state of those locations.

Attending the model of actual physical reality are two counterfactual notions, one local and one global. Firstly, for each location \( a \in L \) there is assumed to be a set \( D_a \subseteq D \) of 'possible states' that the location could be in as well as the actual state \( w(a,t) \) it is in. The set of all states, \( D \), is taken to be the union of all the \( D_a \). Seen this way we can say that any function from \( L \times T \) into \( D \) describes a 'way the world might be', with \( w \) being the way the world actually is. But only some of these functions describe distributions of energy which are physically possible. The distinction is modelled by a subset \( W \) of the functions from \( L \times T \) to \( D \), which provides the second counterfactual aspect to the model. \(^1 \) Elements of \( W \) are called 'possible worlds' and, in particular, \( w \in W \).

Definition 6.1 A world system, \( W = (L, T, D, W_0) \), consists of a set of locations, \( L \), a set of times, \( T \), a state function, \( D \), a set of possible worlds, \( W \), and a distinguished world, \( W_0 \in W \). The state function, \( D \), takes each location, \( a \in L \), to the set, \( D_a \), of possible states it can be in. We abuse our notion by using \( D \) to refer to the set of all possible states, \( \bigcup_{a \in L} D_a \). Each possible world, \( w \in W \), is a function which determines the state, \( w(a,t) \in D_a \), of each location, \( a \in L \), at each time \( t \in T \).

Example For a 'biological' example\(^2 \) of a world system consider Conway's game of life. Suppose \( L \) is a set of cells in an infinite grid which can be either 'alive' or 'dead'. Each generation a cell can come alive, die or stay the same as it was in the previous generation. The life and death of a cell are determined by various simple rules: a living cell dies if it is surrounded by too many or too few other live cells, and a dead cell comes alive if it is surrounded by enough live cells. At the start of a game some cells are chosen to be alive and successive generations are calculated according to the rules.

A game can be modelled as a function \( w : L \times N \rightarrow D \) where \( D \) is the set {'alive', 'dead'} of states a cell can be in and \( N \) is just the natural numbers. So, for example, if in game \( w \) in the 9th generation cell \( c \) was dead then \( w(c,9) = '\text{dead}' \). Suppose that we

\(^1\)Presumably the two notions are connected by the fact that for all \( w' \in W \), \( a \in L \) and \( t' \in T \), \( w'(a,t') \in D_a \).

\(^2\)Thanks to David Israel for suggesting this example.
play the game \( W_0 \). With different initial conditions, we would have had different games: \( W_0 \) is only one of all the possible games we could have had. Let \( W \) be the set of all possible games of life. The structure \((L,N,D,W,w_0)\) is a world system.

Information about the world is modelled by sets of world-time pairs, i.e. subsets of \( W \times T \). Rosenschein calls these world conditions, \( \Phi = \text{pow}(W \times T) \). A world condition is something that the world might satisfy at a given time. The idea is that \((w',t')\) is in world condition \( \phi \) iff \( \phi \) holds in world \( w' \) at time \( t' \). For example, in the game of life world system, at some stage in a game all the cells may be dead. This condition is modelled as the set of pairs \((w,n)\) where \( w \) is a game and \( n \) is a number and in the game \( g \) all the cells are dead at generation \( n \).

World conditions are not simply predicates of the 'total states' of worlds at different times. Whether or not a world \( w \) satisfies a condition \( \phi \) at time \( t \) may depend on more than just the states of all the locations at time \( t \). An example is the condition that all cells have been dead for the last ten generations. It would not be satisfied if the last living cell died out nine generations ago, but would then become satisfied at the next generation even though there would be no change in the instantaneous states of the cells.

The world conditions, \( \Phi \), form a boolean algebra under the usual operations of intersection, union and complementation. This ensures a classical interpretation of various logical operations on world conditions: conjunction, disjunction and negation. One world condition \( \phi_1 \) implies another \( \phi_2 \) just in case \( \phi_1 \subseteq \phi_2 \).

**Definition 6.2** Associated with each state \( v \) that a location \( a \) can be in, there is a condition \( M_a(v) \in \Phi \) which can be thought of as the 'total information content' of \( a \)'s being in state \( v \). It is the set of all world-time pairs \((w',t')\) compatible with \( a \)'s being in state \( v \), i.e.

\[
M_a(v) = \{(w',t') \mid w'(a,t') = v \}
\]

For any world condition, \( \phi \), we say that \( a \)'s being in state \( v \) carries the information that \( \phi \) iff \( M_a(v) \subseteq \phi \).

**Example** Suppose \( \phi \) is the condition that the moon is eclipsed. The interpretation of \( \phi \) consists of all the world-time pairs, \((w',t')\), such that at time \( t' \) in world \( w' \), the earth
is directly in between the moon and the sun. For a location \(a\) in state \(v\) to have the information that the moon is eclipsed it must be the case that \(M_a(v)\) is a subset of the interpretation of \(\phi\). So if, in world \(w'\), \(a\) is in state \(v\) at time \(t'\) (i.e. \(w'(a, t') = v\)) then the earth lies directly between the moon and the sun at time \(t'\). In other words, \(a\) 's being in state \(v\) is a reliable indicator of lunar eclipses. An example might be a location on the 'sunny' side of the moon, which is in state '1' when illuminated by sunlight and in state '0' otherwise. The location's being in state '0' is an indicator of lunar eclipses.

**Example** In the game of life there is very little information carried by a single cell's being alive or dead. Strong conditions like 'all the cells are dead' are of course refuted by a single cell's being alive, but no information is carried about other cells. If \(a\) and \(b\) are cells and \(M_a('alive') \subseteq M_b('alive')\) then \(a = b\). The situation is quite different when we consider collections of cells. If all the cells surrounding cell \(a\) are dead then \(a\) will die out in the next generation, so the surrounding cells' being dead carries the information that cell \(a\) is about to die. But to capture this kind of informational dependency, we must look at aggregate rather than single locations.

### 6.1.1 Aggregate Locations

Rosenschein suggests a way of extending the notion of a state to aggregates of locations. On the computational interpretation, if a single location is a register, then an aggregate location is a collection of registers. The state of the aggregate is determined in terms of the states of its component locations. Rosenschein defines the value of a pair \((a_1, a_2)\) of registers to be the pair of their values, i.e. if \(a_1\) is in value \(v_1\) and \(a_2\) is in value \(v_2\) then \((a_1, a_2)\) is in value \(\langle v_1, v_2 \rangle\). The definition is easily extended to sequences of registers.

On the physical interpretation, the state of a sequence of spatial locations is the sequence of energy levels of the component locations. This may not be specific enough for modelling physical phenomena. It is natural to think of the state of a spatial region as the distribution of the states over the region. For example, if the state of a point in a volume of fluid is given by the concentration of a certain chemical \(X\) at that point, then we would expect the state of the whole volume to be determined by the distribution of \(X\) in the fluid. If we adopted Rosenschein's definition then the state of the volume could only be specified relative to a certain ordering of its points. Not only could two different
orderings give different states, but two different volumes of fluid could judged to be in
the same state, gives suitably chosen orderings, even if the distribution of chemical $X$
in the two volumes was entirely different.

For our purposes Rosenschein's definition is both too specific and not specific enough.
It is too specific since the same set of locations can be said to be in different states
depending on the way the locations are ordered. It is not specific enough since two
aggregates can be in the same state, relative to appropriately chosen orderings, even
though, on our physical interpretation, they have very distinct properties. The reason
for the inappropriateness of Rosenschein's definition in the physical case is that the
notion of a spatial distribution of states takes into account more than just the states of
the component locations. The geometrical relationships between component locations
is also important.

Our approach is to model the distribution of states over an aggregate directly, and define
an equivalence relation between distributions by looking at spatial transformations of
various kinds. Two aggregates can then be said to be 'in the same state' just in case
they have equivalent distributions.

**Definition 6.3** An aggregate location is a set $b \subseteq L$ of locations. Its state distribution,
$\delta_{(b,w,t)}$ in world $w$ at time $t$, is the function from $b$ to $D$ taking $a \in b$ to $w(a,t)$.

The choice of spatial transformations very much depends on the structure of space itself.
Like Rosenschein, we have abstracted away from any specific model of space: $L$ is just a
set of locations, not a vector space nor a topology nor even a partial order. If we were to
give $L$ one of these structures then our choice of transformations of $L$ would consist of
those maps $g : L \rightarrow L$ which 'preserved the structure' of $L$, in an appropriate sense (see
the examples below). In the absence of any specific structure for $L$ we provide a list of
properties that we would expect any suitable set of structure preserving transformations
to have. We call a set of transformations of $L$ which has these properties a set of 'spatial
invariants'.

**Definition 6.4** A set $G$ of bijections $g : L \rightarrow L$ is said to be a set of spatial invariants
iff
CHAPTER 6. WORLDS

1. $G$ contains the identity transformation, is closed under composition and has inverses (i.e. $G$ is a transformation group).

2. Homogeneity. For any two locations $a_1, a_2 \in L$ there is a $g \in G$ such that $g(a_1) = a_2$.

3. If $w \in W$ then for all $g \in G$ the world, $w' : L \times T \rightarrow D$, obtained by transforming $L$ under $g$ (i.e. for $a \in L, t \in T, w'(ga, t) = w(a, t)$) is also possible (i.e. $w' \in W$).

Condition 1. is satisfied by anything which can reasonably be said to be a set of transformations. Condition 2., the `homogeneity' condition, says that space is uniform. From any location you can find a transformation which gets you to any other location, so no location is unfairly privileged. Condition 3. is included to ensure that 'physical laws' which determine the set $W$ of possible worlds are blind to these transformations: if we transform the whole world then what we get is still a physically possible world.

Example Suppose that we model physical space as a three dimensional vector space, $L$. A useful class of invariants of $L$ is the translations, that is the maps

$$\{g_a : L \rightarrow L \mid \forall a \in L \text{ and } \forall a' \in L \ g_a(a') = a + a'\}.$$  
(Here + is vector addition). The translations also capture the homogeneity of $L$, since for any two locations $a_1, a_2 \in L$ there is a translation $g_{a_2-a_1}$ which maps $a_1$ to $a_2$.

Another example is the class of rigid transformations of $L$ (translations plus rotations). We can now say what it is for two aggregate locations, $b_1$ and $b_2$ to be in the 'same state': there must be some spatial invariant which maps elements of $b_1$ to elements of $b_2$ and preserves their elementary states. We describe this more precisely as an equivalence relation between 'distribution' functions.

Definition 6.5 State distributions $\delta_1 : b_1 \rightarrow D$ and $\delta_2 : b_2 \rightarrow D$ are equivalent iff there is a $g \in G$ such that $\delta_2 = \{g(a) \mid a \in b_1\}$ and for each $a \in b_1$, $\delta_1(a) = \delta_2(g(a))$. Two aggregate locations are in the same state iff their distribution functions are equivalent. We model the states of aggregate locations as equivalence classes of distribution functions.

An aggregate $b$ is in state $d$ in world $w$ at time $t$ iff $\delta_{b,w,t} \in d$.

Example
In the case where \( L \) is a vector space and \( G \) is the set of translations, \( g_a \), of \( L \), we can find a canonical representative of each distribution equivalence class. For the finite aggregate \( b \subseteq L \), define

\[
\text{centre}(b) = \frac{1}{|b|} \sum_{b} b
\]

where \(|b|\) is the number of locations in \( L \). Although the definition relies on \( b \) being finite, it can be extended to the infinite case. Informally, \( \text{centre}(b) \) is the geometric centre, or 'balance point' of the aggregate. Each aggregate can be 'normalized' by translating it so that its centre is at the origin of \( L \). We define

\[
\text{norm}(b) = \{a - \text{centre}(b) \mid a \in b\}
\]

Note that if there is a translation, \( g_b \in G \), taking \( b_1 \) to \( b_2 \) (i.e. \( g_b(b_1) = b_2 \)) then \( \text{norm}(b_1) = \text{norm}(b_2) \). Now the equivalence classes of distributions can be canonically represented by those distribution functions with a 0 centre, i.e. \( \{\delta : L \rightarrow D \mid \text{centre}((\text{norm}\delta)) = 0\} \). Aggregate location \( b \subseteq L \) has canonical state \( \delta \) iff \( \delta \) is the distribution function of \( \text{norm}(b) \).

Such canonical representatives of states are not available in the general case. For example, if we include rotations in the set \( G \) then it is difficult to choose a 'preferred' orientation for the representative of a class of rotation equivalent distributions.

Note how the homogeneity condition is important if aggregate states are to be conservative with respect to the elementary states of single locations. If \( a_1, a_2 \in L \) are locations in the same state, \( d \), then the distribution functions of the singleton aggregates \( \{a_1\} \) and \( \{a_2\} \) are just the functions assigning \( d \) to \( a_1 \) and \( a_2 \) respectively. These functions are equivalent iff there is a \( g \in G \) such that \( g(a_1) = a_2 \). The existence of such a \( g \) for each pair \( a_1, a_2 \in L \) is precisely what is guaranteed by homogeneity.

The definition of information content for aggregate locations must be modified to:

\[
M_b(d) = \{(w, t) \mid \delta_{b,w,t} \in d\}
\]

### 6.1.2 Dynamic Locations

The set of locations, \( L \), provides a means of dividing up the world in a uniform way which allows information to flow. Rosenschein's definition of information content compares the
current state of a location with its state at other times (and in other worlds) in order
to constrain the possible states of its environment. It is not the location’s current state
which alone determines this constraint, but all of its possible histories.

On the computational interpretation, a location is a register, and its information con-
tent is the strongest world condition that must be satisfied given only that it is in its
current state. Rosenschein uses this interpretation to ascribe knowledge (that the cur-
rent environment satisfies this world condition) to a robot containing the register. The
identification between location and register works fine as long as the robot is not moving
relative to the environment (or whatever is at rest with respect to the location set L).
If the robot moves then the register will occupy different locations at different
times.

We could avoid this problem by modelling the world from a robot’s eye view, in which
case all locations would be at rest with respect to the robot. But this does not seem
satisfactory since we may want to model several different robots, or even a single robot
with ‘moving parts’.

Our proposal is to consider dynamic locations in addition to the locations in L, which
we now call static. A dynamic location is a partial function \( a : T \rightarrow L \). We write
the collection of dynamic locations as \( L(T) \). Every static location can be recaptured
as a total, constant valued dynamic location. The definition of information content for
dynamic locations must be modified to:

\[
M_s(d) = \{(w, t) \mid t \in \text{dom} \text{ and } w(a(t), t) = d\}
\]

Example Suppose that LISA, the least intelligent situated automaton, is wandering
aimlessly around the world. LISA has one register which is programmed to respond in
a very simple way to the light conditions in LISA’s immediate environment. In well
lit environments LISA’s register is in state ‘1’. In not so well lit environments LISA’s
register is in state ‘0’. As LISA wanders her register flips from ‘1’ to ‘0’ and back again
approximately once a day. We model LISA’s register by the dynamic location \( r : T \rightarrow L \)
which, for each moment \( t \) of LISA’s life, determines the (static) spatial location \( r(t) \)
of the register. Let \( \phi \) be the world condition: ‘in world \( w \) at time \( t \) LISA is in a well

\footnote{We use the notation \( A^{(B)} \) for the set of partial functions from \( B \) to \( A \).}
For LISA to do what she is supposed to do, it must be the case that
\( M_r(1) \subseteq \phi \) and \( M_r(0) \subseteq \neg \phi \).

6.1.3 Behaviour Over Time

Another extension to Rosenschein's model of information concerns the information content of a location's behaviour over time. In many cases of practical interest it is not the state of a location that conveys information, but a pattern of states adopted by a location over a period of time. For example, the instantaneous state of a telegraph wire transmitting Morse code tells us nothing: it is the sequence of states (off, on, off, on, off, on, off, on, on, off, on, off, on, off, on, off, on, off) that conveys a call for help. Similarly, it is the change in height of the column of fluid in a thermometer which carries the information that the temperature is rising.

What is meant by a 'period of time' may well depend on the structure of time itself. For example, if we model time as a linear order we may only consider intervals to be genuine time periods. But, as in the space case, we will abstract over particular models of time and allow any subset \( r \) of \( T \) to be a time period. Over time period \( r \) a (static) location \( a \) may be in various instantaneous states. The question to be answered in this section is how we extend the notion of 'state' to \( a \)'s behaviour over \( r \). The solution we adopt is similar to that for aggregate locations, using distributions over time rather than space and temporal, rather than spatial, invariants.

Definition 6.6 The distribution \( \ell_{(w,r,a)} : r \to D \) of a location \( a \) over a time period \( r \) in world \( w \) is the function taking each \( t \in r \) to \( w(a,t) \).

Definition 6.7 A set \( H \) of bijections \( h : T \to T \) is a set of temporal invariants for \( T \) iff

1. \( H \) contains the identity map, is closed under composition and has inverses.
2. Homogeneity. For any \( t_1, t_2 \in T \) there is an \( h \in H \) such that \( h(t_1) = t_2 \).
3. If \( w \in W \) then for all \( h \in H \) the world, \( w' : L \times T \to D \), obtained by transforming \( T \) under \( h \) (i.e. for \( a \in L \), \( t \in T \), \( w'(a,ht) = w(a,t) \)) is also possible (i.e. \( w' \in W \)).

Example
CHAPTER 6. WORLDS

For example, if we model time as a one dimensional vector space, then we can take $H$ to be the set of translations in $T$; i.e. for each $t \in T$, there is an $h_t \in H$ such that for all $t' \in T$, $h_t(t') = t + t'$. Homogeneity is satisfied since for any $t, t' \in T$, $h_{t+t'}(t) = t'$. Slightly more abstractly, if we model time as an order $(T, \leq)$ then we could take the set of order preserving bijections on $T$ as our set of temporal invariants. The homogeneity condition constrains the kind of order $\leq$ can be.

The definitions of distribution equivalence, state (in this case, 'behaviour') and information content follow the established pattern:

**Definition 6.8** Distributions, $\delta_1 : \tau_1 \to D$ and $\delta_2 : \tau_1 \to D$ are equivalent iff there is an $h \in H$ such that $\tau_2 = \{h(t) \mid t \in \tau_1 \}$ and for each $t \in \tau_1$, $\delta_1(t) = \delta_2(h(t))$. The behaviour of a location $a$ over time period $\tau$ in world $w$ is the equivalence class of $\delta_{(a, \tau, w)}$. The information content of location $a$'s exhibiting behaviour $d$ is given by

$$M_a(d) = \{(w, \tau) \mid \delta_{(a, \tau, w)} \in d\}$$

Note that $M_a(d)$ is not a world condition in the technical sense we have been using so far. The elements of $M_a(d)$ are pairs consisting of a world and a time period. The new set $\Phi^a$ of world conditions are therefore to be thought of as conditions which a world satisfies over a time period. Any 'instantaneous' world condition $\phi \in \Phi$ can be recaptured as the world condition $\{(w, \{t\}) \mid (w, t) \in \phi\}$. It may be that the only conditions of any interest are those which can be reduced to instantaneous conditions in a uniform way, e.g. those $\phi \in \Phi$ for which $(w, \tau) \in \phi$ iff for each $t \in \tau$, $(w, \{t\}) \in \phi$.

But, for the sake of generality, we will not adopt such a reduction here.

