Mass Terms and Plurals in Property Theory

Christopher James Fox

A thesis submitted for the degree of Ph.D.

Department of Computer Science
University of Essex

1993
Abstract

This thesis is concerned with extending a weak, axiomatic theory of Truth, Propositions, and Properties, with fine-grained intensionality (PT), to represent the semantics of natural language (NL) sentences involving plurals and mass terms.

The use of PT as a semantic theory for NL eases the problem of modelling the behaviours of these phenomena by removing the artificial burdens imposed by strongly typed, model-theoretic semantic theories. By deliberately using incomplete axioms, following the example set by basic PT, it is possible to remain uncommitted about: the existence of atomic mass terms; the existence of a ‘bottom element’ (a part of all terms) as a denotation of NL nominals; and the propositionhood (and hence truth of) sentences such as “the present King of France is bald”.

Landman’s theory concerning the representation of individuals under roles, or guises is reappraised in terms of property modifiers. This is used to offer a solution to the problem of distinguishing predication of equal fusions of distinct properties, and the control of distribution into mereological terms, when used to represent NL mass nominals, without assuming an homogeneous extension.

The final theory provides a uniform framework for representing sentences involving both mass and count nominals.
Acknowledgements

Thanks to Ray Turner for his excellent supervision of my research, and to Gennaro Chierchia, for his comments and suggestions on an earlier draft of this work. Anne De Roeck, Camilla Beswick and Richard Ball provided helpful suggestions on the contents of the thesis. Camilla Beswick also offered much support, encouragement and understanding over the years that this work has taken. Part of the research for this thesis was conducted with the aid of a Science and Engineering Research Council Studentship.

This thesis was typeset using Leslie Lamport’s $\LaTeX$. Extra symbols were provided by Jeremy Gibbons and Alan Jeffery’s St Mary’s Road symbol package.
# Contents

1 Introduction ........................................ 1  
   Structure of Thesis .............................. 6  

2 Background ...................................... 8  
   2.1 Plurals ........................................ 8  
   Distributivity and Collectivity .................. 8  
   Bottom, Non-denoting Terms, and Presupposition .. 14  
   Symmetry ......................................... 15  
   Associativity, Reciprocals and Structured Groups .. 15  
   Against Groups .................................. 18  
   Reducing Reciprocals ............................. 25  
   Transitive Verbs ................................ 29  
   Collectives and Intensionality ................. 30  
   Distribution, Scoping, and Logical Complexity ... 34  
   2.2 Link’s Treatment of Plurals ................. 39  
   2.3 Mass Terms ................................... 46  
   2.4 Ontology ...................................... 56  

3 My Approach ................................... 59  
   The Issues Addressed ............................ 59  
   Axiomatics ...................................... 59  
   Outline of Thesis ................................ 60  
   Non-denoting Terms .............................. 61  
   Distribution with Mass Terms ............... 62  
   A Unification .................................. 65  
   Other Issues ................................... 67  

4 Propositions, Properties, and Truth ............. 68  
   General Framework .............................. 69  
   Formal Theory .................................. 71  

5 Plural Terms in PT ............................... 74  
   Presupposition ................................. 85
6 Mereology in PT

7 Intensional Subjects
   Landman’s Formal Theory of Roles ........................................ 101
   Restricted Terms in PT ..................................................... 102
   Further Points ................................................................... 106
   Property Modifiers ............................................................ 108

8 The Final Theory
   Representations of Sentences ............................................... 115
   Types ............................................................................. 116
   Truth Conditions ............................................................... 119
   Property Modifiers ............................................................ 120
   Singulars and Plurals .......................................................... 124
   Examples and Further Axioms .............................................. 125

9 Conclusions
   Summary ........................................................................... 132
   Axiomatics ....................................................................... 133
   Implementation ................................................................... 134
   Further Work .................................................................... 134

10 Afterthoughts
   10.1 Measure and Amount .................................................... 136
   10.2 More on Existence Presuppositions ................................. 140
   10.3 Inheritance Hierarchies .................................................. 143

A A Model of PT with Boolean Terms
   A Model of the λ-Calculus with Summed Terms ....................... 149
   A Model of PT with Boolean Terms ....................................... 150

B A Grammar
   Basic Grammar Rules .......................................................... 156
   Grammar Rules for Ellipsis ................................................... 156
   Categorial Grammar ............................................................ 164
Chapter 1

Introduction

This thesis is concerned with representing the meanings of some sentences of English involving plural\footnote{In this thesis the word plural is often used in a technical sense intended to refer to syntactically singular and plural count terms, and to the representations of count terms.}, and mass terms. Intuitively, these are noun phrases which refer to countable individuals (pencils, or shoes, for example), and uncountable 'stuff' (mud, or footwear, for example) respectively. Some of the work presented here is also discussed in [Fox, to appear].

I am not concerned with the syntax of these expressions but rather their representation, or formal semantics. The aim of this programme can be crudely construed as attempting to devise a logic — a formal system of axioms (statements whose truth is to be assumed) and inference rules (a means of deriving new statements from existing statements) — with the entailment behaviour of natural language (NL). That is, if we provide a mapping from sentences of the fragment of NL under investigation to objects in the logic, then entailments between objects in the logic (performed by the use of inference rules and axioms) should match a native speaker’s intuitions concerning the entailments between the corresponding sentences in the fragment of NL. An additional constraint is provided by requiring a compositional mapping from sentences of the fragment to objects in the logic: the object that a sentence is mapped to must be composed of the objects that parts of the sentence are mapped to. If we take the mapping as attributing meaning to sentences, in some sense, then this can be phrased as the meaning of the whole depends upon the meaning of the parts.