**Example** Suppose that the temperature of a boiler is monitored by a thermometer which lights an LED when the temperature rises above a certain point. The LED goes out when the temperature falls below another (lower) point. If the temperature gets dangerously high, the LED starts to flash on and off to attract the attention of an operator. If we model the location of the LED by $a$, then 'flashing' behaviour can be captured by the equivalence class $[\delta]$ of the distribution $\delta : \{t_1, t_2, t_3\} \to D$ given by $\delta(t_1) = \text{'on'}, \delta(t_2) = \text{'off'}, \delta(t_3) = \text{'on'}$ (for a suitably proximate sequence of times $t_1, t_2, t_3$). Then

---

Note: If time is modelled as a vector space, then we can define canonical representatives of each equivalence class of temporal distributions in the same way as we did for spatial distributions.
6.1.4 Complex Behaviour

We have seen how information can be located across aggregates of locations, in moving ('dynamic') locations and in the behaviour of a location over time. In this section we face the task of analysing the informational content of complex behaviour in terms of these simple cases.

There are several ways of extending the definition of dynamic locations to aggregates of locations. The first is to use aggregates of dynamic locations, i.e. sets \( b \subseteq L^T \). We can define the state distribution of an aggregate \( b \) of dynamic locations in world \( w \) at time \( t \) to be the distribution of states across the aggregate of static locations which \( b \) occupies at time \( t \), i.e. the function with domain \( \{ a(t) : a \in b \text{ and } t \in 4m+3 \} \) mapping each \( a(t) \) to \( w(a(t), t) \). Since we do not require that all the dynamic locations in \( b \) are defined over the same time period, this 'instantaneous' aggregate may vary in size over time.

Example In the game of life there are certain patterns of cells, called gliders, which 'move' across the grid in straight lines: every four generations the pattern is replicated in a slightly different position. Figure 6.1 shows five generations of a glider's life. The labels illustrate one way of modelling the glider as an aggregate \( \{a_1, a_2, a_3, a_4, a_5\} \) of dynamic

Figure 6.1: A 'glider' modelled as the aggregate \( \{a_1, a_2, a_3, a_4, a_5\} \) of dynamic locations.

The flashing LED has information content \( M_q([b]) \) and if \( (w, t) \in M_q([b]) \) then the boiler is dangerously hot in world \( w \) over time period \( t \). In this case the reduction to the instantaneous world conditions mentioned above would be sensible since being dangerously hot over a (short) period of time involves being dangerously hot at each instant of that period.
locations. In each generation the cell marked by \( a_i \) is the instantaneous location of \( a_i \) in that generation. The distribution function in each generation will have the constant value 'live' but with a domain — the set of live cells of that generation — which changes over time. During the period depicted each dynamic location in the aggregate is defined at each generation and so the domain of the distribution function remains of constant size (5 cells).

The second approach uses dynamic aggregate locations, i.e. functions \( b : T \rightarrow \text{pow} D \), which determine an aggregate of 'static' locations \( b(t) \subseteq L \) at each time \( t \). There is no need to use partial functions in this case, since \( b \) can be taken to be 'undefined' when \( b(t) = \emptyset \). The state distribution of a dynamic aggregate location \( b \) in world \( w \) at time \( t \) is the function mapping each \( a \in b(t) \) to \( w(a, t) \).

Example The glider pictured in Figure 6.1 could also be modelled as a dynamic aggregate location, \( b \), which maps each generation \( n \) to the set of live cells \( b(n) \) of that generation. The distribution in each generation can be seen to be exactly the same as in the previous example.

In fact, as things stand, there is little to choose between these two approaches. For each aggregate of dynamic locations, \( b \), there is a dynamic aggregate location, \( \mathcal{V} : T \rightarrow \text{pow} L \), which determines the aggregate \( b(t) = \{ a(t) \in L \mid a \in b \text{ and } t \in \text{dom}\} \) at each time, \( t \).

It is easy to see that \( b \) and \( \mathcal{V} \) have the same distribution function in every world and at every time. In the reverse direction, for a dynamic aggregate location, \( \mathcal{V} \), there may be more than one aggregate of dynamic locations which corresponds in this way, but since our definition of the state of an aggregate only depends on its distribution function there is no real difference.

A third approach is to change the definition of distribution functions for aggregates of dynamic locations. If we take a distribution to be a map from dynamic locations to states, then a subtle difference emerges between the two previous approaches. There is more information associated with an aggregate of dynamic locations, \( b \), than with the corresponding dynamic aggregate location, \( \mathcal{V} \). This fact is illustrated by Figure 6.2. For times \( t_1, t_2 \in T \), there is no connection between locations in the aggregate \( \mathcal{V}(t_1) \) and those in \( \mathcal{V}(t_2) \), whereas \( b \) provides a connection between \( a(t_1) \in \mathcal{V}(t_1) \) and \( a(t_2) \in \mathcal{V}(t_2) \); they are instances of the same dynamic location. We loosely refer to the connections
between location instances in an aggregate of dynamic locations as *dynamic bindings*.

The intuition is that a dynamic location consists of 'static' locations together with information about how they are connected, or 'bound', in time. More precisely, we can think of an aggregate of dynamic locations as inducing a binary relation on space-time:

**Definition 6.9** Let $b$ be an aggregate of dynamic locations. For $l,l' \in L$ and $t,t' \in T$, we say that $(l,t)$ is *bound* to $(l',t')$ in $b$ iff $\exists a \in b$, such that $t,t' \in dom a$, $a(t) = l$ and $a(t') = l'$.

Because of the binding information contained in aggregates of dynamic locations we will use them to model moving aggregate. To make use of the extra information, we define the distribution of states across a moving aggregate to be an assignment of states to dynamic locations.

**Definition 6.10** The distribution $\delta_{b,w} : b \rightarrow D$ of an aggregate $b$ of dynamic locations in world $w$ at time $t$ maps each $a \in b$ to $w(a(t),t)$. Distributions $\delta_1$ and $\delta_2$ are *equivalent* iff there is a function $f : T \rightarrow G$ such that

![Diagram](image-url)
1. \( \text{dom} \delta_1 = \{ a_f | a \in \text{dom} \delta_1 \} \) where \( a_f \) is the dynamic location with \( \text{dom} a_f = \text{dom} a \) and \( a_f(t) = f(t)(a(t)) \) for each \( t \in \text{dom} a \), and

2. for each \( a \in \text{dom} \delta_1 \), \( \delta_1(a) = \delta_1(a_f) \).

The state of an aggregate \( b \) of dynamic locations in world \( w \) at time \( t \) is the equivalence class of \( \delta_{b,w,t} \). If \( b \) is in state \( d \) then it has information content

\[
M_6(d) = \{(w,t) | \delta_{b,w,t} \in d \}
\]

**Example** In the glider example, each of the five 'parts' of the glider are instantiated by different live cells in different generations. The bindings between these cells are represented in Figure 6.1 by the labels \( a_1, a_2, a_3, a_4, a_5 \). If we were to rub out those labels then there would be no means of connecting the live cells of one generation to those of the next. The state distribution of the glider at each generation is just the constant 'live'-valued function with domain \( \{a_1, a_2, a_3, a_4, a_5\} \). Since, according to our final definition, the glider never changes state, it's being in that state carries very little information about its environment.

Using similar techniques we can capture the information content of the temporal behaviour of aggregates and single moving locations, and (finally) the temporal behaviour of moving aggregates. We will spare the reader the details by merely stating and illustrating the notion of equivalence in the most complex case.

**Definition 6.11** The distribution \( \delta_{b,w,r} : L(T) \times T \rightarrow D \) of an aggregate \( b \) of dynamic locations in world \( w \) over time period \( r \) has domain

\[
\text{dom} \delta_{b,w,r} = \{ (a_r, t) | a \in b \text{ and } t \in r \cap \text{dom} a \}
\]

and, for each \( a \in b \) and \( t \in \text{dom} a \cap r \),

\[
\delta_{b,w,r}(a_r, t) = w(a, t)
\]

where \( a_r \) is the restriction of \( a \) to \( r \), i.e. the dynamic location with \( \text{dom} a_r = \text{dom} a \cap r \) mapping each \( t \in \text{dom} a_r \) to \( a(t) \).

Distributions \( \delta_1 \) and \( \delta_2 \) are equivalent iff there is a function \( f : T \rightarrow G \) and an \( h \in H \) such that
1. \( \text{dom} \delta_2 = \{ (a_f, ht) \mid (a, t) \in \text{dom} \delta_1 \} \) where \( a_f \) is the dynamic location with \( \text{dom} a_f = \{ ht \mid t \in \text{dom} a \} \) and \( a_f(ht) = f(t)(a(t)) \) for each \( t \in \text{dom} a \), and

2. for each \( (a, t) \in \text{dom} \delta_1 \), \( \delta_2(a_f, ht) = \delta_1(a, t) \).

The behaviour of an aggregate \( b \) of dynamic locations in world \( w \) over time period \( r \) is the equivalence class of \( b(b, w, r) \). If \( b \) has behaviour \( d \) then it has information content

\[
M_b(d) = \{ (w, r) \mid b(b, w, r) \in d \}
\]

Example Suppose that you and I collect two cars of the same model fresh from the Austin Rover assembly line. We fill them with the same amount of fuel and drive away in different directions. The engines in each of our cars are mechanically identical. Each is composed of a number of moving parts manufactured by the same machine. Each moving part traces out a trajectory in space and time, which we will model as a dynamic location. The engines of our two cars will therefore be modelled as aggregates of dynamic locations. Each part of each engine will be moving both with respect to other parts of the same engine and with respect to the corresponding part in the other engine. Nevertheless, when we are traveling at the same speed, there is a sense in which the behaviour of your car’s engine is the same as the behaviour of mine. From hypothetical positions inside the two engines they would be physically indistinguishable. It is precisely this sense of equivalence which is captured by Definition 6.11.

6.2 Situations and Situation Conditions

In this section we consider how situations can be modelled in Rosenschein’s framework. An event in the physical world consists of something happening at certain locations and certain times. On the physicalist metaphysics adopted in the previous section, anything that happens must involve a change in the state of a location (or dynamic location) over some time period. Rather than capturing the ontology of events, we follow situation theory by using the more neutral term situation to cover both events and states. We model situations by functions which take (dynamic) locations and times to elementary states.

Definition 6.12 A situation is a partial function, \( s : L(T) \times T \rightarrow D \), such that for all \( a, a' \in L(T) \) and \( t, t' \in T \)
1. if \( (a, t) \in \text{doms} \) and \( t' \in \text{doms} \) then \( (a, t') \in \text{doms} \) and \( t \in \text{doms} \), and

2. if \( (a, t), (a', t) \in \text{doms} \) and \( a(t) = a'(t) \) then \( s(a, t) = s(a, t') \).

Associated with each situation, \( s \), is its location, \( \text{loc} \), and its time period, \( \text{time} \). The location of a situation is the aggregate of dynamic locations over which it is defined: \( \text{loc} = \{ a \mid \exists t, \ (a, t) \in \text{doms} \} \). The time period of a situation is the set of points of time at which at least one of its dynamic locations is defined: \( \text{time} = \{ t \mid \exists a, \ (a, t) \in \text{doms} \} \).

Let \( \text{Sit} \) be the collection of all situations in the world system.

Condition 1. ensures that a situation determines the state of every location in \( \text{loc} \) at every time on which that location is defined, i.e. \( \text{doms} = \{ (a, t) \mid a \in \text{loc} \text{ and } t \in \text{doms} \} \).

This imposes an identity condition on situations: two situations are the same if they have the same 'underlying' distribution of states together with the same internal dynamic bindings (see Proposition 6.15).

Condition 2. ensures that there are no 'clashes'. If two dynamic locations in \( \text{loc} \) 'cross', i.e. they occupy the same place at the same time, then \( s \) must be consistent about the state it assigns to them. Another, way of avoiding such clashes would be to insist that the dynamic bindings in the location of a situation never cross, but that seems unnecessarily restrictive.

In comparing situations we must bear in mind that they are defined with respect to dynamic locations. In other words, each situation carries with it a spatio-temporal 'frame of reference' defined by the bindings between instantaneous events that are induced by its dynamic locations. It is quite possible that two situations with disjoint domains could 'overlap': i.e. they could assign the same states to certain 'static' locations at the same points of time. They would differ not because of the physical state of any region of space and time, but because they have different frames of reference.

In saying that one situation is 'part of' another, we will mean that it is fully contained within it, i.e. the smaller situation is made up of instantaneous events, all of which are included in the larger situation. Whatever bindings there are between the instantaneous events are irrelevant to its being a part of the larger situation. On this conception, it is possible for a situation to have a part whose domain is entirely disjoint from its own.

Such a part would be dependent on the same underlying distribution of elementary
CHAPTER 6. WORLDS

states, but would 'cut up the world' differently with its own regimen of bindings between instantaneous events.

It is important to admit this possibility for the sake of a modular understanding of parts. A situation is structured by the bindings it imposes on its elementary parts. We should not expect it to say anything about the structure of those parts. For example, a car engine can be modelled as an aggregate of piston rods, valves, cylinders and other components. In such a model a single piston rod would be represented by the dynamic location which maps out its spatial trajectory. But the same piston rod can itself be modelled as an aggregate of its parts: vibrating molecules or whatever. The situation modelling the internal behaviour of the piston rod is still a part of the situation modelling the engine, despite its distinct component structure.

For this reason, the 'part of' ordering between situations is not determined by set inclusion of their graphs. Instead we use the idea of a 'canonical' situation. A canonical situation is one in which the bindings are just those that are implicit in $L$ itself, i.e. $(I, t)$ is bound to $(I', t')$ iff $I = I'$. We proceed by defining an idempotent operator $*$ on $Sit$ which takes each situation $s$ to a canonical situation $s^*$. The $*$ operator can be thought of as an operator which 'forgets' about the dynamic bindings in a situation. The 'part of' relation is then determined by graph inclusion of canonical situations.

**Definition 6.13** The $*$ operator is defined indirectly using two auxiliary functions, $V$ and $A$.

1. For $s : L(T) \times T \rightarrow D$ let $s^* : L \times T \rightarrow D$ be such that

   (a) $\text{dom } s^* = \{(u(t), t) \mid (u, t) \in \text{dom} s\}$, and

   (b) for each $(u, t) \in \text{dom } s$, $s^*(u(t), t) = s(u, t)$.

   Note that $s^*$ is well-defined by the 'no clashes' condition (2.) of Definition 6.12.

   We call $s^*$ the underlying state distribution of $s$.

2. For $f : L \times T \rightarrow D$ let $f^* : L(T) \times T \rightarrow D$ be such that

   (a) $\text{dom } f^* = \{(a^\bullet, t) \mid (a, t) \in \text{dom } f\}$ where for $a \in L$, $a^\bullet : T \rightarrow L$ such that

   $\text{dom } a^\bullet = \{t' \mid (a, t') \in \text{dom } f\}$ and $a^\bullet(t') = a$ for all $t' \in \text{dom } a^\bullet$, and

   (b) for each $(a, t) \in \text{dom } f$, $f^*(a^\bullet, t) = f(a, t)$.
CHAPTER 6. WORLDS

$f^\wedge$ satisfies conditions 1. and 2. of Definition 6.12 and so is a situation. Note that $f^\wedge = f$.

Now, for a situation, $s$, let $s^* = s^\wedge$ and say $s$ is static iff $s^* = s$, and dynamic otherwise.

Finally, we say that situation $s_1$ is part of situation $s_2$, written $s_1 \sqsubseteq s_2$ iff $\text{dom} s_1^\wedge \subseteq \text{dom} s_2^\wedge$ and for all $(a, t) \in \text{dom} s_1^\wedge$, $s_1^\wedge(a, t) = s_2^\wedge(a, t)$.

Proposition 6.14 For all $s_1, s_2 \in \text{Sit}$,

1. $s^* = s^*$, i.e. $s^*$ is static
2. $s_1 \sqsubseteq s_2$ iff $s_1^\wedge \sqsubseteq s_2^\wedge$
3. $\sqsubseteq$ is a preorder
4. If $s_1 \not\sqsubseteq s_2$ and $s_2 \not\sqsubseteq s_1$ then $s_1^\wedge = s_2^\wedge$
5. $\sqsubseteq$ is a partial order on static situations

Proof: Strightforward application of Definition 6.13.

Proposition 6.15 Two situations are the same if they have the same ‘underlying’ distribution of states together with the same internal dynamic bindings, i.e. $s_1 = s_2$ iff $s_1^\wedge = s_2^\wedge$ and for all $(l, t), (l', t') \in \text{dom} s_1^\wedge (= \text{dom} s_2^\wedge)$, $(l, t)$ is bound to $(l', t')$ in $\text{loc} s_1$ iff they are bound in $\text{loc} s_2$.

Proof: If $s_1$ and $s_2$ have the same internal dynamic bindings then it is easy to show that $\text{loc} s_1 = \text{loc} s_2$. But the first condition on situations in Definition 6.12 then ensures that $\text{dom} s_1 = \text{dom} s_2$. Now for any $(a, t) \in \text{dom} s_1$, $s_1(a, t) = s_1^\wedge(a(t), t) = s_2^\wedge(a(t), t) = s_2(a, t)$ and so $s_1 = s_2$.

Situations can be regarded as partial specifications of Rosenschein’s possible worlds. Not all situations will correspond directly to parts of worlds since they may divide up space using a different ‘grid’ of dynamic locations than $L$. For situation, $s$, the partial function $s^\wedge$ may be extendible to a possible world, but it may not: there are
situations which distribute states in physically unrealizable ways. Conversely, we can lift Rosenschein's worlds to the level of situations using the map \( \Lambda \): for each world \( w \), \( w^\ast \in \text{Sit} \). This allows us to apply the terms 'actual', 'possible' and 'world' to situations in a conservative manner.

**Definition 6.16** If \( s \) is a situation then

- \( s \) is possible iff there is some \( w \in W \) for which \( s \preceq w^\ast \)
- \( s \) is actual iff \( s \preceq w^\ast_0 \)
- \( s \) is a world iff for all \( s' \in \text{Sit} \) if \( s \preceq s' \) then \( s' \preceq a \).

Let \( s^W \) to be the set of possible worlds that it occurs in, i.e. \( s^W = \{ w^\ast \mid w \in W \text{ and } s \preceq w^\ast \} \).

**Proposition 6.17**

1. For \( w : L \times T \rightarrow D \), \( w \in W \) iff \( w^\ast \) is a possible world.

2. For situations \( s_1 \) and \( s_2 \), \( s_2 \in s_1^W \) iff \( s_1 \preceq s_2 \) and \( s_2 \) is a possible world.

**Proof:**

1. If \( w \in W \) then \( w^\ast \) is possible since \( w^\ast \preceq w^\ast \) and \( w \in W \). For any situation \( s \), if \( w^\ast \preceq s \) then \( \text{dom}w^\ast \subseteq \text{dom}^s \) and for all \( (a, t) \in \text{dom}w^\ast \), \( s^*(a, t) = w^\ast(a, t) \). But \( w^\ast w^\ast = w^\ast w^\ast = w^\ast \) so \( \text{dom}w^\ast = \text{dom}w^\ast \) and \( w^\ast \) is maximal for static situations, so \( \text{dom}^s = \text{dom}w^\ast \) and \( s \preceq w^\ast \). Hence \( w^\ast \) is a world.

   Conversely, suppose that \( w^\ast \) is a possible world. Since \( w \) is possible there is a \( w_1 \in W \) such that \( w^\ast \preceq w_1^\ast \). But \( w^\ast \) is a world, so \( w_1^\ast \preceq w^\ast \), so by Proposition 6.14, \( w^\ast w^\ast = w_1^\ast w^\ast \) and so \( w = w_1 \in W \).

2. Suppose \( s_2 \) is a possible world and \( s_1 \preceq s_2 \). Since \( s_2 \) is possible there is a \( w \in W \) such that \( s_2 \preceq w^\ast \) and \( w^\ast \preceq s_2 \) since \( s_2 \) is a world. So, by Proposition 6.14, \( w^\ast w^\ast = s_2^\ast \). Also \( w^\ast w^\ast = w^\ast \), so \( s_2^\ast = w^\ast \). Now, since \( s_1 \preceq s_2 \) we have that \( s_2^\ast \preceq s_2^\ast \) and so \( s_1 \preceq w^\ast \), so \( s_1 \preceq w^\ast \) and so \( s_2 \in s_1^W \).

   Conversely, if \( s_2^\ast \in s_1^W \) then there is some \( w \in W \) for which \( s_2^\ast = w^\ast \) and \( s_1 \preceq w^\ast \), so \( s_1 \preceq s_2^\ast \) and so, by Proposition 6.14, \( s_2 \preceq s_2^\ast = s_2 \) and so \( s_1 \preceq s_2 \). Now for any situation \( s \), if \( s_2 \preceq s \) then \( w^\ast = s_2^\ast \preceq s^\ast \). But \( w^\ast \) is a world, by part 1., so
CHAPTER 6. WORLDS

\[ s^* \not\subseteq w^* \] and so \( s \not\subseteq s_2 \). Hence \( s_2 \) is a world. Finally, \( s_2 \) is possible since \( s_2 \not\subseteq w^* \).

The first part of Proposition 6.17 shows that the definition of ‘possible’ and ‘world’ in Definition 6.16 is justified. It also permits us in identifying \( w \) with \( w^* \) which we will do for the sake of cleanliness. From this point on, except in detailed proofs, we will talk about \( w \) meaning \( w^* \) and vice versa.

Situations have already been studied in Section 6.1.3 under a different name. They are just the state distribution functions of aggregates of dynamic locations over a time period. So, in addition to the \( \not\subseteq \) ordering, we already have a classification of situations into ‘behavioural’ types. The types are just the equivalence classes of the distribution functions under the equivalence of Definition 6.11. However, in order to generalize this classification of situations we need to be slightly more precise about the relationship between equivalent situations. To this end we will define a set of structure preserving maps for \( \text{Sit} \) called the set of situation invariants. A situation invariant will map situations to situations which are equivalent as distributions.

**Definition 6.18** A function \( \epsilon : \text{Sit} \rightarrow \text{Sit} \) is a situation invariant iff there is a function \( f : T \rightarrow G \) and an \( h \in H \) such that, for each \( s \in \text{Sit} \),

1. \( \text{dom} \epsilon = \{ (a_f, t) \mid (a, t) \in \text{dom} s \} \) where \( a_f \) is the dynamic location with \( \text{dom} a_f = \{ t \mid t \in \text{dom} a \} \) and \( a_f(t) = f(t)(a(t)) \) for each \( t \in \text{dom} a \), and

2. for each \( (a, t) \in \text{dom}s, \epsilon(s(a_f, t)) = s(a, t) \).

We say that two situations \( s_1, s_2 \in \text{Sit} \) are equivalent and write \( s_1 \equiv s_2 \) iff there is a situation invariant taking \( s_1 \) to \( s_2 \). A glance at Definition 6.11 will satisfy the reader that we are using our terminology consistently.

The equivalence relation between situations is intended to capture internal physical indistinguishability. Two situations are equivalent just in case it is physically impossible to distinguish between them by ‘looking’ at their internal structure alone: it is only their external relational properties which tell them apart. Internal relational properties, however, may be sufficient to distinguish situations even when they ‘occupy’ the same
physical space. For example, if $s_1 \triangleleft s_2$ and $s_2 \triangleleft s_1$ it may be possible to distinguish between $s_1$ and $s_2$ on the basis of the dynamic bindings in their locations: one will be seen as moving with respect to the other and vice versa. In other words, $s_1 \triangleleft s_2$ and $s_2 \triangleleft s_1$ does not entail $s_1 \equiv s_2$.

The role of a situation invariant is to specify how two situations are internally physically indistinguishable. A situation invariant provides a map between the location of one situation and the location of another situation from which it is indistinguishable. This map associates, point by point, locations and times in one situation with locations and times in another situation so that associated locations and times are assigned the same state in both situations. Moreover, it does this in a way that respects the underlying structure of space and time as represented by the spatial and temporal invariants, $G$ and $H$. The following proposition establishes that the set of situation invariants has all the properties we would expect of a set of invariants on $Sit$.

Proposition 6.19

1. The set of situation invariants contains the identity function, is closed under composition and has inverses.
2. If $e$ is a situation invariant then $s_1 \triangleleft s_2$ iff $e s_1 \triangleleft e s_2$.
3. If $w$ is a world and $s \equiv w$ then $s$ is also a world.
4. If $s$ is possible and $s' \equiv s$ then $s'$ is also possible.

Proof: Messy but straightforward.

Although space and time are taken to be homogeneous, the set of situations is not. Situations have a rich internal structure whereas points of space and time are taken to be atomic. So for situation invariants there is only a trivial version of the homogeneity property enjoyed by $G$ and $H$. Given two equivalent situations there is, by definition, a situation invariant which maps one to the other. The question arises as to whether this invariant is unique. If for each pair of equivalent situations there is only one situation invariant connecting them, then we say that the collection of situation invariants has the Uniqueness Property. Whether or not we have the uniqueness property depends on underlying assumptions about $G$ and $H$. 
Definition 6.20 The set of situation invariants satisfies the Uniqueness Property iff for any situation \( s \) and invariants \( e_1 \) and \( e_2 \), if \( e_1 s = e_2 s \) then \( e_1 = e_2 \).

Example If \( L \) and \( T \) are the euclidean vector spaces \( \mathbb{R}^3 \) and \( \mathbb{R} \), and \( G \) and \( H \) are the sets of translations on \( L \) and \( T \) respectively, then the induced collection of situation invariants has the Uniqueness Property. However, if we allow \( G \) to contain rotations on \( L \), then the Uniqueness Property no longer obtains.

Just as for distribution functions, we define the state of a situation, \( s \), to be the set \([s]\) of situations indistinguishable from it.

Definition 6.21 The collection \( \text{State} \) of situation states is given by

\[
\text{State} = \{[s] \mid s \in \text{Sit}\}
\]

where \([s]\) is the \( \approx \)-equivalence class of \( s \). For \( d \in \text{State} \), we say that a situation \( s \) is in state \( d \) iff \( s \in d \).

We will now extend the notion of 'situation state' to give a broader classification of situations into situation types. A situation type is the union of some set of states. To say that a situation is of a certain type \( d \) is to say that it is in one of the states \( d \subseteq d \).

In general, a situation type is a coarser classifier of situations than a state, although, of course, states are themselves situation types.

Definition 6.22 The collection \( \text{Type} \) of (unary) situation types is given by

\[
\text{Type} = \{U X \mid X \subseteq \text{State}\}
\]

For each \( d \in \text{Type} \) we say that a situation, \( s \), is of type \( d \) iff \( s \in d \).

This classification justifies our use of the term 'indistinguishable' for situations which are \( \approx \)-equivalent, since it is a trivial consequence of the above that for situations \( s_1 \) and \( s_2 \), \( s_1 \approx s_2 \) iff there no \( d \in \text{Type} \) such that \( s_1 \) is of type \( d \) but \( s_2 \) is not.

In the classification of situations into types, we have the beginnings of an account of information content. To know that a situation is of type \( d \in \text{Type} \) is to know something.

If \( d \) is a state then knowing that a situation is of type \( d \) is knowing all we can know about its internal structure, but other aspects of the situation can be revealed through
knowledge of its external properties. A generalization of the notion of a situation type is obtained by considering relational properties of situations which are dependent only on the relative locations of the participating situations (in addition to their internal structure).

**Definition 6.23** The collection \( \text{Type}^r \) of \( n \)-ary relational situation types consists of those \( d \subseteq \text{Sit}^n \) such that for all situation invariants, \( e \),

\[
\text{if } (s_1, \ldots, s_n) \in d \text{ then } (es_1, \ldots, es_n) \in d
\]

It is necessary but simple to check that \( \text{Type}^1 = \text{Type} \).

Relational situation types capture the simplest external properties of situations, but are still very restrictive. There are many sets of situations of interest which are not situation types; for example, the set, \( \{ s \in \text{Sit} \mid \text{time} = \tau \} \), of situations with time period \( r \).

The failure of situation types to characterize sets of situations like this is due to their lack of a 'point of view'. That a pair of situations is related by a type is a disembodied fact: any situation invariant will map the situations to another pair of related situations on the other side of world. It is instructive to note that we can capture the notion of time period as a higher level uniformity across situations since the relation,

\[
\{(s_1, s_2) \mid \text{time}_{s_1} = \text{time}_{s_2}\}
\]

of having the same time period is a binary situation type. This relation differs importantly from the property of having a particular time period \( \tau \) since it is independent of any 'point of view'. If we are to attain the generality of Rosen- schein's 'world conditions' we must go beyond situation types, by re-introducing points of view.

**Definition 6.24** A situation condition is a binary relation between situations. Let \( \Sigma \) be the set of situation conditions. We say that a situation \( s_1 \) satisfies situation condition \( \sigma \in \Sigma \) from the point of view of \( s_0 \) iff \( (s_1, s_0) \in \sigma \).

Pursuing the metaphor of a 'point of view' we arrive at a translation of Rosenschein's 'world conditions', \( \Phi \), into situation conditions. A world \( w \) satisfies world condition \( \phi \in \Phi \) from the point of view of time period \( \tau \) iff \( (w, \tau) \in \phi \). To make this correspondence
more precise, we define a translation, \( \phi \in \Sigma \), of each \( \phi \in \Phi \), viz.

\[
\phi = \{ (\mathbf{w}^\alpha, s) \mid (\mathbf{w}, \mathrm{time}) \in \phi \}
\]

The situation condition which we associate with an aggregate of dynamic locations \( \alpha \) being in a temporal state \( d \) is given by

\[
S^\mathrm{sim}_d(d) = \{ (s_1, s_2) \mid \delta(s_1, s_2, s_1) \in d \}
\]

This is the condition which is satisfied from the point of view of \( s_2 \) by those situations \( s_1 \) in which \( \alpha \) is in state \( d \) now, i.e. over the time period of \( s_2 \). The contribution of the 'point of view' situation is merely that of fixing the time period. This is directly analogous to \( M_\alpha(d) \); in fact it is easily seen that

\[
M_\alpha(d) \subseteq S^\mathrm{sim}_d(d)
\]

or, more precisely, that

\[
M_\alpha(d) = \{ (\mathbf{w}^\alpha, s) \mid (\mathbf{w}^\alpha, s) \in S^\mathrm{sim}_d(d) \text{ and } \mathbf{w} \in W \}
\]

The notion of a situation condition is more general than that of a world condition since, as well as the time, other aspects of the 'point of view' situation may be relevant to the condition. For example we can express the condition, \( S^\mathrm{sim}_d(d) \), which is satisfied from the point of view of \( s_2 \) by those situations \( s_1 \) in which \( \alpha \) is in state \( d \) here, i.e. in the aggregate location occupied by \( s_2 \). Other examples of aspects of the 'point of view' situation which may be called on include line of sight, relative speed, space occupied, etc.

Implication between two world conditions is given simply by set inclusion, but implication between situation conditions is slightly more involved. If one situation condition is contained in another then it will imply the later condition: i.e. from any given point of view, if a situation satisfies the first condition then it will also satisfy the second. But this notion of implication alone is too weak. It does not take into account the global notion of possibility implicit in the world system or anything that corresponds to it. Whereas world conditions are confined to applying to possible worlds (\( \Phi = \text{pow}(W \times T) \))
situation conditions apply indiscriminately to all situations, be they possible or impossible.

One way of rectifying this problem would be to restrict situation conditions to possible situations. But here we wish to consider an alternative to the global notion of possibility on which Rosenschein's account, and others like it, rely. We turn instead to the idea of an 'information channel'.

### 6.3 Information Channels

We introduce the notion of a channel to capture informational dependencies between situations. A channel is something which connects situations in a way which expresses their relative conditions of occurrence. Channels can be active, or not, possibly depending on the environment of the situations they connect. If an active channel links situation $s$ to situation $s'$ then if $s$ occurs, i.e. if the location of $s$ actually has the state distribution given by $s$, then $s'$ also occurs.

For example, consider the simple communication system of Figure 6.3. The system consists of two devices: a 'source' and a 'receiver'. Each device can be in one of six mutually exclusive states: \{a, b, c, d, e, f\} and \{1, 2, 3, 4, 5, 6\} respectively. Signals are transmitted from the source to the receiver, whose state is dependent on the information it receives. In particular, when the source is in state $a$ it sends a signal to the receiver
which reacts immediately by going into state 1. This dependence is captured using channels in the following way. We say that there is a channel which links the situation of the receiver's being in state \( a \) at a particular time, \( t \), with the situation of the source's being in state \( a \) at the same time. If the receiver actually is in state 1 at time \( t \) then, as long as the channel is active, the source will be in state \( a \) at time \( t \). The activity of the channel depends on many factors, such as the proper transmission of signals and the robustness of the receiving device.

One might think that the co-occurrence of any two situations can be 'explained' by the existence of an active channel between them, but this is not so. We require that channels are only sensitive to genuine physical differences. If a channel links the situation \( s \) to the situation \( s' \) and \( s_1 \) is physically indistinguishable from \( s \) then \( s_1 \) must also be linked by the channel to a situation \( s'_1 \) which is indistinguishable from \( s' \). Consequently, if \( s_1 \) is not linked to \( s'_1 \) then \( s \) cannot be linked to \( s' \) either.

In our example, this means that if in a new situation (at time \( t' \), say) the receiver is in a state physically indistinguishable from the state it was in at time \( t \) then this situation must be linked by the channel to a situation in which the source is in a state physically indistinguishable from the state it was in at time \( t \). Moreover, the relation between this situation and the situation of the source at time \( t \) must be 'parallel' to the relation between the situation of the receiver at times \( t' \) and \( t \): in other words, the channel must link the situation of the receiver at time \( t' \) to the situation of the source at time \( t' \).

Before we go any further a potential misconception should be dismissed. It is tempting to identify information channels with the physical objects which enable them to work: the wires and radar links of modern electronic communication systems. We avoid this indentification because, amongst other reasons, it is in general very difficult to locate a physical object (a wire, say) which is solely responsible for the flow of information from one situation to another, and because, even when we can identify such an object, there are usually complex aspects of the environment which play a supporting role in ensuring that the 'wire' performs as it should.

We model information channels as partial functions from situations to situations. The sense in which they respect the 'parallel' indistinguishability criterion is made precise by our notion of a situation invariant. In our example, the situation invariant which
connects the source's two situations, in virtue of which they are indistinguishable, must be the same as the invariant connecting the receiver's two situations. In this case, the situation invariant is induced by a simple translation from $t$ to $t'$ along the time line.

**Definition 6.25** $c : \text{Sit} \rightarrow \text{Sit}$ is an information channel iff for all situations $s \in \text{dom}c$ and situation invariants $e : \text{Sit} \rightarrow \text{Sit}$ with $e(s) \in \text{dom}c$ and $ecs = ces$. In other words, channels commute with situation invariants. Let $\text{Channels}$ be the collection of all situation channels in the world system.

We can use channels to describe the information which "flows" between situations on the basis of their classification into types. To do this we mimic Situation Theory's constraints. For a channel $c$ and situation types $t_1$ and $t_2$ with $t_1 \subseteq \text{dom}c$ and $\text{ran}c \subseteq t_2$ we say that $t_1$ involves $t_2$ just in case $c$ is active. From this we can derive the usual condition for constraint satisfaction used in Situation Theory: if $s$ is a situation of type $t_1$ and $t_1$ involves $t_2$ then there is some other situation of type $t_2$. Moreover, with the dependency expressed as a channel we can be more precise, since the posited situation of type $t_2$ is determined by the channel to be the situation $cs$.

Note that for any two situations there is a channel which links them. In fact, if we have the Uniqueness Property for situation invariants, then there is a 'minimal' channel linking any two situations, $s_1$ and $s_2$, namely the channel $c : [s_1] \rightarrow [s_2]$ defined for $s \in [s_1]$ by $cs = c_{s_1} s_2$ where $c_s$ is the unique situation invariant mapping $s_1$ to $s$. This channel is minimal in the sense that its graph is contained in any other channel linking $s_1$ to $s_2$. The channel we appealed to in our discussion of the source/receiver system was an example of such a channel.

This abundance of channels may at first seem somewhat worrying, but all that should be concluded is that, potentially, there are dependencies between any two situations. Only when a channel is active will there be consequences of any substance. Clearly then, much of the work needed to make sense of an analysis of information flow in terms of channels will go in to saying what it is for a channel to be active.

As we have hinted, whether or not a channel works for a particular situation in its domain can, in general, depend on the environment in which that situation occurs. For example, the receiver's being in state 1 may be linked by a channel to the source's being
in state $a$ in virtue of a cable connecting the source to the receiver. Information only flows down the cable if certain conditions are met; the temperature must be within a reasonable range, the cable must not be fractured, there must be no strong external electromagnetic forces, etc. These are conditions which must be satisfied by the environment in which the situation of the receiver's being in state $1$ occurs.

A first attempt at an analysis of channel activity may proceed by looking at the behaviour of a channel in the actual world. We may require of active channels that they work ubiquitously in the actual world. For this to be the case the channel must map each actual situation in its domain to an actual situation. In other words, if $s \in \text{dom}C$ and $s \leq s_0$ then $cs \leq s_0$. For a general account channel activity, this is both too restrictive and too lax. It is too restrictive because it does not account for channels that work in some actual circumstances but not in others, and it is almost certain that any channel of interest will fail sometimes. It is too lax in that it counts a channel as active even if it links situations which only co-occur by accident: if every actual situation of a certain type is linked to an actual situation in a regular way (a 'parallel' way) then the channel will be deemed to be active even if this correlation occurs 'by chance'.

Admittedly, for the kinds of channels we have been considering, this possibility is somewhat remote. If the cable between source and receiver was broken, but the state of the receiver continued to be correlated with the state of the source, then the cable is unlikely to have been responsible for the flow of information in the first place. This is further support for our wariness of identifying channels with physical objects.

There is another, potentially more worrying, reason why the account of channel activity is too lax. If $d$ is a type of situation which does not occur in the actual world (i.e. there are no actual situations of type $d$) then any channel with domain $d$ will be deemed active.

The traditional approach to these kind of problems (which arise in attempting to analyse law-like dependencies of any kind) is to appeal to a global notion of possibility. In this case, a putative solution would be to count as active those channels which only link situations $s_1$ and $s_2$ which co-occur in all possible worlds, i.e. for which $s_1^W \subseteq s_0^W$.

**Definition 6.26** For world system, $W$, let $Ch_w^C$ be the set of channels $c$ such that for every $s \in \text{dom}C$, $s^W \subseteq cs^W$. 

Taking \( \text{Ch}_{Wv} \) to be the active channels solves part of the problem of laxity. 'Accidental' correlations cease to be active, as do channels from non-occurring types, but the latter problem re-emerges since all channels from impossible types will still be active. Also the account is still overly restrictive. Only channels which are genuinely ubiquitous are active.

Our approach to the problem is to reverse the direction of explanation. Instead of trying to explain the law-likeness of information flow along channels in terms of a global notion of possibility, we will reconstruct possibilities from the notion of an active channel. Instead of supposing that certain worlds are, as a matter of fact, possible whilst others are not, will will suppose that, in certain situations, certain channels are active whilst in other situations they are not.

To demonstrate the coherency of this proposal we will show how, given just the set \( \text{Sit} \) of situations, the set of situation invariants and taking the set of channels \( \text{Ch}_{Wv} \) to be ubiquitously active, we can recover all of the theory presented in the previous section. Moreover, the use of channels will enable us to state the relation of implication between situation conditions in a way that is conservative with respect to implication between world conditions.