Both count and mass nominals display certain, related, patterns of entailment. For example:

John died.
Mary died.

together entail the sentence:

John and Mary died.
CHAPTER 1. INTRODUCTION

In the mass domain:

Oil is liquid.
Water is liquid.

entail:

Oil and Water are liquid.

These inferences are both examples of cumulative behaviour. The complementary inferences are examples of distributive behaviour:

John and Mary died.

entails:

John died and Mary died.

In the mass domain we might have:

Water is liquid.

entail:

Hot water is liquid and cold water is liquid.

It would be convenient to be able formally to distinguish count and mass terms. For a preliminary indication of the distinction, the following, although potentially misleading, will suffice:

“Syntactically, mass nouns do not have a singular and plural form, they do not take number words, and take the determiners ‘much’, ‘little’; count nouns on the other hand, take the determiners ‘many’ and ‘few’ but resist a number of quantifiers such as ‘a great deal of’, ‘a little bit of’. Semantically, the basic difference is that count terms typically refer to discrete objects, while mass terms typically refer to entities conceived as continuous. These differences apply not only to bare plurals but also to more complex phrases like ‘orange juice’, ‘cream in your coffee’, ‘twentieth century Dutch poetry’, etc.” [Bunt, 1985].

Using examples from Bunt, the following four sentences seem acceptable:

---

2 Care should be taken with this last example. These may be regarded as generic statements: they may be interpreted as expressing a behaviour generally consented to, but permit counter-examples [Carlson, 1977]. Such a reading is the only one possible when considering the sentence: “all birds fly”. When attempting to capture the entailment behaviour of NL sentences, such generic interpretations are problematic. To simplify matters, the representations I give to sentences, when true, do not permit counter-examples.
CHAPTER 1. INTRODUCTION

There’s a lot of water in Holland.
How much luggage do you have?
Birdie eats 20 rice grains a day.
The second symphony is fantastic.

and the next four do not:

There’s a lot of canal in Holland.
How much suitcase do you have?
Birdie eats 20 rice a day.
The second music is fantastic.

suggesting that — as they occur in these sentences — ‘water’, ‘luggage’, ‘rice’ and ‘music’ are mass nouns, and that ‘canal’, ‘suitcase’, ‘grain’ and ‘symphony’ are count nouns (although the first sentence in the last group of four has a colloquial acceptability to some). It is hard to provide a reliable test as to which category a noun belongs to. In particular, nouns removed from the context of use cannot be simply classified: “generally speaking the count/mass distinction is not between words” [Bunt, 1985]. In this vein, Pelletier devised the notion of a Universal Grinder [Pelletier, 1975], which would take the object referred to by a count term n, and grind it up, thus creating a context in which it would be appropriate to use n as a mass term. Bunt also considers the converse of this machine, a Universal Sorter [Bunt, 1985] which packages the referent of a mass term m according to various properties the substance might have, colour, or composition for example, and thus allows m to be used as a count term. Note that mass terms can also be pluralised to refer to corresponding kinds.

Three muds were mixed.

can be uttered, meaning “three kinds of mud were mixed”.

The difficulty in providing a definitive test for plurals and mass terms may suggest that semantically there is little difference between them apart from count terms providing a ‘built-in’ means of individuation that appears lacking in mass terms [Quine, 1960]. Pelletier’s Universal Grinder removes this built-in individuation, and Bunt’s Universal Sorter adds it. This might indicate that, apart from this distinction, their semantics can be treated within the same framework. This is what I seek to achieve in the final theory of the thesis.

In many cases, existing treatments of formal semantics can be described as model-theoretic. Such an approach is exemplified by Montague Semantics [Dowty et al., 1981], where English sentences are translated into a formal language, and this formal language is, in turn, ‘interpreted’ by translating it into some set-theoretic construction, which provides a model of the language. The intermediate formal language is almost superfluous with this approach. When attempting to formalise the semantics of particular phenomena, the
CHAPTER 1. INTRODUCTION

model-theoretic approach is to find some mathematical object which seems
to have the desired behaviour, and add this to the model theory. The inter-
mediate formal language inherits the behaviour of the model by insisting that
every inference in the model is a valid inference in the formal language. This
is known as completeness. Further, to avoid inconsistency, every inference in
the language must be valid in the model. Model-theoretic semantics typically
provide no proof theory, thus it is not possible to demonstrate, within the
representation, that the semantics has the appropriate entailment behaviour.
It might be argued that this approach to semantics embodies a confusion be-
tween obtaining the appropriate entailment behaviour, and demonstrating
soundness.

In contrast, the approach to formal semantics in this thesis is axiomatic.
The aim of axiomatic semantics is to translate NL into a formal language,
as before, and then offer rules that appear to produce the intuitively correct
entailment behaviour directly in the chosen representation. Models may be
provided to show that there are mathematical objects which satisfy all of the
axioms. This demonstrates that the axioms are consistent (provided that
mathematics is consistent). However, since the behaviour is given by the
axioms, not the model, there is no need to insist that the representation is
complete with respect to the chosen model: inferences are only valid if they
follow from the axioms, not if they obtain from some artificial mathematical
construction, whose existence is independent of the phenomena in question.
This gives the flexibility to formulate axioms that are incomplete with respect
to certain issues, should there be no evidence for the existence of a ‘correct’
treatment. It also avoids the tendency to reduce concepts to perhaps inap-
propriate mathematical notions. For example, properties can seem to behave like
sets. Despite evidence to the contrary, a model-theoretic approach encour-
gages the view that they actually are sets, rather than primitive objects that
can be given an axiomatic characterisation. Model-theoretic formalisations
that use set theory are, at some level, extensional theories: a set is defined by
its members (its extension). This has undesirable consequences: such theo-
ries are generally too strong for the semantics of NL (see [Chierchia, 1982a,
Chierchia, 1984, Chierchia and Turner, 1988] for example). For this reason, I
use an essentially intensional theory — property theory — where propositions
and properties are taken to be primitive notions, not necessarily reducible
to some set-theoretic characterisation. Briefly, this allows two properties to
have the same members — and two propositions always to be true together
— without being equated. It should be stressed, however, that the adopted
approach is nevertheless model-theoretic in the sense that the semantics is
given in terms of properties and propositions and other ontologically prim-
itive notions, rather than sets. In the end, set theory is itself an axiomatic
theory, but it is not clear that its ontology is directly appropriate for NL
semantics.

Existing set-theoretic treatments of mass terms sometimes assume an
CHAPTER 1. INTRODUCTION

extensional ontology: they take the referents of NL nominals to be equated with the sum of their physical parts. This can be construed as a form of nominalism, which denies the existence of primitive Platonic objects — terms independent of material manifestation. Such a nominalistic ontology causes philosophical and logical problems. As an example of a logical problem, we might wish to have distributive properties — properties that when applied to an object also apply to the parts of that object — to explain the distributive entailments exemplified above. With an extensional ontology where terms are equated with their physical extension, this leads to inappropriate inferences as such properties will distribute to all the physical parts of a term.

The philosophical problems of nominalism are exemplified by the metaphysics of change: an object may survive a particular physical make-up. Even if a more intensional, or Platonic, ontology is adopted to avoid some of the philosophical problems, the logical problems may remain: there are cases where we might consider Platonic objects to be the sum of their Platonic parts. Muddy water might have mud as a part. If we take “is water” to be distributive, the sentence “Muddy water is water” would, on the construal of some existing theories, entail “Some mud is water”. This could be avoided by choosing an even more intensional ontology — forcing muddy water to only have muddy, watery parts, in the above example — this is at odds with any effort to maintain a distinction between philosophical and essentially semantic choices.\footnote{However, if it were the case that no formal representation could be devised for particular phenomena, without taking a particular ontological commitment, then this would be evidence in favour of a certain philosophical choice.} This suggests that for the formal semantics of NL, a more intensional notion of distribution is required.

In this thesis, I develop an axiomatic theory that remains incomplete with respect to certain controversial issues, such as the denotation of the so-called ‘non-denoting’ definite descriptors (like “the present king of France”) and the status of sentences involving them (such as “the present king of France is bald”), as well as the existence of minimal or atomic mass terms. This is sympathetic to the axiomatic version of property theory I shall use (PT) — an axiomatisation of Aczel’s Frege Structures [Aczel, 1980] due to Turner [Turner, 1992]\footnote{This is similar to the slightly stronger theory TP [Turner, 1990, Chapter 5]. Later, when considering presuppositions, it appears that the stronger theory may be required.} — and captures the undecided nature of our intuitions on such points. Further, the weak typing of PT allows a certain simplicity in the formal treatment which is not available in more extensional, set-theoretic formalisms, as there is no necessity to resort to type-shifting rules within the semantics. By using property theory, it should be possible to integrate the treatment of plurals and mass terms with the representation of intensional phenomena in natural language, such as generics and dispositional attitudes. The weak treatment of non-denoting terms is compatible with a treatment both of presuppositions and of helpful answers to questions.
CHAPTER 1. INTRODUCTION

To control distribution with mass terms, I develop a representation where the relevant property — which the parts that are distributed to must have — is effectively recorded. For this I make use of some of Landman’s insights regarding individuals under a role or guise [Landman, 1989]. I reappraise his theory in terms of property modifiers, which is compatible with the representation of NL modifiers such as adjectives and adverbs, and avoids positing extra ‘intensional’ individuals.

Structure of Thesis

The background literature is discussed in Chapter 2. Issues concerning plurals (count terms) and mass terms are presented. Some ontological points are also raised. Following the discussion of natural language plurals, one of Link’s formal theories, concerning the representation of plurals, is presented in some detail. This is in order to give direct motivation for the semantic theories developed in subsequent chapters, and to present some of the formal issues without the notational distractions of the intensional theory that I adopt.

I give an outline of my approach to the representation of plurals and mass terms in Chapter 3.

The basic theory of properties and truth I shall use, PT, is presented in Chapter 4. In Chapter 5 I add operators and axioms to PT so that it embodies some of the insights of Link’s theory. The plurals are axiomatised in a way that avoids the problems of non-denoting definite descriptors. At the end of this chapter, an indication is given as to how this may be compatible with a formalisation of presuppositions.