The implications between situation conditions are dependent on which channels are active in the following way. Situation condition \( \sigma_1 \) implies situation condition \( \sigma_2 \) if, from every point of view, if \( \sigma_1 \) satisfies \( \sigma_1 \) then there is an active channel linking \( \sigma_1 \) to a situation \( \sigma_2 \) satisfying \( \sigma_1 \) from the same point of view. To capture this dependence, we relativize implication between situation conditions to a set of channels. By doing this we mean to determine those implications that would exist between situation conditions on the assumption that the relativizing set of channels are all active.

**Definition 6.27** Let \( C \subseteq \text{Channels} \). For \( \sigma_1, \sigma_2 \in \Sigma \), we say that \( \sigma_1 \) \( C \)-implies \( \sigma_2 \) iff for all \( \sigma_0 \in \text{Sit} \), if \( \langle \sigma, \sigma_0 \rangle \in \sigma_1 \) then there is some \( c \in C \) for which \( \langle cs, \sigma_0 \rangle \in \sigma_2 \).

Now, supposing that the situations \( S \subseteq \text{Sit} \) all occur and the set \( C \) of channels are ubiquitously active, we know at least that the situations \( cs \), for each \( s \in S \cap \text{some} \) also occur. But do they occur in \( S' \)? We say that \( S \) is \( C \)-closed if they all do. More precisely,
CHAPTER 6. WORLDS

Definition 6.28 Given a collection $S$ of situations and a channel $c$, we say that $S$ is closed under $c$ if for all $s \in S \setminus \dom c$, $cs \in S$. If $C \subseteq \text{Channels}$, we say that $S$ is $C$-closed just in case $S$ is closed under each $c \in C$.

Proposition 6.29 If $s$ is $C$-closed and $s \approx s'$ then $s'$ is $C$-closed.

Proof: Suppose $s$ is $C$-closed and $s \approx s'$. For $c \in C$, we must show that $s'$ is closed under $c$. Since $s \approx s'$ there is a situation invariant, $e$, such that $es = s'$. But $e^{-1}$ is also a situation invariant, so by Proposition 6.19, for each $s'' \preceq es$, $e^{-1} s'' \preceq e^{-1} es = s$. Now if $s'' \in \dom c$ then, since $c$ is a channel and $e^{-1}$ is a situation invariant, $e^{-1} s'' \in \dom c$ and $e^{-1} s'' = e^{-1} cs''$. But $s$ is closed under $c$, so $e^{-1} cs'' \not\subseteq s$, and, by Proposition 6.19 again, $cs'' = ee^{-1} cs'' \subseteq es = s'$. So $s'$ is also closed under $c$.

We extend the notion of $C$-closure to single situations. A situation is $C$-closed just in case the set of all its parts, $s_k = \{s' | s' \preceq s\}$, is closed. If a situation is $C$-closed then it is informationally isolated with respect to the channels $C$. We now show that the notion of 'possible world' can be recaptured as closure under $Chw$.

Proposition 6.30

1. Every $Chw$-closed situation is possible.
2. Every possible world is $Chw$-closed.
3. A world is possible if and only if it is $Chw$-closed.

Proof:

1. Suppose that $s$ is a $Chw$-closed situation and, for contradiction, that $s$ is not possible, i.e $s^W = \emptyset$. Let $E$ be a set of situation invariant such that $[s] = \{es | e \in E\}$ and for $e, e' \in E$, $es \not\equiv es'$. Let $c : [s] \rightarrow [uq]$ be the channel linking $s$ to $uq$, such that for each $e \in E$, $ces = eq$. This is a well-defined channel (and, given the Uniqueness Property, is independent of $E$).

We show that $c \in Chw$. For each $s' \in [s]$, $s \approx s'$ so, by Proposition 6.19, $s'$ is not possible, i.e $s'^W = \emptyset$. Now $s' = es$ for some $e \in E$, so $ces = ces = eq$. But
CHAPTER 6. WORLDS

\[ \text{ew}_0 \approx w_0, \text{ so by Proposition 6.19, } \text{ew}_0 \text{ is possible, so } cs^{\text{ew}_0} = \text{ew}_0^W \neq \emptyset. \text{ Hence } s^{\text{ew}_0} \subseteq cs^{\text{ew}_0} \text{ and so } c \in \text{Chw}. \]

Now since we have supposed \( s \) to be \( \text{Chw} \)-closed, \( cs \not\subseteq s \). But \( cs = w_0 \text{ so } w_0 \not\subseteq s \). Since \( w_0 \) is a world it follows that \( s \not\subseteq w_0 \), so \( w_0 \in s^W \). Hence \( s \) is possible.

2. Suppose that \( w \) is a possible world and \( c \in \text{Chw} \). Now if \( s \not\subseteq w \) and \( s \in \text{domc} \) then \( s^W \subseteq cs^W \). But, by Proposition 6.17, \( w^* \in s^W \text{, so } w^* \in (cs)^W \), and so \( cs \not\subseteq w \).

Hence \( w \) is \( \text{Chw} \)-closed.

3. follows from 1. and 2.

Next we show that \( \text{Chw} \)-implication between situation channels is conservative with respect to implication between world conditions. We make use of the translation of world condition \( \phi \) into situation condition \( \hat{\phi} \) that was defined at the end of Section 6.2.

**Proposition 6.31** For world conditions \( \phi_1 \) and \( \phi_2 \), \( \phi_1 \) implies \( \phi_2 \) iff \( \hat{\phi}_1 \text{Chw-implies } \hat{\phi}_2 \).

**Proof:** Suppose \( \phi_1 \) implies \( \phi_2 \). Then, by definition, \( \phi_1 \subseteq \phi_2 \), so \( \hat{\phi}_1 \subseteq \hat{\phi}_2 \). So, for all \( s_0 \in \text{Sit}, (s, s_0) \in \phi_1 \text{ then } (s, s_0) \in \phi_2 \). But the identity map on \( [s] \) is clearly in \( \text{Chw} \), so \( \hat{\phi}_1 \text{ Chw-implies } \hat{\phi}_2 \).

Conversely, suppose \( \hat{\phi}_1 \text{ Chw-implies } \hat{\phi}_2 \). Then for all \( s_0 \in \text{Sit} \), if \( (w, s) \in \phi_1 \) then there is a \( c \in \text{Chw} \) such that \( (cw, s) \in \phi_2 \). But \( \phi_1 \) is a world condition, so \( w \) is a possible world and so is \( \text{Chw} \)-closed, by Proposition 6.30. So \( cw \not\subseteq w \). But \( \phi_2 \) is also a world condition, so \( cw \) is also a possible world and so \( w \not\subseteq cw \), by the definition of 'world'. Moreover, \( w = w^* \text{ for some } w_1 \in W, \text{ so } w^* = w^{cw}_* = w^* = w \), so \( w \) is static.

Similarly, \( cw \) is static, so, by Proposition 6.14, \( w = cw \). Hence \( \phi_1 \subseteq \phi_2 \), so \( \phi_1 \subseteq \phi_2 \), and so \( \phi_1 \) implies \( \phi_2 \).

The above two propositions show that Rosenschein's model of information, which appeals to the notion of 'possible world' can be replaced by one which uses situations...
and information channels. Admittedly, the definition of \( Ch_W \) appeals the set \( W \) of possible worlds, but \( Ch_W \) is just a set of channels. The question arises as to whether a set of possible worlds can be induced by an arbitrary set of channels. Motivated by Proposition 6.30, we define, for \( C \subseteq \text{Channels} \),

\[
\text{Worlds}(C) = \{ w \mid w \text{ is a world and } w \text{ is } C\text{-closed} \}
\]

Now for the set of channels, \( Ch_W \), we have the property that

\[
Ch \text{Worlds}(Ch_W) = Ch_W
\]

i.e. the set of channels deemed active with respect to \( \text{Worlds}(Ch_W) \) is just the set \( Ch_W \).

In other words, if, as a matter of fact, the channels, \( Ch_W \), were the active ones then, taking \( \text{Worlds}(C) \) to be our set of possible worlds, the possible world account of channel activity would be correct. In general, then, we can say that a possible worlds account of channel activity is forthcoming iff the set of all active channels, \( C \), has the property

\[
Ch \text{Worlds}(C) = C
\]

For each \( C \subseteq \text{Channels} \), let \( C^\circ = Ch \text{Worlds}(C) \). This defines a closure operator, i.e. \( C \subseteq C^\circ, C^{\circ\circ} \subseteq C^\circ \text{ and } C_1 \subseteq C_2 \text{ then } C_1^{\circ\circ} \subseteq C_2^{\circ\circ} \). Whether or not we can provide a possible worlds account of channel activity depends on whether or not the set of active channels, \( C \), is closed with respect to this operator, i.e \( C^\circ = C \). We should examine some of the consequences of a set of channels being closed. The proofs of the following properties of closed sets of channels are omitted owing to their monotonity and messiness.

The importance is rather in the picture of closed sets of channels that emerges.

**Proposition 6.32** Let \( C \) be a closed set of channels, i.e. \( C^\circ = C \).

1. For each \( d \in \text{Type} \), the identity function on \( d \) is in \( C^\circ \).
2. For situations \( s, s' \in \text{Sit} \) such that \( s' \preceq s \) the channel with domain \( [s] \) taking \( es \) to \( es' \) for each situation invariant, \( e \), is in \( C^\circ \). (Assumes Uniqueness Property).
3. If \( c, c' \in C \) and \( \text{dom} c' \subseteq \text{ran} c \) then \( c \circ c' \in C^\circ \).
4. If \( c, c' \in C \) then the channel with domain \( \text{dom} c \cap \text{dom} c' \) taking \( s \) to \( cs \cup c's \) (which is defined) is in \( C^\circ \).
5. If \( X \subseteq C \) then the channel with domain \( \bigcap X \) taking \( s \) to \( \bigcup_{\epsilon \in X} \epsilon s \) (which is defined) is in \( C^a \).

For an arbitrary closed set \( C \) of channels, we can reconstruct the notion of a possibility as 'C-coherence', which we now define.

**Definition 6.33** Given a collection \( S \) of situations and a set of channels \( C \subseteq \text{Channels} \), we say that \( S \) is \( C \)-coherent just in case there is a \( C \)-closed situation, \( s \in \text{Sit} \) such that \( S \subseteq s \). A situation \( s \in \text{Sit} \) is \( C \)-coherent iff \( s \) is.

**Proposition 6.34** If \( C = C^a \) then a situation is \( C \)-coherent iff it is possible with respect to \( \text{Worlds}(C) \).

**Proof:** If \( s \) is \( C \)-coherent then there is a \( C \)-closed situation, \( s' \), such that \( s \subseteq s' \), i.e. such that \( s \not\subseteq s' \). Now \( C = C^a = \text{Ch}_{\text{Worlds}(C)} \), so by Proposition 6.30.1, \( s' \) is possible with respect to \( \text{Worlds}(C) \). Hence \( s \) is also possible.

Conversely, if \( s \) is possible with respect to \( \text{Worlds}(C) \) then \( \text{Worlds}(C) \neq \emptyset \), i.e. there is a world \( w \in \text{Worlds}(C) \) such that \( s \not\subseteq w \). By Proposition 6.30.2, \( w \) is \( \text{Ch}_{\text{Worlds}(C)} \)-closed. But \( \text{Ch}_{\text{Worlds}(C)} = C^a = C \), so \( w \) is \( C \)-closed, and so \( s \) is \( C \)-coherent.

Finally, the notion of \( C \)-coherence enables us to define a negative implication, or 'preclusion', between situation conditions.

**Definition 6.35** Let \( C \subseteq \text{Channels} \). For \( \sigma_1, \sigma_2 \in \Sigma \), we say that \( \sigma_1 \) \( C \)-precludes \( \sigma_2 \) iff for all \( s_0 \in \text{Sit} \), if \( (s, s_0) \in \sigma_1 \) and \( (s, s_0) \in \sigma_2 \) then \( s \) is not \( C \)-coherent.

In this section we have proposed that informational dependencies between situations should be modelled directly, by channels, rather than indirectly in terms of a set of possible worlds. Relative to a set \( C \) of channels, which we suppose to be ubiquitously active, we defined notions of \( C \)-closure, \( C \)-coherence, \( C \)-implication and \( C \)-preclusion.

\( C \)-implication between situation conditions proved to be an appropriate generalization of implication between world conditions, whilst \( C \)-preclusion was new to this setting.

We suggested that situations which are \( C \)-coherent can be regarded as 'possible' and
CHAPTER 6. WORLDS

situations which are $C$-coherent can be regarded as 'possible worlds'. This correspondence with the terminology of Section 6.2 was shown to be exact when $C = Chw$ and, more generally, for any $C$ such that $C^{o} = C$.

We can conclude, on the assumption of ubiquity, that if the set $C$ of active channels is such that $C^{o} = C$ then it is amenable to analysis by the set $Worlds(C)$ of possible worlds. However, if $C^{o} \neq C$ then no such analysis is possible. In the next section we set out to question the assumption of ubiquity: can we analyse channel activity by supposing that there is a fixed set of active channels, or is it that channels are active on some occasions and not on others?

6.4 Channel Conditions and Conditional Channels

Suppose that we are concerned with a particular communication system, a telephone network, say, and that there is a set $C$ of channels which are active in the network. These channels link situations in which sounds are produced by the receiver of a telephone, $A$, with situations in which sounds are made in the vicinity of the mouthpiece of another telephone; the telephone $B$ to which $A$ is connected by the network at that time.

What are the consequences of the channels $C$ being active? For each channel $c \in C$ we can be sure that for all $w \in Worlds(C)$ and $s \in dom(c)$, if $s \not\subseteq w$ then $cs \not\subseteq w$. In particular, if $w_0 \in Worlds(C)$ and $s$ is actual then $cs$ is actual. But, as we have been suggesting, this is too restrictive. For example, under normal circumstances, if I am speaking to you on the telephone and I hear nothing through my receiver then you are not saying anything; i.e. the silent situation in the vicinity of your telephone is actual. But if there is some technical problem with the telephone wires connecting us, then this may no longer be the case. There are certain conditions which must be satisfied for the channel to transmit information.

The set of worlds, $Worlds(C)$, contains those worlds in which the telephone network always works perfectly. For each of these worlds it must at least be the case that the telephone network exists. Worlds in which all telephones were fakes, or executive toys with no function, would be excluded from $Worlds(C)$. But these worlds are clearly not impossible in any physical sense, so that $Worlds(C) \neq W$. 
CHAPTER 6. WORLDS

More seriously, in all worlds in $Worlds(C)$ the telephone network must remain in working order for all time. This strongly suggests that the actual world is not in $Worlds(C)$. Consequently, if situations $s_0$ and $s_1$ are linked by an active channel $c$ and $s_0$ is actual, we cannot decide whether or not $s_1$ is actual, since, in all probability, the actual world $s_0$ will not be closed under $c$.

The problem is that worlds are too big to capture the conditions under which information flows, so let us consider the set of C-closed situations, rather than just the set of C-closed worlds. We define,

$$Site(C) = \{ s \mid s \in Sit and s \text{ is C-closed} \}.$$ 

Now if $s \in Site(C)$ occurs and $s' \not\subseteq s$ then for appropriate $c \in C$, $cs' \not\subseteq s$ and so $cs'$ also occurs. In other words, inside a C-closed situation information is guaranteed to flow from $cs$ to $s$. The advantage of considering $Site(C)$ over $Worlds(C)$ is that there may be actual C-closed situations even though the actual world is not itself C-closed.

From an epistemological point of view we are no better off. If we know that $s$ occurs then we can only be sure that $cs$ occurs if we know that some $s' \in Site(C)$ occurs for which $s \not\subseteq s'$. But if we knew that $s'$ occurs then we would know that $cs$ occurs anyway, since $cs \not\subseteq s'$. Nevertheless, we need not be dismayed by this conclusion. As Dretske, and others, have argued at length, it is not necessary that we have such meta-knowledge about our environment. It need only be the case that a channel is working when we rely on it to gain knowledge. When we use a telephone, we are almost always ignorant of the state of the telephone network we rely on\(^{1}\), but we still succeed in gaining information about what is said or unsaid at the other end of the line. The success of our acquisition of information depends not on the guarantee of certainty, but on our behaviour. If we consistently confused real telephones with toy telephones, we would not gain information very effectively. But we don’t, so we do.

Nevertheless, an account of the role of channels in supporting information flow would be incomplete without an analysis of the conditions under which channels work. The need

\(^{1}\)In practice, we are not usually kept in ignorance for long if a telephone line breaks down. Apart from our knowledge of human behaviour – you are unlikely to remain silent for too long – the telephone would probably make an unpleasant sound indicating to me that something is wrong. But this is a fact which helps us recover from epistemic mistakes, not a precondition for the possibility of knowledge.
for such an analysis is most acute in the explanation of implication between situation conditions. We have relativized implication to a set of channels $C$ such that $\sigma C$-implies $\sigma'$ just in case situations satisfying $\sigma$ from some point of view are linked by channels in $C$ to situation satisfying $\sigma'$ from the same point of view. As above, the passage from $\sigma$ satisfaction by an actual situation to $\sigma'$ satisfaction is dependent on certain conditions. We capture this dependence in the following definition.

**Definition 6.36** Let $C \subseteq \text{Channels}$ and $d \in \text{Type}$. For situation conditions $\sigma_1$ and $\sigma_2$, we say that $\sigma_1 C$-implies $\sigma_2$ given $d$ if for all $s_0, s \in \text{Sit}$, if $(s, s_0) \in \sigma_1$ and there is a $s' \in d$ such that $s \leq s'$ then there is a $c \in C$ such that $(c, s_0) \in \sigma_2$. $d$ is called a background type.

The inclusion of a background type in statements of implication between situation conditions expresses at least some of the conditionality implicit in the channels on which the implication relies. Of course, we will only fully capture the background conditions in implications of the form $\sigma_1 C$-implies $\sigma_2$ given $\text{Sit}(C)$. These implications are covered by our definition since

**Proposition 6.37** $\text{Sit}(C)$ is a situation type.

**Proof:** If $s \in \text{Sit}(C)$ and $s' = s$ then, by Proposition 6.29, $s' \in \text{Sit}(C)$ and so $[s] \subseteq \text{Sit}(C)$. So $\text{Sit}(C) = \bigcup_{s \in \text{Sit}(C)} [s]$, and so $\text{Sit}(C)$ is a situation type.

We now turn to a potentially more serious problems with the notion of an active channel. We have seen how the $C$-closure of a containing situation ensures the proper functioning of $C$ channels, and this enables us to capture conditions under which a channel is active. But do these conditions adequately capture the conditionality of real informational dependencies? We will look at an example.

A channel makes an informational link between situations in a way that is dependent on their relative locations in time and space. If my telephone is connected to yours by the telephone network then, given that the network is working properly, situations in which the sound of a cough is emitted by my receiver are linked by an active channel to situations in which you cough (at more or less the same time). Moreover, if my
receiver emits a second coughing sound later in our conversation, this new situation will be linked to another situation in which you cough (at the later time). In this case the relative locations of the cough sounds to the coughs remain constant: cough sounds are produced by my telephone and coughs are coughed at your telephone at the same time. The definition of a channel requires that this relationship be maintained.

However, it may be that the the informational connection between situations depends on more than their relative locations in time and space. For example, suppose we are speaking on the telephone. I am in my office and you are traveling down the M6. Your car telephone is therefore in motion relative to my office phone. Suppose I hear a cough in my receiver at 1.00pm whilst you are traveling past Carlisle. The situation of my receiver emitting the coughing sound is presumably linked by a channel to the situation of your coughing in your car at more or less the same time. Moreover, we would suspect that this channel is active: there is a genuine connection between the two situations.