The axioms for plurals in PT can be modified to give a potentially atomless part-whole structure amongst some of the terms. This can be used for an extensional treatment of mass terms, or to characterise mereological stuff. Such a theory is formalised in Chapter 6. Some ontological issues are also discussed here. To explain certain inferences, typified by distribution, this theory is not really adequate.

A model of PT extended with plural/mereological terms is given in Appendix A, to demonstrate the soundness of these two systems.

To allow choices of ontology where substances may be part of one another (so the result of predicating a distributive property of a term depends upon how that term is described) and where count individuals may possess different properties, again depending upon how they are described, it is desirable to add some intensionality to the representation. In Chapter 7, Landman’s theory concerning individuals under different roles, or guises, is introduced. This is reappraised in terms of property modifiers. A more direct PT implementation of Landman’s ideas is also given.

The final theory is presented in Chapter 8. This gives a unification of the mass and count domains, and makes use of existing notions in theories
of adjectives to give property modifiers a useful behaviour. Some examples of the application of the theory are described.

Chapter 10 presents some remaining points and issues concerning measures and amounts; existence presuppositions with definite descriptors; and speculation about a possible connection between property inheritance and Landman’s roles.

A context-free grammar is given in Appendix B. It is presented to demonstrate the compositionality of the final theory by producing the semantic expressions used in the examples of Chapter 8.
Chapter 2

Background

2.1 Plurals

Standard interpretations of NL nominal expressions (e.g. Montague’s [Dowty et al., 1981]) have them denoting objects in a domain of ordinary individuals like “John” and “chair42”. This is not generally sufficient to cope with plurals and mass terms.

Distributivity and Collectivity

Some sentences with plural and conjoined nouns can be represented with ordinary individuals. For example, the verb “die” is said to be distributive. From the sentence “John and Mary die”, it is possible to infer that “John dies and Mary dies”, effectively distributing the verb over the conjunction. Similarly with syntactic plurals, when we say “the boys died”, we interpret this distributively, so that it is roughly equivalent to “every boy died”. As it is possible to infer “John and Mary die” from “John dies and Mary dies”, we can also say that the verb “die” is cumulative. These entailments allow such examples to be represented with ordinary (singular) individuals. However, there are some cases where such equivalences do not hold. For example, the verb “meet” is not distributive. Given the sentence “John and Mary met”, it is not legitimate to say “John met and Mary met”. The verb “meet” is said to be collective. With syntactic plurals, as in “the boys met”, it is usual to interpret this as a collective activity (or perhaps a collection of collective activities). Similarly, if “John and Mary bought a boat together” the thing that bought the boat is the collection or sum “John and Mary”. To represent such collective readings, the terms of a formal representation language must be enriched to represent plural objects.

We can show that collective verbs, like “meet”, are cumulative: from “the boys met and the girls met”, we may infer “the girls and the boys met”. The conclusion is weaker than the antecedents — more circumstances may satisfy the final sentence than satisfy the first two. Cumulativeness is a property of
many, but not all, verb phrases. According to Schwarzschild [Schwarzschild, 1990], if we have:

The students in Mr. T’s shop class are all of the same sex.
The students in Miss Murphy’s home economics class are all of the same sex.

then we cannot infer

The students in Mr. T’s shop class and the students in Miss Murphy’s home economics class are all of the same sex.1

Given that the theory seems to need collections, one objective of a compositional treatment of these phenomena should be to explain the relationship between collective and distributive interpretations of sentences. Distribution should not only have its occurrence restricted to certain properties, but it should also be to an appropriate level: if John danced, it does not imply that John’s intestines danced. This inference is normally blocked by assuming that individuals, like John, are atomic, and independent of their physical parts, at least with respect to distributive inferences.

Typically, theories concerning the representation of plurals allow individuals to be joined to create new plural objects, or sums. So, “John and Mary”, in the representation, form a new plural object that is the sum of the representations of “John” and “Mary”. This sum could then be the argument of the representation of a verb such as “meet” or “die”. Most of the literature uses entirely extensional, set-theoretic languages for the representation of natural language count terms.

Russell had the first proposal for the extension of first-order logic to second and higher order logics [Russell, 1903]. These were introduced into the formal semantic analysis of NL to cope with the intuitions behind “and” by Bartsch and Bennett [Bartsch, 1973, Bennett, 1975]. They amend the extensional treatment of semantics so that terms may be individuals, as before, or sets. Thus if the “meaning” (denotation) of the word boys is the set of $a, b, c$, that is:

$$[\text{boys}] = \{a, b, c\}$$

then:

The boys bought a boat.

1I find the conclusion acceptable if “are all of the same sex” is interpreted as distributing over the conjunction. The strong sense of non-cumulativity in this example might be attributed to the presence of an implicit existential quantifier, introduced by the phrase “the same sex”, that outscopes the subject noun phrase. That is, we might interpret “N are all of the same sex” as entailing: “there is a sex which all of N have”.
is translated into:

$$\exists x (\text{boat}'(x) \& \text{bought}'(\{a, b, c\}, x))$$

which can be paraphrased as: “there is a boat, and \(a, b, c\) collectively bought it”.

This has problems, many words must now be multiply ambiguous depending on whether individuals, or sets may appear in a particular argument position. There would be four representations for “bought”, but “who bought the boat?” does not seem intuitively to be ambiguous.