Yet there is a problem with the way we have defined channels so far. If the situation in my office, \(s_0\), is to be linked by \(c\) to the situation \(s_1\) in your car, then it must be the case that every situation which is physically indistinguishable from \(s_0\) is linked by \(c\) to a situation 'parallel' to and indistinguishable from \(s_1\). But consider the situation invariant \(e\) which maps situations three hours forward in time. If you cough whilst passing Manchester at 4.00pm, three hours after coughing in Carlisle, then I will hear a coughing sound in my receiver at more or less the same time. So the situation \(e s_0\) occurs. But \(c\) is required to link the situation \(e s_0\) at my receiver to the situation \(e s_1\). This is quite wrong since \(e s_1\) is the situation of your coughing in Carlisle and neither you nor your car phone are within a hundred miles of Carlisle at 4.00pm.

The problem resides in the fact that the relative locations of situations linked by a channel may not be the factor which determines where information flows. In the above example, it should be the relative locations of our telephones which determine which situations are linked. The situation that \(s_0\) is linked to should be dependent on where your telephone is.

Our solution is to specify the location of an informational source (the situation at your telephone) relative to the background situation in which the information flow occurs (the situation of the telephone network). Indirect informational connections of this
kind are captured by 'conditional channels':

Definition 6.38 A conditional channel is a partial function $c : \text{Sit} \times \text{Sit} \rightarrow \text{Sit}$ such that for each situation invariant $e$ and $(s_1, s_2) \in \text{dom} c$, $c(es_1, es_2) = ec(s_1, s_2)$. Let $C\text{Channels}$ be the set of conditional channels.

The informational connection between my office phone and your car phone is expressed by a conditional channel, $c$, which maps pairs $(s_1, s_2)$, where $s_1$ is the situation at my telephone and $s_2$ is the network situation (which specifies the location of your telephone), to the situation $c(s_1, s_2)$ at your telephone.

Note that in a particular 'environment' $s$, a conditional channel $c$ determines an unconditional channel $c'$ with domain $\{s' | (s', s) \in \text{dom} c\}$ and such that for each $s' \in \text{dom} c'$, $c'(s') = c(s, s')$. Also, the set of unconditional channels are naturally embedded in the set of conditional channels by allowing the second (environment) argument to range freely over $\text{Sit}$. It is now easy to extend the definition of $C$-implication and $C$-closure to conditional channels by quantifying over the second argument:

Definition 6.39 Let $C \subseteq C\text{Channels}$. For $s_1, s_2 \in \Sigma$, we say that $s_1$ $C$-implies $s_2$ iff for all $s_0, s_1 \in \text{Sit}$, if $(s_1, s_0) \in s_1$ then there is some $c \in C$ for which $(c(s_1, s), s_0) \in s_2$ for all $s \in \text{Sit}$.

Definition 6.40 Given a collection $S$ of situations and a conditional channel $c$, we say that $S$ is closed under $c$ iff for all $s, s' \in S$, if $(s, s') \in \text{dom} c$ then $c(s, s') \in S$.

We extend the definition to $C$-closure, $C$-coherence and $C$-preclusion for a set $C$ of conditional channels in exactly the same way as we did in the unconditional case.

Note that the conditionality of conditional channels is very different from the conditionality of conditional constraints. All channels are dependent on implicit 'channel conditions' for their activity. So all channels induce conditional, rather than unconditional constraints. Conditional channels, even when active, depend on a 'background' situation to determine which situations they link.

6.5 Perspectives on the World

As an application of the theory developed in the last two sections, we will show how to model perspectives in a world of situations and channels. A perspective, as you will
CHAPTER 6. WORLDS

recall, is a part of reality seen from some point of view. Within the current model, a part of reality is a (physical) situation, \( s \). Given a point of view, \( a_0 \) (another situation), and a situation condition, \( \sigma \), we can say which parts of \( s \) satisfy \( \sigma \) from the point of view of \( a_0 \); i.e. those situations \( s_1 \in s \) such that \( \langle s_1, a_0 \rangle \in \sigma \). This provides the basis for a classification of the parts of \( s \) by the conditions they satisfy from the point of view of \( a_0 \).

For example, take my office to be the situation, \( a \); the part of reality being classified. The white-board on the wall defines a situation, \( s' \), consisting of the locations it occupies (over time) and their physical state. The white-board situation, \( s' \), is a part of my office situation, \( s \). Now \( s' \) has certain geometrical properties. It has a well defined shape which is constant over time. We might say that its internal shape is square, meaning that in its primary plane (which in this case can be thought of as the wall) the locations it occupies at any time fill a square region of that plane.

We wish to model the (visual) perspective of my office from my (visual) point of view. My visual point of view, we will suppose, is given by the situation of my eye\(^4\), \( a_0 \), which is also a part of \( s \). My eye, \( a_0 \) has various internal properties, one of which is the inclination of the plane of my retina\(^5\). That is, \( a_0 \), is related by a certain relation, \( d \), to a situation \( a_0' \); its retinal plane. The relation, \( d \), need not be restricted to eyes; an appropriately shape region of empty space can determine a plane in a similar fashion. However, \( d \) is a binary situation type, since physically indistinguishable eyes will have physically indistinguishable retinal planes.

One condition that the white-board satisfies from my point of view is that its edges look parallel. We can make this more precise. We say that situations \( s' \) and \( s'' \) satisfy the condition \( \sigma_p \) iff \( s' \) has a well defined polygonal internal shape and opposite edges of the geometric projection of \( s' \) onto the retinal plane of \( s'' \) (the unique \( s'' \) such that \( \langle s', a_0 \rangle \in d \)) are parallel. In the current example, the white-board, \( s' \), does indeed satisfy \( \sigma_p \) from my point of view, \( a_0 \). Strictly, we should refine the definition of \( \sigma_p \) further to account for changes in the position of either my eye or the white-board, but

\(^4\)we will ignore the problems of binocular vision.
\(^5\)although my retina is in fact curved, we can take its plane to be defined in a suitable way, so that the line from the centre of my pupil to the mid-point of my retina is orthogonal to it.
for the moment we will ignore the temporal dimension supposing $s'$ to be a temporal slice of the white-board.

Following this approach, we could arrive at a classification of parts of my office according as to whether or not they satisfy $\sigma_p$ from my point of view. Some parts of the room will fail to satisfy this condition because they fail to have a well-defined polygonal internal shape. Others will fail because they have opposite edges which are not parallel, or because their primary plane is not parallel to the retinal plane of my eye. As I move around the room, the relative inclination of my retinal plane to the primary planes of various parts of the room will change and so the $\sigma_p$ classification of them will also change. This is the phenomenon of visual perspective.

However, there is something missing from the story so far. If we extended $s$ to include the whole building, rather than just my office, there will be white-boards in other offices which satisfy $\sigma_p$ from my point of view even though I cannot see them. Extending $s$ even further, across the globe, there will be white-boards in California which still possess the right geometric properties to satisfy, $\sigma_p$. We are failing to capture the important constraint which, in visual terms, amounts to the existence of a field of view. My visual perspective on the world around me is not only relative to my position in the world, but it is also constrained by what I can see.

A visual perspective exists because of light. That the geometric projection of the white-board onto my retina has a certain shape is a physical fact. It is the reflected light from surfaces in my office which carry information about their geometric properties. In the terms of the theory we have been developing, there are active information channels linking the situation in my eye to other parts of the room. But there are no such channels linking my eye to parts of other rooms in the building, or to offices in California. My field of view is defined by the range of a certain set of (visual) information channels.

To model this dependence on information channels, we suppose that there is some set $C$ of (conditional) channels active in $s$. These are the channel concerned with the informational dependencies enabled by reflected light in the room. It is not that there are no other active channels between situations in the room, but the set of channels $C$ are those which give rise to visual perspectives. Other channels would define other kinds of perspective.
Although we will not demand that $C^c = C$, we will need to ensure that $C$ is at least closed under ‘iterated composition’, i.e. for $c, c' \in C$, the channel $c; c'$ with domain

$$\{(s', s) \mid (s', s) \in \text{dom} c \text{ and } (c(s', s), s') \in \text{dom} c'\}$$

and mapping each $(s', s) \in \text{dom} c; c'$ to $c'(c(s', s), s)$ is also in $C$.

Now, for a situation, $s_1 \notin s$, to be classified at all, $s_1$ must be ‘visible’ from the point of view of $s_0$, so there must be a (conditional) channel in $C$ which links $s_0$ to $s_1$ in $s$.

**Definition 6.41** For $s_1 \not\subseteq s$, we say that the channel $c_1 \in C$ locates $s_1$ from $s_0$ in $s$ iff $(s_0, s) \in \text{dom} c_1$ and $c_1(s_0, s) = s_1$.

In the visual case we have been discussing, there would seem to be no need for a channel to locate the situation of the white-board since it is in a fixed location in the room. But in the general case, when either surfaces in the room are moving, or I am moving, we need a channel to locate the part of the world about which my eye is obtaining information. This can be seen in the example of my telephoning you whilst you are traveling down the M6. The channel must fix your location if the information I gain is to apply to you (or rather to your telephone).

To classify $s_1$ by some situation condition, $\sigma$, we also require that the channel $c_1$ which locates $s_1$ from $s_0$ in $s$ is a channel along which information about $\sigma$ satisfaction flows. That is, if the room outside $s_0$ was different from the way it is, say $s'$, then, supposing $c_1$ to be active, the situation located by $c_1$ from $s_0$ in $s'$ would still satisfy $\sigma$ from the point of view of $s_0$.

**Definition 6.42** For situation condition $\sigma$, we say that the (conditional) channel $c \in C$ carries $\sigma$ to $s_0$ iff for all $s' \in \text{Sit}$, if $(s_0, s') \in \text{dom} c$ then $(c(s_0, s'), s_0) \in \sigma$.

Put more simply, this definition says that $c$ carries $\sigma$ to $s_0$ just in case whatever is at the ‘end’ of $c$ satisfies $\sigma$ from the point of view of $s_0$.

Now we can use situation conditions carried by channels in $C$ to classify situations located by channels in $C$. A perspective structure is added by considering the informational dependencies between the situation conditions, given by $C$-implication and $C$-preclusion.
CHAPTER 6. WORLDS

Definition 6.43 Given a set $C$ of (conditional) channels which is closed under iterated composition, a set $\Sigma$ of situation conditions and $C$-coherent situations $s$ and $s_0$, we define the structure $\text{Persp}(C, \Sigma, s, s_0) = (s_0, \Sigma, \Rightarrow, \bot)$ (where $s_0$ is the set of parts of $s$) by

1. $s_1 : \sigma$ iff there is a $c_1 \in C$ which locates $s_1$ from $s_0$ in $s$ and which carries $\sigma$ to $s_0$.
2. $\sigma \Rightarrow \sigma'$ iff $\sigma$-$C$-implies $\sigma'$
3. $\sigma \perp \sigma'$ iff $\sigma$-$C$-precludes $\sigma'$

Proposition 6.44 $\text{Persp}(C, \Sigma, s, s_0)$ is a perspective.

Proof:

1. Facticity. Suppose $s_1 : \sigma$ and $\sigma \Rightarrow \sigma'$. Then there is a channel $c_1 \in C$ such that $c_1(s_0, s) = s_1$ and for which $s_1, s_0 \in \Sigma$. Also $\sigma$-$C$-implies $\sigma'$, so there is a channel $c \in C$ such that $c(s_1, s_0) \in \sigma'$. Let $s_2 = c(s_1, s)$. We will show that $s_2 : \sigma'$.

Firstly, $c_1; c \in C$ as $C$ is closed under iterated composition. So $c_1; c$ locates $s_2$ from $s_0$ in $s$, since $c_1(s_0, s) = c(c_1(s_0, s), s) = c(s_1, s) = s_2$.

Secondly, we show that $c_1; c$ carries $\sigma'$ to $s_0$, from which it then follows that $s_2 : \sigma$. We have to show that for all $s' \in \text{Sit}$, if $s_0, s' \in \text{dom} c_1$ then $(c_1; c(s_0, s'), s_0) \in \sigma'$.

But if $(s_0, s') \in \text{dom} c_1; c$ then $(s_0, s') \in \text{dom} c_1$, so $(c_1(s_0, s'), s_0) \in \sigma$ since $c_1$ carries $\sigma$ to $s_0$. Also, $(c_1(s_0, s'), s') \in \text{dom} c$ and $c$ was chosen to witness that $\sigma$-$C$-implies $\sigma'$, so $(c_1(s_0, s'), s_0) \in \sigma'$. But $(c_1; c(s_0, s'), s_0) = (c_1(s_0, s'), s')$ by definition of iterated composition and so we are done.

2. Xerox. If $\sigma \Rightarrow \sigma'$ and $\sigma' \Rightarrow \sigma''$ then there are $c, c' \in C$ such that for all $s', s'', s''' \in \text{Sit}$,

if $(s'', s') \in \sigma$ then $(c(s'', s'), s'') \in \sigma'$, and
if $(s'', s') \in \sigma'$ then $(c'(s'', s'), s'') \in \sigma''$.

But $c; c' \in C$ and for all $s', s'', s''' \in \text{Sit}$, if $(s'', s') \in \sigma$ then $(c(s'', s'), s'') \in \sigma'$, so $(c; c'(s'', s'), s'') = (c(c(s'', s'), s''), s'') \in \sigma''$. Hence, $\sigma \Rightarrow \sigma''$. 
3. Local Preclusion. Suppose $\sigma \perp \sigma'$ and $s_1 : \sigma$ and, for contradiction, $s_1 : \sigma'$. Then there are channels $c, c' \in C$ such that $c(s_0, s) = c'(s_0, s) = s_1$ and for which $(s_1, s_0) \in \sigma$ and $(s_1, s_0) \in \sigma'$. But, since $\sigma \perp \sigma'$, it follows that $s_1$ is not $C$-coherent, contradicting the fact that $s_1$ is part of the $C$-coherent situation $s$.

4. Mutual Preclusion. Straightforward, since $C$-preclusion is clearly symmetric.

Concluding Remarks

In this chapter we have drawn certain connections between Situation Theory and the model of information used by Rosenschein in specifying and designing robots. With hindsight, the essential questions which needed answering were the following:

- How is information located in the physical world?
- How can information in one place be linked to information in another place?

On Rosenschein's account the information that $\phi$ is modelled as the condition the world must satisfy (at a time) in order for $\phi$ to be true. Technically this is captured by a world condition. The first question is answered by defining the information content $M_a(d)$ of a location $a$'s being in state $d$ in a natural way. The information $\phi$ is located at $a$ just in case $a$ is in state $d$ and $M_a(d) \subseteq \phi$. Implication between $\phi_1$ and $\phi_2$ is given straightforwardly by the subset relation. This enables one to give a simple answer to the second question.

On the present account, certain modifications and generalizations have been made. Firstly, the sense in which information can be located was extended to aggregates of locations, dynamic locations, and temporal behaviour. Rosenschein's suggestion of modelling distributed information by indexed sets (or tuples) of locations was abandoned in favour of distribution functions under various notions of equivalence based on spatial and temporal invariants.

The answer to our first question was given by combining these into the definition of a situation. On our account, a situation just is the sort of thing in which information can
CHAPTER 6. WORLDS

be said to be located. Situations were ordered by the part-of relation $\subset$ and classified into types in a way commensurate with the underlying notion of physical indistinguishability. Given this ordering, the usual definitions of possible and actual situation and of a world were forthcoming.

Next, world conditions were generalized to situation conditions to suit the new setting. Rosenschein's definition of implication between world conditions was seen not to generalize in a straightforward way to implication between situation conditions. To fill the gap, the idea of an information channel (from Communication Theory) was introduced. The set $W$ of possible worlds was used to define a collection $Ch_W$ of active channels relative to which a sensible definition of $Ch_W$-implication between situation conditions could be given. $Ch_W$-implication between situation conditions was seen to be conservative with respect to implication between world conditions.

We then moved on to question whether channel activity could really be determined by the set of possible worlds. Arguing that it could not, we relativized implication and the notions of 'possible situation' and 'world' to $C$-implication, $C$-coherence and $C$-closure for an arbitrary set $C$ of channels. The necessary theoretical apparatus can thus be obtained by supposing channel activity to be a unanalysed theoretical primitive and dispensing with the notion of a possible world. Moreover, an analysis of channel activity by possible worlds was seen to be unavailable in the general case.

Nevertheless, some attempt was made to analyse the conditions under which a channel becomes active. An approach reminiscent of Situation Theory's background types, was deemed unsatisfactory and we fell back on the trick of including the activity conditions of a channel in the channel itself. This leads us to the definition of a conditional channel.

Finally, we gave a model of information flow within a situation $s$ 'seen' from a point of view $s_0$, classified by situation conditions $\Sigma$ with respect to (conditional) channels $C$. This model was shown to be a perspective in the technical sense of Chapter 4. This construction shows how perspectives can be introduced into a perspectiveless world. Starting from a Rosenschein world system with active channels $Ch_W$ and world conditions $\Phi = \text{pow} W \times T$, we can build perspectives $\text{Pers}(Ch_W, \Phi, s, s')$ for each pair of actual situations $s, s'$. This collection of perspectives on actual situations can be seen as a metaphysical alternative to the collection of possible worlds $W$ in Rosenschein's
Nevertheless we have not fully established the feasibility of this alternative. A collection of perspectives generated from a world system is in some sense consistent in virtue of being so generated, but this cannot be said of arbitrary collections of perspectives. In Chapter 5 we showed how a collection of perspectives on an ordered set of situations (a perspectival domain) can give rise to notions of object and property as uniformities within perspectives and across perspective ‘shifts’. Although providing a definition of predication in some simple cases, the account lacked a notion of consistency between perspectives to make sense of predication in general. Now we have a characterization of a class of perspectival domains which can be said to be internally consistent: those generated from Rosenschein’s world systems and, more generally, those generated from (almost) arbitrary collections $C$ and $E$, of (conditional) channels and situation conditions respectively. Much remains to be done in clarifying this connection. In particular, a place has to be found for objects and properties in the bodiless world of states and space-time.
Appendix A

Classification Systems

This appendix is intended as a supplement to the theory of classifications presented in Chapter 1. In it we provide a range of examples and applications of the theory which do not fall inside the scope of the main text of the thesis.

In Section A.1 we give examples of four systems of classification which are often used in practice: taxonomies, state systems, feature systems and attribute-value systems. Each one seems to illustrate some aspect of classification in general. Various themes such as the distinction between partial and total specification, coherency and incompatibility are picked out from these examples.

In Section A.2 we explore some of the relationships between the different systems of the previous section. We provide several constructions which show how one kind of classification system can be built out of another and vice versa. Both feature systems and state systems can sometimes be thought of in terms of attributes and values, whilst the general relationship between featural and taxonomic systems of classification is unclear.

We look at a system of classification which relates a feature system to an arbitrary classification, then explore the specific connection between featural and taxonomic classifications which would help to solve the problem of feature-based recognition. This involves us in some discussion of different principles of 'inheritance'.

Finally, in Section A.3 we develop a specific application of our approach to classification: that of classifying words according to the syntactic relationships between them. We intend this section to be read as an exercise in the application of our approach to a novel domain rather than a significant contribution to linguistic theory. Our starting point
APPENDIX A. CLASSIFICATION SYSTEMS

Figure A.1: The subtype relation in a leaf taxonomy.

is the approach to linguistics known as 'dependency grammar' in which dependency relations between words (rather than sequences of words) are taken to be the primary focus of linguistic analysis.

A.1 Homogeneous Classification Systems

In this section we give four examples of systems of classification. The list is by no means exhaustive since we intend it to be read mainly as an amplification of the ideas developed so far. Taxonomies are fairly rigid classification systems in which the extensions of types are partitioned by the extensions of their subtypes. State systems have the property that two tokens in the same state are indistinguishable. Feature systems are concerned with partial rather than complete classification of tokens: the notions of coherence and compatibility are central. Attribute-value systems are derived from attribute-value structures, which are widely used classificatory structures in Linguistics and Computer Science. We show how our systems are related to Johnson's (1988) study of these structures.

A.1.1 Taxonomies

Our first example of a classification is that of a taxonomic hierarchy of the sort that occurs in school biology. Here individuals of some domain (leaves, say) are classified into types which are ordered by the subtype relation, as in Figure A.1. Some leaves are pointy, others are rounded. The pointy ones may have prickly points, like holly leaves, or
they may not. The prickly leaves are a subtype of the pointy leaves. Of course, a more informed biological classification of leaves would be based on more robust biological invariants such as the species of the plant that the leaves came from, but the principle is still the same — only the generality of the classification is weaker.

**Definition A.1** A *taxonomy* is a structure \( \langle D, T, \prec \rangle \) where \( T \) is a \( D \)-classifier and \( \prec \) is a binary relation on \( |T| \) such that, for each \( d \in D \), \( \overline{d} \) is a *maximally coherent closed* set of types. A set \( X \subseteq |T| \) is

- **coherent** iff \( \forall t_1, t_2 \in X \text{ if } \exists t \in |T| \text{ such that } t_1 \prec^* t \text{ and } t_2 \prec^* t \text{ then } t_1 = t_2 \), and
- **closed** iff \( \forall t_1 \in X \text{ if } t_1 \prec t_2 \text{ then } t_2 \in X \).

where \( \prec^* \) is the transitive closure of \( \prec \). If \( t_1 \prec t_2 \) we say that \( t_1 \) is a *subtype* of \( t_2 \).

Subtypes of a type are taken to be incompatible and exhaustive. If a token \( d \) is of type \( t \) and \( t \) has subtypes \( \{t_1, t_2, t_3, t_4\} \) then \( d \) must be one of (and no more than one) subtype; either \( d : t_1 \) or \( d : t_2 \) or \( d : t_3 \) or \( d : t_4 \). The classification must have this property if it is to sort the tokens into distinct classes (which we take to be the function of a taxonomy).

For example, in Figure A.1 it is understood that leaves cannot both be prickly and have soft ends, and that rounded leaves are all either prickly or have soft ends. If we find a leaf which does not satisfy these conditions then we either take it as falling outside the scope of the taxonomy or else as an example motivating a change. We could introduce a third subtype of rounded leaves which picks out some property identifying this kind of leaf.

In informational terms, we can regard the closed coherent sets of types in a taxonomy to be *approximations* to the taxonomic categorisation of tokens. Each such set \( X \) provides partial information about the possible categorisation of a token, approximating the complete categorisation of \( d \) when \( X \subseteq \overline{d} \). In particular, if \( d : t \) then the set \( \{t' \in |T| \mid t \prec t' \} \) is an approximation to \( \overline{d} \). If, for example, we know that a leaf is dark green but have not noticed whether or not it is waxy then it has a partial specification in the taxonomy of Figure A.1, namely \{green, dark green\}, but we cannot decide on its complete categorisation. The distinction between partial and complete information in a classification will be explored further over the next few sections.
APPENDIX A. CLASSIFICATION SYSTEMS

The complete categorisation of tokens by a taxonomy is enforced by the requirement that type classes must be maximally coherent closed sets. This also has the consequence that every token is classified. There is no place for moles or hedgehogs in a leaf taxonomy.

It may have been noticed that the definitional conditions of a taxonomy do not constrain the subtype relation at all. Any binary relation is allowed so long as the classification respects our interpretation of 'subtype' as sorting tokens of a type into mutually exclusive and exhaustive sets. However, a consequence of this is that certain types will be incapable of classifying tokens at all if the subtype relation is not well behaved. We show the consequences of various properties of the subtype relation and define two special kinds of taxonomy.

Proposition A.2 For a taxonomy, \( (D, T, \prec) \), and \( t_1, t_2, t_3, \ldots, t_n \in |T| \) \( (n \geq 3) \),

1. if \( t_1 \prec t_2 \) and \( t_2 \prec t_3 \) then either \( t_1 = t_2 \) or \( t_1 \cap t_2 = \emptyset \),
2. if \( t_1 \prec t_2 \) and \( t_2 \prec t_1 \) then \( t_1 = t_2 \),
3. if \( t_1 \prec t_2 \), \( t_2 \prec t_3 \), and \( t_3 \prec t_1 \) then either \( t_3 = t_2 \) or \( t_1 \cap t_2 = \emptyset \), and
4. if \( t_1 \prec \ldots \prec t_n \) and \( t_1 \prec t_n \) then \( t_1 = \emptyset \)

Definition A.3 A taxonomy, \( (D, T, \prec) \), is hierarchical iff every coherent closed set \( X \subseteq |T| \) is a tree (i.e. every initial segment in the transitive closure of \( \prec \) restricted to \( X \) is linear). It is strictly hierarchical iff \( |T| \) is an upside down tree (i.e. every final segment in the transitive closure of \( \prec \) is linear).

The example depicted by Figure A.1 is a hierarchical taxonomy but it is not strictly hierarchical since the final segment generated by speckled, namely \{speckled, light green, green, yellow\} is not linear.

Proposition A.4 A taxonomy, \( (D, T, \prec) \), is hierarchical iff \( \prec \) is asymmetric.

A.1.2 State Systems

A common way of classifying things in science is by their state. For some domains we may have a very clear-cut way of distinguishing tokens, either motivated by our theoretical concerns or by the practical limitations of measurement. In a classification
based on the definition of an object’s state, we appeal to such a method for distinguishing tokens. If any two tokens have the same state they are considered indistinguishable as far as we are concerned, i.e. relative to the scientific concerns we have in classifying them in that way.

For example, we may know from our theory that the only relevant factor effecting the trajectory of a ball is its velocity (the mass of the ball and gravitational force being constant, and other factors such as drag being ignored as negligible). The ball’s velocity would provide an adequate definition of it’s state for the purpose of predicting it’s trajectory or performing any other calculations that such a model is suitable for. In practice, an observer tracking meteors can only plot the angle of elevation of the falling object, having no direct access to information about it’s velocity. Nevertheless, the plot the observer makes can provide a perfectly good criterion for specifying the state of a meteor given suitable assumptions about the atmosphere and overall orientation of the meteor (perhaps provided statistically).

What distinguishes the state of an object from any of its other properties is that it provides a complete specification of the object on which calculations can be performed, predictions made or explanations given. The colour of the ball or the meteor’s luminosity may not be part of such a specification, but only because they are irrelevant to the purposes of the scientist. If the observer is interested in explaining other aspects of the meteor’s descent then it’s luminosity could be irrelevant and then would be included in or calculable from the meteor’s state.

Definition A.5 A type s in a D-classifier is a state iff for all \(d_1, d_2 \in D\),

\[
\text{if } d_1, d_2 : s \text{ then } d_1 \approx d_2
\]

In other words, if \(d_1\) and \(d_2\) are both of type \(s\) ("in state \(s\)) then there is no other type that distinguishes them, i.e. no \(t\) such that either \(d_1 : t\) and \(d_2 \not\approx t\) or \(d_1 \not\approx t\) and \(d_2 : t\).

There are no examples of states in Figure A.1 since a complete specification always requires our saying whether or not the leaf is speckled, clean or waxy, and prickly or has soft ends. Nevertheless, any two tokens with the same classification by these types are indistinguishable. If we restrict the types to one of the two disconnected hierarchies then all of these types are states.
Proposition A.6 In a connected hierarchical taxonomy the states are the -minimal types.

States have disjoint (or equal) extensions, but they need not be disjoint from the extensions of other types. Types which are not states can therefore be regarded as approximations to states. But it is not necessary for there to be a state to which a type, or sequence of types approximate even in a connected hierarchical taxonomy. An example of such a taxonomy is the approximation of irrational numbers by rational binary divisions of an interval. The open interval (0,1) classifies all irrational numbers between 0 and 1. The subintervals (0,1/2) and (1/2,1) can be regarded as subtypes of (0,1) since they split the irrationals classified by (0,1) into two disjoint classes. Continuing in this fashion we get a taxonomy of irrationals between 0 and 1 which has the shape of a downwardly growing binary tree with infinite branches. Moreover, it is a connected strictly hierarchical taxonomy. Each interval in the construction approximates some (in fact many) irrationals, but there are no states since none of the intervals are minimal subtypes: we can always get a better approximation.

Definition A.7 A (D, S), is a state system iff S is a D-classifier all of whose types are states. |S| is called the state space of (D, S).

Proposition A.8

1. If s is a state in T and T ′ ⊆ T and s ∈ |T ′| then s is also a state in T ′.
2. T ′ \ {s | s is a state in T } is a state system.
3. If S is a state system and S ′ ⊆ S then S ′ is a state system.

In both of our examples, the state of an object was given by a single quantity: the velocity of the ball and the elevation plot of the meteor. More complex systems will use a number of quantities to specify a state. For example, we may be interested in the behaviour of a container of fluid, whose temperature, pressure, volume, are being continuously monitored. The state of the system at a given time can be represented by a vector, (67,42,26), whose components are the measurements of the water’s temperature, pressure and volume at that time. If we consider V T, V P, V to be the sets of
possible temperatures, pressures and volumes that may be recorded by the measuring instruments, the state of the fluid at a certain time is given a single point in the space $V_x \times V_p \times V_v$. In studying the behaviour of the fluid we can use this state space to classify individual time-slices of the fluid. The behaviour of the fluid over a period of time is then given by a path through the space.

In Chapter 6, we will be concerned to develop the theory of state systems relative to certain invariants on the token domain. A state distribution function is a function from a set of tokens to the state space. A definition of state for sets of tokens is given as equivalence classes of the sets under the invariants.

A.1.3 Feature Systems

In contrast to a classification by taxonomies or by states in which there is the supposition that a complete specification is being provided, a classification by features is supposed to be incomplete. We classify things by a list of features (it was brown, rectangular and moving very fast) when we do not know a precise enough description or even what would be relevant. A set of features is often an approximation to a complete specification, and so is used with the understanding that it could be improved if more were known (either about the token being classified, or about which features are relevant to the use which is being made of the classification). Although incomplete as a specification, a classification by features does provide some information about the tokens classified, since some features are incompatible with others. Incompatible features are ones that cannot occur together. If a token $t$ is classified as having the set $G$ of features then it may well have more features not appearing in $G$. But if a feature, $f$, is incompatible with $G$ then $t$ certainly does not have feature $f$. Knowing that something is rectangular is not much, but at least you know it’s not circular.

Definition A.9 A feature system, $(D, F, C)$, consists of a $D$-classifier, $F$, together with a collection $C$ of subsets of $|F|$ such that

1. for all $x, y \subseteq |F|$, if $x \subseteq y$ and $y \in C$ then $x \in C$, and
2. for all $f \in |F|$, $\{f\} \in C$
APPENDIX A. CLASSIFICATION SYSTEMS

Figure A.2: Incompatibility in a hierarchical taxonomy.

The set \(|F|\) is called the feature space of \((D,F,C)\) and its elements are called features. The sets of features in \(C\) are called the coherent sets of features. \((D,F,C)\) is said to be of order \(n\) if \(n\) is the smallest number for which the following condition is satisfied.

\[
\text{if } z \subseteq |F| \text{ and, } \forall f_1, \ldots, f_n \in z, \{f_1, \ldots, f_n\} \in C \text{ then } z \in C.
\]

If this condition is not satisfied for any \(n\) then \((D,F,C)\) is said to be of infinite order.

The most simple feature systems are generated from a notion of incompatibility between pairs of features. Given a \(D\)-classifier, \(F\), and a symmetric, irreflexive binary relation \(\perp\) on \(|F|\) such that for all \(d \in D\) and \(f, f' \in |F|\),

\[
\text{if } d \perp f \text{ and } f \perp f' \text{ then } d \notin f'.
\]

We define a feature system \((D,F,C_\perp)\) by

\[
C = \{z \subseteq |F| | \forall f_1, f_2 \in z, f_1 \perp f_2\}
\]

This feature system is of order at most 2 since coherence only depends on the lack of binary incompatibility between features. Order \(n\) feature systems can be defined in a similar way using \(n\)-ary incompatibility relations.

Definition A.10 For feature system \((D,F,C)\) we say that \(f, f' \in F\) are incompatible and write \(f \perp f'\) iff for all \(z \in C\) if \(f \in z\) then \(f' \notin z\).

Proposition A.11 A feature system \((D,F,C)\) is of order 2 at most iff \(C = C_\perp\).

Example In a taxonomy \((D,T,\prec)\) we define a binary incompatibility relation between types by
APPENDIX A. CLASSIFICATION SYSTEMS

<table>
<thead>
<tr>
<th>size</th>
<th>small</th>
</tr>
</thead>
<tbody>
<tr>
<td>activity</td>
<td>nocturnal</td>
</tr>
<tr>
<td>diet</td>
<td>insectivore</td>
</tr>
<tr>
<td>coat</td>
<td>spiny</td>
</tr>
<tr>
<td>face</td>
<td>pig-snouted</td>
</tr>
</tbody>
</table>

\[
t \perp t' \iff \text{there are distinct types } t_1, \ldots, t_n, t'_1, \ldots, t'_m, t'' \in [T] \text{ such that } t \equiv t_1, \\
t' = t'_1 \text{ and } t_1 < \ldots < t_n < t'' \text{ and } t'_1 < \ldots < t'_m < t''.
\]

For example, in Figure A.2 \( t_1 \perp t_2 \), but \( t_2 \not\perp t_3 \) and \( t_3 \not\perp t_4 \). Intuitively, two types are incompatible if they are sister subtypes of the same type or else have ancestors which are sister subtypes. In terms of classification relation, two types are incompatible just in case no token could possibly be of both types. So the use of the term 'coherent' used in the definition of taxonomies coincides with its use in the present example.

There is no natural notion of state in an arbitrary feature system \((D,F,C)\) because of the partial nature of classification by features. But if we lift a feature system to the level of coherent sets of features we can define a \(D\)-classifier with types \(C\) and \(d : x \iff d \vdash f \text{ for all } f \in x\). The maximally coherent sets of features are states in this classifier.

A.1.4 Attribute-Value Systems

A common means of classification in the cognitive sciences is by 'attribute-value' structures. The mathematical properties of attribute value structures has been extensively studied, but we base our account on Johnson (1988).

Figure A.3 shows what is known as a 'matrix representation' of two attribute-value structure. Matrix (a) says of what it is describing that it is of small size displaying nocturnal activity and has a pig-snouted face and a spiny coat and whose diet is of the kind insectivore. The items on the left side of the '=' sign represent 'attributes' whilst the items on the right side indicate the 'values' of those attributes. So (a) classifies small nocturnal pig-snouted spiny insectivores, like hedgehogs, for example.
Matrix (b), copied from Johnson's book, is more complex in two ways. Firstly, it has values which are themselves attribute-value structures. The matrix classifies the noun phrase "the man that I saw", saying of it that it is + definite, has a predicate man with a complement which is an attribute-value structure itself. Secondly, it has circled '1's (called co-indices) which indicate that, in this case, the value of the attribute obj is given by the whole matrix, making the structure circular. There can be further complexities, like attributes being values and whole attribute-value structures being attributes. Johnson gives a general definition of an attribute-value structure which incorporates all these possibilities:

**Definition A.12** An attribute-value structure, \((F, C, \delta)\) consists of a set, \(F\), called the set of attribute-value elements, a subset \(C \subseteq F\), called the set of constant elements, and a partial function \(\delta : F \times F \rightarrow F\) such that \(C \times F \cap \text{dom} \delta = \emptyset\).

Intuitively, as Johnson explains, \(\delta(f, a)\) is the value associated with attribute \(a\) in the attribute-value element \(f\). Attributes are themselves attribute value elements. In fact, any attribute value element can act as an attribute. The constant elements are special attribute-value elements which have no attributes. All the words in Figure A.3 represent constant elements. It is important to note that it is the attribute-value elements in an attribute-value structure that are depicted by matrices or parts of matrices. An attribute-value structure contains many attribute-value elements, some of which are values of other elements, but others may only share values, or may even be completely unconnected.

In using attribute-value structures to classify tokens of one kind or another we must be more precise than we have been about the relationship between the tokens classified and the structures which classify them. Just to say that a token, a hedgehog, say, is classified by a certain structure \((F, C, \delta)\) is not enough since it is not clear which attributes the hedgehog is supposed to have. Which attribute-value element \(f\) in the structure do we look at to find the attributes \(a\) for which a value \(\delta f, a\) is given? In certain cases there may be a clear answer. For example, if \((F, C, \delta)\) is rooted (i.e. there is a unique \(f_0 \in F\) not in the range of \(\delta\)) then the root element would seem to be an obvious choice. Structures like the one depicted in Figure A.3 (b) present an (albeit surmountable) difficulty, since they are not rooted but there is a clear choice as to the
APPENDIX A. CLASSIFICATION SYSTEMS

element we are to choose. In general, however, an attribute-value structure may present no clear candidate: there may be many partially or entirely disconnected elements in $F$.

A plausible solution is to classify tokens with attribute value elements in a structure rather than with the structures themselves. A single structure would then be a classifier whose elements are the types. The problem with this approach is that the types would not be individuated by what we intuitively take to be their classificatory role. If we take two ‘isomorphic’ rooted attribute-value structures $(F_1, C_1, A_1)$ and $(F_1, C_1, A_1)$ then in their disjoint union the two roots would appear to have identical classificatory roles. A response to this problem might be to restrict our attention to extensional attribute-value structures in which no two distinct elements generate ‘isomorphic’ substructures.

But in doing this we loose the ‘shared value’ distinctions implicit in the co-indexing convention of the matrix pictures.

Our approach to these problems is to classify tokens directly by attributes-value pairs in a context in which the connections between values are given relationally. Starting with a $D$-classifier, $T$, we think of the types of $T$ as single attribute-value pairs, like *face*, *pig-snouted* or *pred., man*. The type class of a token in $D$ is therefore a collection of such pairs. Each type must therefore indicate an attribute and a value. When the value is complex the type ‘points’ to the set of types making up the complex value. There is no problem involved in a type pointing to a set of types which contain it, so the circular structures discussed by Johnson are incorporated.

**Definition A.13** An attribute-value system is a structure, $(D, T, e_A, e_V, \alpha, \mu)$, where $e_A$ and $e_V$ are equivalence relations, and $\alpha$ and $\mu$ are binary relations on $|T|$ such that, for all $x_1, x_2, y_1, y_2 \in |T|$,

1. if $(x_1, x_2) \in e_V$ then for all $y \in |T|$, $(x_1, y) \in \mu$ iff $(x_2, y) \in \mu$,

2. if $(x_1, y_1), (x_1, y_2) \in \alpha$ then $(x_1, x_2) \in e_A$ iff $(y_1, y_2) \in e_V$,

3. $\forall x \in |T|, (y | (x, y) \in \mu)$ is coherent, and

\footnote{What isomorphism amounts to in this setting is not entirely clear, since, for example, it may or may not be the case that we want the sets of constants $C_1$ and $C_2$ to be identical.}

\footnote{This distinction is studied explicitly in Reape (1990) using a modal logic interpretation.}
APPENDIX A. CLASSIFICATION SYSTEMS

4. Vd ∈ D, z is coherent,

where X ⊆ |T| is coherent iff ∃ z1, z2 ∈ X, if (z1, z2) ∈ eA then z1 = z2.