Subsequent authors have tried to improve this approach by making singular and plural terms (both lexical, and structural plurals) objects of the same type. It is possible to “elevate” the representation of singular terms to singleton sets [Landman, 1989, Scha, 1981]. In his PhD thesis, Schwarzschild takes an alternative route of equating singleton sets with their members, so \(a = \{a\}\). This he calls Quine’s initiative [Schwarzschild, 1990]. Such a structure of sets and singleton sets models a complete atomic Boolean algebra (see [Halmos, 1963] for example). Conjunction between definite noun phrases can then be modelled by the \textit{met} operator in the Boolean algebra. This operator allows the sum of two terms to be created. The definite descriptor can be interpreted by the supremum operator\(^2\) of the Boolean algebra as used by Link for example [Link, 1983]. This is summarised in the following table:

<table>
<thead>
<tr>
<th>Natural Language</th>
<th>Representation Language</th>
<th>Set-theoretic Interpretation</th>
<th>Lattice-theoretic Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>(j')</td>
<td>({j'})</td>
<td>(j')</td>
</tr>
<tr>
<td>John and Mary</td>
<td>(j' \oplus m')</td>
<td>({j'} \cup {m'})</td>
<td>(j' \sqcup m')</td>
</tr>
<tr>
<td>the men</td>
<td>(\sigma x \text{men}'x)</td>
<td>({x : x \in \text{men}'})</td>
<td>(\sqcup_{x \in \text{men}'x})</td>
</tr>
</tbody>
</table>

The sentence:

John and Mary met.

would be represented as:

$$\text{met}'(j' \oplus m')$$

which can then be given either a set or lattice-theoretic interpretation.

It might seem desirable to represent the sentence:

John and Mary died.

\(^2\)The definite descriptor could be modelled by an operator that picks the greatest element in the extension of the predicate. Singular nouns can be represented by predicates that range over atoms in the Boolean algebra, and plural nouns by predicates that range over the elements of a Boolean algebra (generated by the atoms in the singular property). If a predicate has no extension, there is no greatest element, and if there is more than one element in the denotation of a singular predicate \(p\), then ‘the \(p\)’ is badly formed, as again there is no greatest element [Scha, 1981].
directly as:

died (j') \& died (m')

The distinction between the collective and distributive readings would then effectively reside in the scope of the conjunction. With the collective reading, conjunction would have narrow scope (it is within the scope of the representation of the verb phrase), and in the distributive example, it would have wide scope (outscoping the representation of the verb phrase). With this approach, it is sometimes said that the representation of the distributively read noun phrase is lifted to produce the collective form. This is equivalent to taking distribution over conjoined noun phrases to be a variety of ellipsis, where:

John and Mary died.

is analysed as a reduced form of:

John died and Mary died.

Ellipsis occurs with other categories, as in the noun phrase:

the man and woman

where the determiner distributes over the conjunction.\(^3\) Distributive inferences with noun phrases are more general than this, as they can involve syntactic plurals, as in:

The men died.

The scoping ambiguity could be generalised to cover the implicit conjunction present in syntactic plurals [Lakoff, 1970].\(^4\) Representing the collective/distributive distinction in this way results in collectively read noun phrases always having narrow scope. With examples that only use intransitive verbs, this might not seem problematic. However, with the following sentence:

\(^3\)This ellipsis can be treated syntactically: in categorial grammars, these scoping effects are accounted for using type-lifting (or type-raising) [Dowty, 1988]. Perhaps this should be distinguished from type-lifting used merely as a means of addressing problems that arise in strongly typed semantic theories, where a term cannot belong to two different types.

\(^4\) On a quantifier scoping account of the collective/distributive distinction, we might represent the collective interpretation by having the verb phrase outscope (the intension of) the noun phrase. The distributive interpretation would arise when the verb phrase has narrow scope. For example: \(\exists x (\text{men}'x \& \text{die}'x)\) would represent “some men died”, whereas the sentence “some men met” would be represented with: \(\text{met}'(\lambda y \exists x (\text{men}'x \& yx))\). With a strongly typed semantic theory, the types of the verb phrase, and/or noun phrase would have to be altered, so that the representation of the noun phrase can appear as an argument of the verb phrase. The representation of the noun phrase in the distributive representation is said to be ‘lifted’ to produce the collective representation.
Five insurance associations gave a $25 donation to several charities.

It is possible to obtain an interpretation where “five insurance associations” is given a collective interpretation, which outscopes a distributive interpretation of “several charities”, so that each charity received $25, but there are only five associations [Roberts, 1988]. This suggests that the distributive v. collective distinction should not, in general, be equated with a wide v. narrow scope interpretation, and that collections of individuals are things in themselves, and not a simple artifact of scoping. It may still be legitimate to treat distribution over conjunction as part of a theory of ellipsis, rather than exclusively in a semantic theory of plurals and mass terms.\(^5\) We shall return to the use of scoping and lifting later in this chapter, in the sections “Associativity, Reciprocals and Structured Groups”, and “Distribution, Scoping and Logical Complexity”.\(^6\)

If the difference between collectively and distributively read sentences is not, in general, a scoping effect — residing in the relationship between the interpretation of the verb phrase and noun phrase — then the question arises as to whether the distinction resides in the semantics of the verb phrase, or that of the subject noun. If we follow Link’s account [Link, 1984, Link, 1991a], the distinction is in the predicate representing the verb phrase.\(^7\)

According to this view, some predicates are always distributive, such as died. The sentence:

\[
\text{John and Mary died.}
\]

could be interpreted as:

\[
died' (j' \oplus m')
\]

As “died” is taken to be distributive over the conjoined noun phrase, then it holds of the conjuncts, and it is possible to infer that:

\[
died' (j') \& died' (m')
\]

For predicates that can be read ambiguously as either collective, or distributive, an operator D is optionally used to mark a predicate as being read distributively. So:

---

\(^5\) The grammar given in Appendix B treats distribution of intransitive verbs across conjoined noun phrases as a scoping effect in the grammar, by taking it to be a form of ellipsis. This complements the final theory (Chapter 8) which, to avoid combinatorial complexity, does not cover distributive inferences over conjoined noun phrases.