Intuitively, the equivalence relations eA and ev capture the relationship of having the ‘same attribute’ and ‘same value’ respectively. When a type has a complex value, the relation µ connects it to each of the types in the set representing that complex value. For example, if a type t has the value depicted by Figure A.3 (a) then (t, t') ∈ µ just in case the attribute and value of t' are those depicted by one of the entries in the matrix. When a type has a constant value then it is not related to anything by µ (although other types may be related to it). The relation α indicates which values correspond to which attributes: if (t1, t2) ∈ α then the attribute of t1 is the value of t2. Johnson allows attributes to occur as values and α is needed to ensure that this can be given an interpretation within the classifier. Although we speak of types ‘having’ attributes and values, they have no internal structure: everything about them is determined by the relations on |T|.

Example An attribute-value system is naturally thought of as a feature system of order 2. Two attribute-value pairs are incompatible if they share the same attribute but have different values. In fact, given an attribute-value system (D, T, eA, ev, α, µ) the feature system (D, T, C1), where

\[ t_1 \perp t_2 \iff (t_1, t_2) \in eA - ev, \]

has sets X ∈ C1 iff X is coherent in the sense of Definition A.13.

We now show how an attribute-value structure can be constructed out of an attribute-value system.

From Systems to Structures: Given an attribute-value system, \((D, T, eA, ev, \alpha, \mu)\), we define an attribute-value structure, \((F, C, \delta)\), by letting

\[ F = |T|/ev, \]

\[ C = \{ c \in F \mid \forall z \in c, x' \in |T|, (x, x') \notin \mu \} \]

and \( \delta : F \times F \rightarrow F \) be defined as \( \delta^1(f_1, f_2) \) when this is non-empty and undefined otherwise. \( \delta^1 : F \times F \rightarrow F \cup \{\emptyset\} \) is defined by
\[ \delta^+(f_1, f_2) = \bigcup \{ \text{ev}(x) \mid (x, y) \in \mu \} \]

where \( \text{ev}(x) = \{ x' \mid (x, x') \in \text{ev} \} \). To show that \((F, C, \delta)\) is an attribute-value structure we have to show that \(\delta^+(f_1, f_2)\) is either empty or an ev equivalence class for all \(f_1, f_2 \in F\), and that \(\delta^+(c, f) = \emptyset\) for \(c \in C, f \in F\).

Firstly, if \(\delta^+(f_1, f_2) \neq \emptyset\) then \(\exists x_1 \in \delta^+(f_1, f_2)\). So \(\exists x'_1 \in [T]\) such that \((x_1, x'_1) \in \text{ev}\) and \(\exists y_1, y_2 \in f_1, f_2\) such that \((x'_1, y_2) \in \alpha\) and \((y_1, x'_1) \in \mu\). We will show that \(\delta^+(f_1, f_2)\) is the ev equivalence class of \(x_1\). Clearly the equivalence class of \(x_1\) is contained in \(\delta^+(f_1, f_2)\). We need to show the converse.

If \(x_2 \in \delta^+(f_1, f_2)\) then \(\exists x'_2 \in [T]\) such that \((x_2, x'_2) \in \text{ev}\) and \(\exists y'_1, y'_2 \in f_1, f_2\) such that \((x'_2, y'_2) \in \alpha\) and \((y'_2, x'_2) \in \mu\). Now \((y'_2, y'_2) \in \text{ev}\), since \(f_1\) is an ev equivalence class, so by Definition A.13.1,

\[ \{ y \in [T] \mid (y_1, y) \in \mu \} = \{ y \in [T] \mid (y'_1, y) \in \mu \}. \]

Also, \((y_2, y'_2) \in \text{ev}\), \((x'_1, y_2) \in \alpha\) and \((x'_2, y'_2) \in \alpha\). So, by Definition A.13.2, \((x'_1, x'_2) \in \epsilon_A\), and thus \(x'_1 = x'_2\) by Definition A.13.3., and so \((x_1, x_2) \in \text{ev}\). Hence \(\delta^+(f_1, f_2)\) is contained in the ev equivalence of \(x_1\). So, when \(\delta(f_1, f_2)\) is defined it is an element of \(F\).

It only remains to show that \(\delta\) is not defined for \((c, f)\) when \(c \in C\). But if \(c \in C\) then for all \(x \in c\) and \(x' \in [T]\), \((x, x') \notin \mu\) so, for any \(f \in F\), \(\delta^+(c, f) = \emptyset\) and so \(\delta\) is undefined for \((c, f)\).

The construction of an attribute-value system from an attribute-value structure cannot be complete because, as we have noted, the latter does not contain enough information about how to classify tokens. Nevertheless, we will give a construction which determines the type structure of an attribute-value system without saying anything about the classification relation.

**From Structures to Systems** If \((F, C, \delta)\) is an attribute-value structure then we define an attribute-value system, \((D, T, \epsilon_A, \text{ev}, \alpha, \mu)\) modulo the classification relation \(\gamma\).

The construction is somewhat indirect since we first extract sets \(A\) and \(V\) of attributes
APPENDIX A. CLASSIFICATION SYSTEMS

and values from \((F, C, \delta)\). The attributes are defined to be the elements \(a\) of \(F\) for which \(\delta\) is defined on \((f, a)\) for some \(f\), i.e. those elements which act as attributes in \((F, C, \delta)\). The values are defined to be those elements that act like values in \((F, C, \delta)\) i.e. those in the range of \(\delta\). To avoid constructing too many types, we restrict ourselves to pairs \((a, v)\) where \(v\) is a value that \(a\) has in the structure. Thus we also need to define the sets \(V_a\) of values of \(a\).

\[
A = \{a \in F \mid \exists f \in F | (f, a) \in \text{dom} \delta \}.
\]

For \(a \in A\), \(V_a = \{\delta f, a \mid (f, a) \in \text{dom} \delta\} \).

\[
V = \bigcup_{a \in A} V_a.
\]

\[
[T] = \Sigma_{a \in A} V_a \ (= \{(a, v) \mid a \in A, v \in V_a\}).
\]

The equivalence relations \(e_A\) and \(e_V\) are the relations of 'same attribute' and 'same value', i.e.

\[
e_A = \{(a, v), (a, v') \mid a \in A, v, v' \in V_a \}
\]

\[
e_V = \{(a, v), (a', v) \mid a, a' \in A, v \in V_a \cap V_{a'} \}
\]

These are clearly equivalence relations. Since we are modelling attributes and values by their equivalence classes we need the additional relation \(a\) to capture the way in which elements can act as both attributes and as values. This is given by

\[
a = \{(a, v), (a', a) \mid a' \in A, a \in A \cap V_a, v \in V_a \}
\]

Finally, the relation \(\mu\) captures the distinction between complex and atomic elements by relating each attribute-value pair \((a, v)\) to the set of attributes-value pairs which \(\delta\) determines at \(v\), i.e.

\[
\mu = \{(a, v), (a', \delta(v, a')) \mid a \in A, (v, a') \in \text{dom} \delta \}
\]

We must check that each of the first three conditions of Definition A.13 hold. In each case the check is simply a matter of inspecting the relevant definitions.
APPENDIX A. CLASSIFICATION SYSTEMS

A.2 Heterogeneous Classification Systems

In this section we see how the classification systems of Section A.1 can be related together, providing a basis for heterogeneous classification systems (i.e. ones made up of systems of different kinds). First we show how feature systems and state systems can sometimes be represented in attribute-value systems and how, when represented, the relationship between features and states can be characterized. Second, we explore the properties of a mixed system involving a feature system and an arbitrary classifier. The results we obtain are put as solutions to the problem of relating feature to taxonomic classifications in a general setting. Pursuing the problem of 'feature-based recognition' a bit further, we reject the type-feature system approach in favour of a more flexible but related approach involving 'inheritance principles'.

A.2.1 Attribute-Value Representations

As was observed in Section A.1.4, an attribute-value system can be thought of as a feature system of order 2. The question arises as to when a feature system can be represented as an attribute-value system.

Definition A.14 An attribute-value representation of a feature system \((D,F,C)\) is given by an attribute-value system \((D,T,\varepsilon_A,\varepsilon_V,\alpha,\rho)\) and a \(D\)-classifier isomorphism \(\rho : |F| \rightarrow |T|\) such that, for \(x \subseteq |T|\),

\[ x \in C \text{ if } \{\rho(t) : t \in x\} \text{ is coherent in } (D,T,\varepsilon_A,\varepsilon_V,\alpha,\rho). \]

Proposition A.15 If a feature system has an attribute-value representation then it is of order 2 at most.

However, not all order 2 feature systems have attribute-value representations.

Proposition A.16 If \((D,F,C)\) has an attribute-value representation then

- for distinct features \(f_1, f_2, f_3 \in |F|\), if \(f_1 \perp f_2\) and \(f_1 \perp f_3\) then \(f_2 \perp f_3\).
- if \(x \in C\) and \(x \cup \{f\} \notin C\) then there is a unique feature \(f' \in x\) such that \(f \perp f'\).

Proof: Let \(\rho\) be a representation of \((D,F,C)\) in the attribute-value system \((D,T,\varepsilon_A,\varepsilon_V,\alpha,\rho)\).

By Propositions A.15 and A.11, \(C = C_L\).
APPENDIX A. CLASSIFICATION SYSTEMS

1. If \( f_1 \perp f_2 \) and \( f_1 \perp f_3 \) then \( \{\rho f_1, \rho f_2\} \) and \( \{\rho f_1, \rho f_3\} \) are not coherent in \((D,T,e_A,\alpha,\mu)\). So \( \{\rho f_1, \rho f_2\}, \{\rho f_1, \rho f_3\} \in e_A - ev \) (see Example A.1.4), so \( \{\rho f_1, \rho f_2\} \in e_A \).

Now suppose, for contradiction, that \( f_2 \not\perp f_3 \). Then \( \{\rho f_2, \rho f_3\} \in C \) and so \( \{\rho f_2, \rho f_3\} \) is coherent in \((D,T,e_A,\alpha,\mu)\). By the definition of coherence in an attribute-value structure, this implies that \( \rho f_2 = \rho f_3 \), contradicting the fact that \( \rho \) is injective.

2. If \( x \in C \) and \( x \cup \{f\} \notin C \) then the existence of an \( f' \in x \) such that \( f \perp f' \) follows directly from the fact that \( C = C_A \). Uniqueness follows, since if there were distinct \( f_1, f_2 \in x \) such that \( f \perp f_1 \) and \( f \perp f_2 \) then, by 1., \( f_1 \perp f_2 \), which contradicts the assumption that \( x \in C = C_A \).

\[ \square \]

States can also be thought of in terms of attributes and values. If the state of a physical system is being measured by instruments \( m_1, \ldots, m_n \), then a natural way of representing a particular state is by a sequence \( (m_1, \ldots, m_n) \) consisting of the values \( m_i \) indicated by each instrument \( m_i \). A state may therefore be represented by a coherent set of attribute-value pairs, one for each attribute in a fixed set of attributes specifying the system.

**Definition A.17** A attribute-value representation of a state system \((D,S)\) consists of an attribute-value system \((D,T,e_A,\alpha,\mu)\) and a \( D \)-classifier isomorphism \( \sigma : |S| \rightarrow |T| \upharpoonright M \) where

\[
M = \{ x \in |T| \mid x \text{ coherent}, \forall t \in |T| \exists x' \in x, (x,x') \in e_A \}
\]

The setting of attribute-value representations provides a convenient way of comparing feature systems with state systems. For a feature system, \((D,F,C)\), the notion of state arises by considering maximally coherent sets of features. If \((D,F,C)\) has a representation \( \rho : |F| \rightarrow |T| \) in the attribute-value system \((D,T,e_A,\alpha,\mu)\) then the state system, \((D,S)\), with state space

\[
|S| = \{ x \in C \mid \forall y \in C \text{ if } x \subseteq y \text{ then } z = y \}
\]
APPENDIX A. CLASSIFICATION SYSTEMS

and \( d \colon x \iff \exists z = x \), has the representation \( \sigma : |S| \to |T \uparrow M| \), where \( M \) is as in Definition A.17 given by, for all \( x \in |S| \),

\[
\sigma(x) = \{ \rho(f) \mid f \in z \}
\]

That \( \sigma(x) \) is coherent in \( (D,T,\varepsilon_A,\varepsilon_V,\alpha,\mu) \) follows from the fact that \( \rho \) is a representation of \( (D,F,C) \). To check that \( \sigma(x) \in M \) we must show that for any \( t \in |T| \) there is an \( t' \in \sigma(x) \) such that \( (t,t') \in \varepsilon_A \). But \( \rho \) is surjective, so \( t = \rho(f) \) for some \( f \in |F| \) and, since \( z \) is maximally in \( C \), either \( f \in x \) and we are done, or \( x \cup \{ f \} \notin C \). But then, by Proposition A.16 there is a (unique) \( f' \in z \) such that \( f' \perp f \). So, following Example A.1.4, \( (t',t) \_ \_ (\rho f',\rho f) \in e_A - ev \) as required.

If we start from a state system, \( (D,S) \), then features are naturally modeled as invariant aspects of the states. If \( (D,S) \) has a representation \( \sigma : |S| \to |T \uparrow M| \) in the attribute-value system \( (D,T,\varepsilon_A,\varepsilon_V,\alpha,\mu) \) (\( M \) as before) then we can define a map \( \rho : |T| \to \mathcal{P} |S| \) by

\[
\rho(t) = \{ s \in |S| \mid t \in \sigma(s) \}
\]

for \( t \in |T| \), which is injective. (To show that \( \rho \) is indeed injective amounts to the observation that if \( t \neq t' \in |T| \) then there is an \( m \in M \) such that \( t \in m \) but \( t' \notin m \)).

Then the feature system \( (D,F,C_\perp) \) with \( |F| = \mathcal{P} |D| \) and, for \( d \in D \) and \( f \in |F| \),

\[
d : f \iff d : s \text{ and } f \in \rho s \text{ for some } s \in S,
\]

which is well-defined since \( \rho \) is injective, and

\[
\rho t \perp \rho t' \iff (t,t') \in \varepsilon_A - ev
\]

for \( t, t' \in |T| \) has the representation \( \rho^{-1} : |F| \to |T| \) in \( (D,T,\varepsilon_A,\varepsilon_V,\alpha,\mu) \)

Moreover, these two constructions are mutually inverse up to isomorphism.

A.2.2 Type-Feature Systems

One might think that the results of Section A.2.1 show that feature systems and state systems are best thought of in terms of their attribute-value representations. But Proposition A.15 tells us that we cannot have attribute-value representations for feature systems of order greater than 2. This means that such representations are only good for
studying pairwise incompatibility of features. In practice, however, we may wish to account for higher orders in incompatibility. For example, a feature system for a domain of polygons may have features like having equal sides, having a right angle, having \( n \) sides, and having equal angles. There are no pairwise incompatibilities amongst the features \{ equal sides, right angle, 3 sides \} but they cannot co-occur.

We already know how to generalize the notion of incompatibility in a feature system to arbitrary orders using the idea of coherent sets of features. In this section we explore mixed systems of classification in which the compatibility of features in a feature system are determined by a map \( \rho \) from a separate classifier, \( T \), mapping types of \( T \) to sets of features. The coherent sets of features are defined as those in the range of \( \rho \).

**Definition A.18** \((D,F,T,\rho)\) is a type-feature system iff \((D,F,\text{ran} \rho)\) is a feature system, \( T \) is a \( D \)-classifier and \( \rho : [T] \rightarrow \text{pow} [F] \) is a \( D \)-classifier morphism from \( T \) to \( F \cap \text{ran} \rho \). This definition is illustrated in Figure A.4.

**Proposition A.19** Given a type-feature system \((D,F,T,\rho)\),

1. every token of \( D \) is classified in \( T \), and,
2. for \( d,d' \in D \), if \( d \equiv_F d' \) then \( d \equiv_T d' \).

**Proof:**

1. Since \((D,F,\text{ran} \rho)\) is a feature system, each \( d^F \in \text{ran} \rho \). But \( \rho \) is surjective so there is a \( t \in [T] \) such that \( \rho t = d^F \) and \( d \Rightarrow_T t \), i.e. every element in the domain is classified by \( T \).
2. For $d, d' \in D$, if $t \in d^F$ then $pt \subseteq d^F$. If $d \approx_F d'$ then $d^F = d'^F$ so $pt \subseteq d^F$ and so $t \in d^F$. Hence $d^T \subseteq d^F$ and, similarly, $d'^T \subseteq d'^F$ so $d \approx_T d'$.

We can now express the general relationship between states and features within the composite system.

**Proposition A.20** In a type-feature classification, $(D, F, T, \rho)$, if $x \in \text{ran} \rho$ is maximal $(\forall y \in \text{ran} \rho \text{ if } x \subseteq y \text{ then } x = y)$ there is a state $s$ of $T$ such that $\rho s = x$.

**Proof:** If $x \in \text{ran} \rho$ then there is a $t \in |T|$ such that $pt = x$. For any $d \in D$, if $d \supseteq t$ then $x = pt \subseteq d^F$. But $d^F \in \text{ran} \rho$ so, by the maximality of $x$, $x = d^F$. So for any $d_1, d_2 \in D$, if $d_1, d_2 \supseteq t$ then $d_1^F = d_2^F$ and so

$$d_1^T = \{ t \in |T| \mid pt \subseteq d_1 \} = \{ t \in |T| \mid pt \subseteq d_2 \} = d_2^T.$$ 

Hence $d_1 \approx_T d_2$ and so $t$ is a state in $T$.

The converse implication does not hold. If $s$ is a $T$ state then $\rho s$ may not be maximal in $\text{ran} \rho$. It may be the case that every token with features $\rho s$ also has other features. In the attribute-value representations it is tacitly assumed that attributes are independent, but this may not be the case. In our example of the container of fluid being monitored for temperature, pressure and volume, there is a dependency between these three attributes, which allows any pair of them to completely specify the system. Similarly, in the case of the ball's trajectory, the mass of the ball is assumed constant and so its velocity alone is sufficient to define its state, relative to the concerns of the scientist. In a classification of many balls with different masses, the velocity of a ball would no longer be adequate for specifying its state. Nevertheless, a weaker condition does hold, namely: if $s$ is a $T$ state and $d_1, d_2 \supseteq s$ then $d_1 \approx_F d_2$, i.e. any tokens of type $s$ are featurally indistinguishable.

**A.2.3 Feature Based Recognition**

One important way in which the more serious biological classification differs from our school room example is that the words we use to name the types are usually semantically
Figure A.5: Feature Recognition

redundant, artificial and in Latin. To the expert they can contain information about the their position in the hierarchy ('Ilex α' and 'Ilex β' are species of the same genus), but they are rarely as descriptive as 'pointy' or 'prickly'. Because the definition of species is essentially an historical one based on ancestry, there are little or no necessary connections to non-historical properties that the tokens of a certain type might have, such as being pointy.

A development of the idea of a taxonomy allows features to be associated with its types. Given a taxonomy \((D, T, \prec)\) and a feature system \((D, F, C)\), we introduce an association, \(t \text{ has } f\) between the types \(t \in [T]\) and the features \(f \in [F]\), as depicted in Figure A.5. The key question, of course, is what constrains the three relations \(\prec, \prec_F\) and \(\text{has}\)? Answers to this question are guided by different kinds of feature recognition strategy. Given a token \(d\), we can approach the problem from two angles:

- if \(d\) has certain features then what type is it, and
- if \(d\) is of a certain type then what features should we expect it to have.

The simplest answer to these questions is provided by the type-feature classifications of Definition A.18. If \((D, F, T, \rho)\) is a type-feature system for which \((D, T, \prec)\) is a taxonomy then we say that \(t \text{ has } f\) \iff \(f \in pt\). There is a simple relationship between \(\prec, \prec_F\) and \(\text{has}\) given by \(d \prec t \iff pt \subseteq \mathcal{D}^f\), or in answer to our questions:
APPENDIX A. CLASSIFICATION SYSTEMS

- if \( t : f \) for every \( f \) such that \( t \text{ has } f \) then \( t \rightarrow t \) (i.e. a token is of a type if it has all the features associated with that type)

- if \( t \rightarrow t \) and \( t \text{ has } f \) then \( t : f \) (i.e. if a token is of a type then we expect it to have all the features associated with that type)

We also note that for any \( d \in D \) there is a state \( s_d \in |T| \) such that \( d : s_d \), i.e. the taxonomy is so finely graded that there is a complete specification of each token up to featural indistinguishability.

This simple relationship is all very well when every coherent set of features is represented by a type in the hierarchy. In general, we want a taxonomy to ignore some features and favour others in order to classify tokens in an economical way. A taxonomy can only be the basis for generalizations if it classifies more coarsely. As a step towards this we might consider a subclassifier, \( H \preceq T \), as the basis for the taxonomy. The set of features of a token, although coherent, need not to be represented by a type in the subclassifier.

Nevertheless, we can define a set of features \( I_H(d) \) which measures the "featural degree" to which a token \( d \) is classified by the subclassifier \( H \), i.e. such that \( I_H(d) \subseteq \Delta^F \) and if any other token \( d' \) has \( I_H(d') \subseteq \Delta^F \) then \( d \approx_H d' \).

**Definition A.21** Given a type-feature system, \( (D, F, T, p) \), and a taxonomy, \( (D, H, \prec) \) such that \( H \preceq T \), the features inherited by a token \( d \in D \) are given by

\[
I_H(d) = \bigcup \{ pt \mid d : H t, t \in |H| \}
\]

Note that for each token \( d \) there is a \( t \in |T| \) such that \( pt = I(d) \), but \( t \) need not be in \( |H| \).

The solution offered by using subclassifiers is not sufficient to account for all aspects of feature-based recognition. The most common problem is that of defaults. We may have \( t_1 \prec t_2 \) but \( pt_1 \cup pt_2 \notin \text{map} \). For example, leaves of the species 'Ilex a' are dark green, prickly, smooth, smaller than your hand and have a waxy texture. But the subspecies 'Ilex b' has leaves which are yellow. Clearly no subclassifier will respect the subtype relation for such a hierarchy, since if \( d \rightarrow t_1 \) and \( d \rightarrow t_2 \) then \( pt_1 \subseteq \Delta^F \) and \( pt_2 \subseteq \Delta^F \), so \( pt_1 \cup pt_2 \subseteq \Delta^F \notin \text{map} \) and so, by Definition A.9.1, \( pt_1 \cup pt_2 \notin \text{map} \).
APPENDIX A. CLASSIFICATION SYSTEMS

Generalizing, we need a taxonomy $H$, with $|H| \subseteq |T|$, which respects an inheritance principle:

$$\text{for all } d \in D, I_H(d) \subseteq F$$

and if $d' \in D$ and $I_H(d) \subseteq F$ then $d \sim_H d'$.

To account for defaults we follow the general form of this principle whilst changing the kind of inheritance employed.

Definition A.22 Given a type-feature system, $(D, F, T, \rho)$, and taxonomy $(D, H, \sim)$, with $|H| \subseteq |T|$, the features default inherited (top-down) by a token $d$ are given by

$$D, I_H(d) = \bigcup_{d \in H} \{ f \in \rho t \mid \forall t' \in F \ \text{either} \ t < t' \ \text{or} \ \rho t' \cup \{ f \} \in \text{map} \}$$

In other words, if a token is of type $t$ then it has any feature, $f$, associated with $t$ unless it is also of a subtype $t'$ of $t$ associated with features incompatible with $f$. Features associated with a type are regarded as defaults which can be overridden by features associated with its subtypes.

The definition of the default inheritance map is top down in that it supplies independent criteria for the inheritance of features of each type. The criterion for a feature $f$ associated with type $t$ (i.e. $f \in \rho t$) to be (default) inherited by token $d$ is that $d \sim_H t$ and $f$ is compatible with all the features associated with all the types in $F$ other than those associated with types of which $t$ is a subtype. In other words, for types $t, t' \in D$

- if $t < t'$ then the features of $t$ override any features of $t'$ with which they are incompatible.
- if neither $t < t'$ nor $t' < t$ then only the mutually compatible features of $t$ and $t'$ can be inherited.

The later condition shows that we are taking a weak line\(^3\) on the problem of multiple inheritance. The default principle resolves conflicts in inheriting features of types in favour of the subtype. But when the conflict arises between the features of unrelated types, there seems to be no principled way of deciding which type should be given precedence. The condition above ensures that at least all those features which are mutually

\(^3\)This is the generally accepted line that is taken in AI. See, for example, Touretzky (1986).
compatible get inherited, whilst allowing the possibility that the conflict between incompatible features may be resolved in some way. Note, however, that this solution is only possible for feature systems of order 2. In the general case, \( I(d) \) is not guaranteed to be coherent.

When \( H \) is hierarchical and the initial segments in \( \bar{d}^H \) are well-ordered in addition to being totally ordered (i.e. each subset has a least member), a bottom-up default inheritance map can be defined recursively over the order type of the initial segments:

**Definition A.23** Given a type-feature system, \( \langle D, F, T, \rho \rangle \), and taxonomy \( \langle D, H, \prec \rangle \), with \( |H| \subseteq |T| \) and such that, for each token \( d \in D \), the initial segments of \( \bar{d}^H \) are well-ordered, we define the \( D \)-classifiers \( H_\alpha \), for each ordinal \( \alpha \), by

\[
d : \cdot d_\alpha \text{ t } \text{ if } d : \cdot d_\alpha \text{ t and the initial segment of } t \text{ in } \bar{d}^H \text{ is of order type } \alpha
\]

The features \( \alpha \)-inherited by a token \( d \) are recursively specified by

\[
D_2 I(d) = \bigcup_{\alpha \in \alpha} D_2 I(d) + T \bigcup_{\alpha \in \alpha} \{(f) \in \rho(d) \text{ if } d : \cdot d_\alpha \text{ t} \text{ then } d \in \rho_\alpha \}
\]

where \( A + T \text{ } B = A \cup \{ f \in B \} \cup A \in \rho_\alpha \). The features, \( D_3 I(d) \), default inherited bottom-up by \( d \) are those that are \( \alpha \)-inherited for some ordinal \( \alpha \).

Even if initial segments of type classes are well-ordered, the two definitions of default inheritance do not in general coincide. An illustration of this is given in Figure A.6. If there are types \( t < t' \prec t'' \in \bar{d}^H \) and features \( f_1 \in \rho t, f_2 \in \rho t', f_3 \in \rho t'' \) such that \( f_1 \downarrow f_2 \) and \( f_2 \downarrow f_3 \) then \( \text{ceteris paribus} \) both \( f_2 \) and \( f_3 \) are inherited: \( f_2 \) because it conflicts with \( f_1 \) and \( f_3 \) because it conflicts with \( f_2 \). On the top-down definition of inheritance, \( f_3 \) is not be inherited since it is incompatible with \( f_2 \). But on the bottom-up definition \( f_2 \) is not inherited since it is incompatible with \( f_1 \), and so there is nothing to stop \( f_3 \) from being inherited. Nevertheless, it is worth remarking that the two principles agree when \( F \) has an attribute-value representation, i.e. for all \( d \in D \), \( D_1 I(d) = D_3 I(d) \) (see Proposition A.16).

Finally, we note that the principles espoused in this section can be applied to inheritance of types in any classifier that can be given a feature system structure. For example, we observed in Section A.1.3 that a notion of incompatibility can be given for the types of a taxonomy. Applying the inheritance principles from this section we would be able to develop a recognition strategy connecting classifications by different taxonomies.
A.3 Classifying Words

Modern linguistic theories use a variety of mathematical structures to model their syntactic analyses of linguistic expressions. Perhaps the most popular structure in use is the tree. Many grammatical formalisms take words, phrases and sentences to be organised in a tree-like way. In general, words occur at the leaves of the tree and the nodes of the tree represent various combinations of words into phrases. The root of the tree usually represents a whole sentence.

Even in formalisms which do not explicitly use trees, there is an implicit tree in any approach which takes sequences of words as the basic unit of analysis: namely the tree given by the subsequence ordering. However, in what follows we will focus on the classification of words rather than sequences of words. We will be concerned with relationships between words, but these relationships will in general not form trees. There are two reasons for this choice. Firstly, the approach to classification we have been advocating takes the tokens of a classification to be unstructured. It is instructive to see how far this ‘purist’ attitude can be taken in practice. Secondly, the word-based approach is characteristic of the linguistic tradition of ‘dependency grammar’ and, more
specifically, 'Word Grammar' (see e.g. Hudson 1984, Volino 1990), but little work has been done in formalising these grammars.

By a 'word' we mean a particular instance a type of word, i.e. an occurrence of 'mole' rather than the abstraction across all such occurrences. In practice, the segmentation of an utterance into words may be ambiguous or, at least, highly context dependent. But then, so is the individuation of many useful environmental invariances, physical objects included.

All grammars use a classification of individual words at at least one level of analysis: the level of lexical decomposition. A particular word token is classified as being of a certain grammatical category (noun, verb, adjective, etc.) and having certain features (case, number, gender, etc.) in virtue of its lexical type. Depending on the flavour of the linguistic theory, more or less information about the word may be specified in this kind of classification, which is known as a lexicon.

We model lexicons using type-feature systems \((W, F, L, \rho)\) where \(W\) is a domain of word tokens. The types, \(|L|\) form the basic classificatory framework in which grammatical constraints operate, and each lexical type is related by \(\rho\) to a coherent set of lexical features. We can think of the feature space \(|F|\) as including all the familiar agreement features of case, number and gender as well as grammatical categories like noun and verb. Other features like subject and object may well also be included, although they need not classify any words in the lexicon. So for each lexical type \(t\), \(pt\) consists of sets of lexical features which 'go together' like \{ noun, accusative, plural \} but not \{ verb, noun \}. The relation \(\lambda\) is taken to be modelling just this level of incompatibility.

**Definition A.24** A \(W\)-lexicon, \((W, F, L, \rho)\) is a type-feature system.

It should be noted that 'lexical types' do not distinguish words below the featural level (Proposition A.19). They may or may not distinguish occurrences of the word 'John' from occurrences of the word 'Peter', depending on which aspect of linguistic analysis is being modelled. For certain purely syntactic matters, the distinction between the two may not be important, whereas it is clearly significant at some level for an understanding of how these words work.
APPENDIX A. CLASSIFICATION SYSTEMS

A small mole burrowed furiously
(a)

a small mole burrowed furiously
(b)

Figure A.7: A Phrase Structure Tree and a Dependency Structure.

The notion of dependency in linguistics is an old one (dating back to the earliest grammarians, (Pāṇini c. fourth century B.C.) Certain words seem to depend on other words (their heads) for their occurrence in linguistic expressions. Since we can say "the hungry hedgehog snorted" and "the hedgehog snorted" but not "the hungry snorted" we conclude that 'hungry' is a dependent of 'hedgehog' rather than the other way round. Mere occurrence, or the lack of it, is not a good enough criterion for identifying the dependency relation which, like many linguistic primitives, is not amenable to precise definition. Figure A.7 shows the structures obtained from the simple English sentence "a small mole burrowed furiously" by typical linguistic analyses based on (a) phrase structure and (b) dependency.

In a domain of word tokens, \( W \), a dependency structure is given by an asymmetric relation \( \alpha \) between tokens of \( W \), for which \((w_1, w_2) \in \alpha \) is interpreted as modelling the fact that word \( w_2 \) is a dependent of word \( w_1 \) (alternatively, \( w_1 \) is the head of \( w_2 \)). For example, the arrows in Figure A.7 go from head \( w_2 \) to dependent \( w_1 \) so depicting the dependency relation: \{ burrowed, a \},( burrowed, furiously ),( a, mole ),( mole, small )

Within linguistic expressions, especially sentences, there is always a particular word, called the root, which is not the dependent of any other word in the expression. In
APPENDIX A. CLASSIFICATION SYSTEMS

Figure A.7 we can see that the root of the sentence is the main verb, 'burrowed'. Our definition of a 'dependency structure' requires that every structure has a unique root.

Definition A.25 A $W$-dependency structure $\alpha$ is an asymmetric binary relation on $W$ together with a distinguished word $o_0 \in W$, called the root of $\alpha$, such that $dom o_0 \subset \{o_0\}$. The extent, $|\alpha|$, of $\alpha$ is defined to be $\{o_0\} \cup ran o_0$.

Dependency structures are partially ordered by $\alpha \preceq \alpha'$ if $\alpha \subseteq \alpha'$, $|\alpha| = |\alpha'|$ and $o_0 = o'_0$.

The 'rules' of a dependency grammar are captured by grammatical constraints. These are binary relations on the set $|L|$ of lexical types which are interpreted in the following way:

if a word is classified by $L$ as being of type $t \in |L|$ and $t$ is related by a grammatical constraint to $t' \in |L|$ then the word must have a dependent of type $t'$.

Grammatical constraints can thus be thought of as requirements that a word places on its dependents in virtue of the way it is classified by the lexicon. These requirements must be satisfied by the dependency relation between words in any expression in which they occur.

It is not just lexically determined features that may enter into grammatical constraints. For example, in the sentence "Marta frightened the mole", the verb may require that one of its dependents has the feature 'nominative', a feature which is not assigned to any of the words by the lexicon. In a proper lexical analysis of the sentence, the feature must classify the word "Marta" if the requirements of the verb are to be met, even though "Marta" is not classified by this feature in the lexicon.

Definition A.26 A lexical analysis of a $W$-dependency structure, $\alpha$ in a $W$-lexicon $(W,F,L,\rho)$ is a function $l : |\alpha| \rightarrow |L|$ such that, for each $w \in |\alpha|$, $WF \subseteq \rho w$.

In other words, a lexical analysis of a dependency structure is a map assigning a lexical type to each word in the extent of the structure in a way that respects the lexicon. A lexical analysis can therefore assign types which have more but not less features than the types assigned by the lexicon. We can now say what it is for a dependency structure to satisfy a grammatical constraint with respect to some lexical analysis.
APPENDIX A. CLASSIFICATION SYSTEMS

Definition A.27 A \( W \)-dependency structure \( \alpha \) satisfies a grammatical constraint \( g \) with respect to a lexical analysis \( l \) if, for all \( w \in |\alpha| \), if \( (t, t') \in g \) and \( pt \subseteq plw \) then there is some \( w' \in |\alpha| \) such that \( (w, w') \in \alpha \) and \( pt' \subseteq plw' \).

Definition A.28 A dependency structure, \( \alpha \), is a solution to a grammatical constraint, \( g \), if there is some lexical analysis \( l \), with respect to which \( \alpha \) satisfies \( g \).

Proposition A.29 If a dependency structure, \( \alpha \), is a solution to grammatical constraints \( g \) then

1. if \( \alpha \) is also a solution to \( g' \) then \( \alpha \) is a solution to \( g \cup g' \), and
2. if \( g' \subseteq g \) then \( \alpha \) is a solution to \( g' \), and
3. if \( \alpha' \) is a dependency structure and \( \alpha \leq \alpha' \) then \( \alpha \) is a solution to \( g \).

Proposition A.29.3 says that if we have a solution and we add a new dependency link to it then the resulting structure is still a solution. Certain links will have been forced by the grammatical constraint, but others may be redundant. In other words, we may have described certain words as dependents of other words when there is no real dependency between them. We are most interested in finding solutions which lack this redundancy, i.e. solutions for which the removal of any link results in a structure which fails to satisfy the requirements of the constraint.

Definition A.30 A dependency structure, \( \alpha \), is a minimal solution to a grammatical constraint, \( g \), if for any solution, \( \alpha' \), to \( g \), if \( \alpha' \leq \alpha \) then \( \alpha' = \alpha \).

A dependency grammar is a set of grammatical constraints which are treated as options. The distinction between optional and obligatory constraints is very important in dependency theory. For example, although every verb needs a subject, it need not have an adverbal modifier. Although this is permitted, it is not necessary. We can say “a small mole burrowed” as well as “a small mole burrowed furiously”. In fact, we treat all the constraints in a dependency grammar as optional, but Proposition A.33 shows us how to account for obligatory constraints.

Definition A.31 A dependency grammar, \( G \), over a lexicon, \( (W, F, L, p) \), is a set of grammatical constraints over \( L \).
APPENDIX A. CLASSIFICATION SYSTEMS

Definition A.32 A dependency structure, $\alpha$, is a solution in a dependency grammar, $G$, iff it is a minimal solution to some $g$ in $G$.

Proposition A.33 If $G_{\text{oblig}}$ and $G_{\text{opt}}$ are sets of grammatical constraints then there is a dependency grammar $G$ such that $\alpha$ is a solution in $G$ iff there is some $X \subseteq G_2$ such that $\alpha$ is the smallest dependency structure which is a solution to all the constraints in $G_1$ and $X$.

Proof: Take $G = (\bigcup (G_1 \cup X) | X \subseteq G_2)$. The rest follows from Proposition A.29.

Example Simple Sample Dependency Grammar Let $|F| = \{d,n,nm,ac,pn,cm,a,v,t,i,av\}$ be the set of lexical features. For simplicity we will identify lexical type $l \in |L|$ with its associated set of features pl. Then $|L|$ is the set containing all subsets of the sets $\{n,nm,d\}$, $\{n,nm,pn\}$, $\{n,ac,d\}$, $\{n,ac,pn\}$, $\{n,cm\}$, $\{n,i\}$, $\{v,t\}$, $\{a\}$ and $\{av\}$. Together with the following table these define a lexicon (subject to an ability on our part to use the table to classify individual word tokens).

| she         | (pn,nm)     | small  | (a) |
| her         | (pn,ac)     | a      | (d) |
| Marta       | (n)         | burrowed | (v,i)|
| mole        | (cn)        | swallowed | (v)  |
| furiously   | (av)        |        |     |

To specify the dependency grammar $G$ we will use the construction of Proposition A.33. We use the notation $x,y,z \Rightarrow a,b,c$ to denote the pair $\{(x,y,z),(a,b,c)\}$.

$G_{\text{oblig}} = \{(v \Rightarrow n, nm), (v, i \Rightarrow n, ac), (d \Rightarrow cm)\}$

$G_{\text{opt}} = \{(v \Rightarrow av), (cn \Rightarrow a)\}$

The dependency structure depicted by Figure A.7 is a solution in $G$ since the lexical analysis

\[
a \quad (d,n,nm)
\]
APPENDIX A. CLASSIFICATION SYSTEMS

small {a}
mole {cn}
burrowed {v,i}
furiously {av}

respects the lexicon and ensures that the dependency structure satisfies each constraint in G_{abig} and both constraints in G_{opt}. It is a minimal solution since if we removed any of the links then one of the constraints would fail to be satisfied (in fact the relation would cease to give us a dependency structure since there would be two roots).

The above fragment is extremely simple, but it indicates the main principles of the classifier approach to dependency. Work in progress with Max Volino has shown that these principles can be applied to many other kinds of grammatical relation as well as some novel approaches to word order. There is also an implemented parser for grammars specified in this way.
Bibliography


208
BIBLIOGRAPHY