\(^6\) If collections of individuals are not scoping artifacts, it should, at least in principle, be possible to quantify over them. However, universal quantification over arbitrary collections has an impact on the logical complexity of the representation.

\(^7\) Later, I shall discuss theories where there is an ambiguity in the noun phrase, and examples that suggest the ambiguity must be in the verb phrase.
CHAPTER 2. BACKGROUND

John and Mary bought a boat.

could be represented as one of:

\[ \text{bab}'(j' \oplus m') \]
\[ \text{D} \text{bab}'(j' \oplus m') \]

where from the latter, it is possible to infer that bab' holds of the individuals, or atoms of the sum, thus entailing:

\[ \text{bab}'(j') \land \text{bab}'(m') \]

If the inference process is run the other way, so that from:

\[ p(x) \land p(y) \]

we infer:

\[ p(x \oplus y) \]

then the verb (or predicate) is said to be cumulative. Sometimes an operator distinct from \( \text{D} \) — * for example [Lønning, 1989, Link, 1991b] — is used to indicate that the predication arose out of cumulative behaviour. Given:

\[ \text{bab}'(b') \land \text{bab}'(j' \oplus m') \]

we may infer:

\[ ^* \text{bab}'(b' \oplus j' \oplus m') \]

which is indeterminate as to which parts the unadorned property applies.

Verbs (or predicates) which are both fully distributive and cumulative (like “died”) are said to be homogeneous: their cumulative and distributive forms may be equated.

Most of the treatments follow broadly similar lines. Massey applied the Calculus of Individuals [Leonard and Goodman, 1940] to the problem [Massey, 1976]. This is an axiomatisation of mereology — part and whole structures, especially physical — where there is nothing that is a part of everything, in opposition to set theory, where the empty set is a part of all sets. Bunt developed ensemble theory in which to represent plurals and mass terms [Bunt, 1985]. It appears to be a variant of mereology, except there is a bottom element (a term that is a part of all other terms). He had discrete ensembles to represent plurals, and continuous ensembles for mass terms. His work is discussed in more detail in §2.3, the section on mass terms. Blau used a primitive notion of collections given in terms of a semantic analogue of the NL definite descriptor [Blau, 1981]. He criticises Massey’s use of Leonard and Goodman’s mereology as too extensional since objects are equated with their physical parts. He also criticises Bartsch’s re-appraisal of this [Bartsch, 1973] — where the part-whole structure is not that of physical manifestation
— because the notion of ‘part’ does not have an analogue in NL. The ‘part of’ relation need not be taken as primitive, however.

In one paper Link uses a complete join semi-lattice (actually the same as a complete lattice) for plurals [Link, 1987]; in others, he gives models equivalent to complete atomic Boolean algebras [Link, 1983, Link, 1984, Link, 1991b, Link, 1991a].

The strength of a complete atomic Boolean algebra prevents perverse results allowed in a weaker lattice, where, for example, an individual may be part of a conjunction, without being a part of either of the conjuncts [Link, 1991a].

Lönning sought to provide a unified framework in which to compare these approaches [Lönning, 1989]. He takes the “standard” approach to be typified by [Leonard and Goodman, 1940, Massey, 1976, Link, 1983, Link, 1987, Link, 1991b, Schä, 1981]. These all assume the associativity, symmetry and idempotence of conjunction. This, as Lönning points out, is not universally accepted. In addition, there is some dispute about whether there should be a bottom element in the representation. These issues are discussed below.

**Bottom, Non-denoting Terms, and Presupposition**

In a Russellian treatment, we could say that ‘non-denoting’ definite descriptors, like “the present king of France” denote bottom — indeed, this is the natural consequence of adopting a Boolean algebraic model — and predication of the bottom element leads to falsity. This can be viewed as a simple formal treatment of presupposition with definite descriptors: any true predication of a term presupposes that the term exists, if the term does not exist, then any (atomic) predication of it cannot be true.

Lönning suggests the possibility of assigning false to any sentence containing a predication of a non-existent term, so both “the present king of France is bald” and “the present king of France is not bald” are false sentences [Lönning, 1989]. This is a contradiction in classical logic, as it can be seen to involve the assertion $A \& \sim A$. Such a contradiction could be avoided if we reject bivalence, and say that these sentences are not true, as opposed to false. This can capture the idea that a term’s existence is presupposed by both positive and negative attributions of properties. Blau adopts a non-classical logic which explicitly rejects bivalence by adopting three truth values, the value ‘unknown’ being assigned to predication of non-existent terms [Blau, 1981].

An alternative is to adopt a free logic, where such terms ‘denote elsewhere’ and are thus not available for predication within the theory [van Fraassen, 1966, Link, 1991a]. Massey, Bunt and Link seem to think there should not be a bottom element [Massey, 1976, Bunt, 1985, Link, 1991a] which has the

---

8 Note that there was a delay of many years between the writing of [Link, 1991b] and its publication. This paper is often referred to by others as [Link, to appear].
effect of excluding non-denoting terms from the theory. The issues concerning bottom; non-denoting terms; and the logical aspects of presuppositions with definite descriptors are taken up when I outline my approach in Chapter 3 (in the section Non-denoting Terms) and in Chapter 5 where I give an account of non-denoting terms, which is potentially compatible with a formal theory of presupposition.9

Symmetry

There are examples where it appears that the summation operator cannot be symmetric. In English, we can easily say:

John and Mary are husband and wife (respectively).

It sounds rather odd, however, to say:

Mary and John are husband and wife (respectively).

Schwarzschild suggests that this asymmetry can be accounted for in the dynamic component of meaning [Schwarzschild, 1990, Schwarzschild, 1992]. That is, the term denoted by the conjoined noun phrase is the same, irrespective of the order of the conjuncts, but the linear order has an effect on the discourse structure: for example the terms “Mary” and “John” become available for anaphoric reference, and the order in which they come available is recorded in some representation of the discourse structure [Kamp, 1981, Heim, 1982]. The “respectively” construction can then be interpreted at some meta-discourse level.

Associativity, Reciprocals and Structured Groups

There are readings of sentences which call into question the associativity of the conjunction of nouns. Some verbs can be ambiguous between distributive and collective interpretations. From the sentence “John and Mary bought a boat” we might conclude either that they bought a boat together, or separately. When there are more than two members of a collection, then there may be further readings, depending upon how the part of the collection are grouped. From the sentence:

John, Mary, Peter and Sharon bought a boat.

it is not clear how many boats were bought. There could be as many as four (the distributive reading), or as few as one (the collective reading), or some number in between. These other readings may be termed intermediate distributive readings. When a distributive inference is performed in these

9I do not seek to give a linguistic appraisal of presupposition, nor a complete formalisation.
cases, the result is some collective readings of sub-collections of the original noun phrase. It may seem desirable, initially, to allow the noun phrase to indicate which are the relevant sub-collections by adding some structure to its representation.

Intermediate distribution seems to be related to the representation of reciprocal sentences. Reciprocal readings occur with predicates that require a collection of at least two individuals — the subject NP must allow for plural number agreement — such as “meet” when used as an intransitive verb. Such sentences often involve explicitly reflexive properties: properties modified with a reflexive pronoun such as “each other”. If structure is added to the noun phrase for intermediate distribution — in order to indicate the relevant objects to which a property should distribute — then perhaps a similar structure on the noun phrase could be used to indicate the relevant sub-collections involved in reciprocal sentences, the difference being that the reciprocal property does not distribute down to these collections (unless they are collections of structured collections), but instead indicates some relationship between them.

Lønning thinks that reciprocals are the best starting point for considering structured noun phrases. If we take the reciprocal sentence:

\[(\text{Wellington and Blüchner}) \text{ and Napoleon fought against each other at Waterloo.}\]

only the indicated ‘bracketing’ of the noun phrase (NP) matches the historical facts [Hoeksema, 1983]. The representations obtained by this explicit bracketing of the noun phrase in the semantics of sentences involving reciprocals, are termed structured group readings. Lønning considers it an integral part of considering structured groups to explain why some sentences, when given a collective interpretation, cannot have such readings, as in:

The Swedish Players and the Dutch Players carried the piano upstairs.

If this sentence is interpreted collectively, then any structure added to the representation of the noun phrase — to block distribution — is irrelevant, as is any structure added to indicate the appropriate sub-collections for reciprocal readings. The latter structures would be relevant only for predicates that cannot be ascribed to individuals as well as collections [Lønning, 1989]. Lønning suggests that there may be no need for structured collections: it might be possible to derive these readings when mapping from syntax to semantics [Lønning, 1987a]. In the general case, the intuitive meaning of reciprocals is often unclear, and does not lend itself to some easy reduction. Later I shall examine evidence that adding structure to the representation of the subject noun phrase does not account for all the data, and I shall explore some of the problems encountered in attempting a reductive analysis of reciprocal sentences.
CHAPTER 2. BACKGROUND

Typically, solutions to the problem of representing structured groups effectively involve the use of non-associative brackets, or non-associative conjunction — which has the effect of introducing non-associative brackets — sometimes together with an associative conjunction (although this can introduce an ambiguity in the interpretation when there is no intuitive justification). Various means of obtaining the effect of a non-associative conjunction have been tried. Hoeksema proposed a non-associative conjunction [Hoeksema, 1983]. However, he did not consider plural definite descriptors nor a distributive operator. Link obtained structured and intermediate distributive readings by taking conjunction to be ambiguous between lattice join, and a structured group operator [Link, 1984]. Thus:

\[ [A \text{ and } B] \]

becomes either:

\[ \langle A \rangle \oplus \langle B \rangle \]

or:

\[ A \oplus B. \]

In that paper, Link says his conjunction is ambiguous between \( x \oplus y \) and \( \langle x \rangle \oplus \langle y \rangle \). Others have taken the definite descriptor as ambiguous between \( x \) and \( \langle x \rangle \) [Hoeksema, 1987, Landman, 1987, Landman, 1989, Roberts, 1987]. But the definite descriptor is not usually taken to be ambiguous when not conjoined — except for [Landman, 1989] — as over-generation can occur [Hoeksema, 1987]. Distribution also goes wrong with an unambiguous \( \oplus \): distribution across \( \langle x \rangle \oplus \langle y \rangle \) reaches the terms \( \langle x \rangle \) and \( \langle y \rangle \) as pointed out by [Landman, 1989], whereas it is usually required to reach \( x \) and \( y \).

The new conjunction is n-ary, only goes to one level, and always has wider scope than standard sum. Such a theory still over-generates. Taking the previous sentence:

The Swedish players and the Dutch players carried the piano upstairs.

has a collective reading, a distributive reading, and an intermediate distributive reading. Yet Link’s theory also produces an irrelevant structured group reading [Lønnig, 1989]. If we take \( P \) to be the representation of “carried the piano upstairs”; and \( a, b \) are the Swedish players and \( c, d \) are the Dutch players, then the useful readings produced are:

- Collective: \( P(a \oplus b \oplus c \oplus d) \)
- Distributive: \( ^D P(a \oplus b \oplus c \oplus d) \)
- Intermediate distributive: \( ^D P(\langle a \oplus b \rangle \oplus \langle c \oplus d \rangle) \)

where:

\[ ^D P(a \oplus b) \leftrightarrow P(a) \& P(b) \]
However, the theory additionally produces a structured-group collective reading which is not relevant for “lift”:

\[
P((a \oplus b) \oplus (c \oplus d))
\]

If we accept the the need for many levels of conjunction, the question of how many there are in English, for example, remains. In one of his papers, Link is confident there are only two [Link, 1984]. Lønning thinks there may be up to three (possibly more) [Lønning, 1989]. He does not like non-associative “and”, as used by [Link, 1984, Hoeksema, 1987], because other uses of “and” in English appear to be associative, and Lønning would rather keep a uniform interpretation.

Lønning considers using type-lifting to achieve the intermediate distributive readings [Lønning, 1989]. Schwarzschild also explores type-lifting to capture structured group readings in reciprocal sentences [Schwarzschild, 1990, Schwarzschild, 1992]. His arguments against their use in reciprocal sentences — and against the need for such structured representations — can be taken as arguments against attempting to represent the intermediate distributive readings. Following [Partee and Rooth, 1983], Lønning argues that the ambiguity is not between two collective conjunctions, but between collective term conjunction, and conjunction of terms ‘lifted’ to quantifiers, in which case conjunction is interpreted as set intersection. The claim is that lifting is independently motivated by such phrases as:

John and every other man.

where it is assumed that “John” is of type \(e\), and “every other man” is of type \(\langle e, \langle e, t \rangle \rangle\) and the objects must be of the same type to be conjoined. Thus, “John” must be lifted to type \(\langle e, \langle e, t \rangle \rangle\). Lifting occurs only if necessary, and conjunction occurs at the lowest possible level, except that definite NPs may be lifted anywhere to achieve the distributive/collective ambiguity. This account is similar to the treatment of ellipsis in categorial grammars [Dowty, 1988].

Lønning combines this with the distribution operator to obtain intermediate distributive readings. He argues that this explains why there are two sorts of conjunction: there is term conjunction and set intersection. Both of these conjunctions are associative.

Against Groups

Lønning’s type-lifting solution, together with Link’s and Hoeksema’s theories [Lønning, 1989, Link, 1984, Hoeksema, 1987] cannot easily cope with examples where there is a conjoined VP and the NP is to be interpreted distributively with respect to one VP and collectively with respect to the other:
The boys and the girls had to sleep in different dorms, met in the morning at breakfast, and were then wearing their blue uniforms.

Lønning expresses some doubts about the examples of this presented in [Landman, 1987], because some of them require the presence of modifiers such as "different" which have never been investigated in detail. However, Schwarzschild gives some examples that do not use "different":

Ray and Tess awoke early and met in the ballroom.

which is equivalent to:

Ray and Tess awoke early.
Ray and Tess met in the ballroom.

where the first sentence is read distributively, and the second is read collectively [Schwarzschild, 1990]. This might suggest that there is an ambiguity in the the verb phrase as well as the noun phrase, so that the collective verb phrase can be 'lifted' to accept a structured noun phrase, and so ignore that structure. I shall comment on Schwarzschild's arguments against this theory below.

These attempts at obtaining structured group readings and intermediate distribution by positing an ambiguity in conjunction fail to account for sentences such as:

Rodgers, Hammerstein and Hart wrote musicals.

where the intended interpretation is:

Rodgers and Hammerstein wrote musicals together, and Rodgers and Hart wrote musicals together.

The formal reason is that an ambiguous conjunction can only account for ambiguity amongst partitions\footnote{In set theoretic terms, a partition of a set $S$ is a set $P$ of non-empty subsets of $S$ where members of $P$ are disjoint, and the union of the members of $P$ is $S$. The partitions of $\{a, b, c\}$ are $\{\{a, b, c\}\}; \{\{a, b\}, \{c\}\}; \{\{a, c\}, \{b\}\}; \{\{a\}, \{b\}, \{c\}\}.$} of the collections of individuals, whereas the correct reading here requires the entire noun phrase to be taken as ambiguous between the possible covers\footnote{Set-theoretically, a cover is a more general than a partition: there is no requirement for members of a cover to be disjoint. The covers of $\{a, b, c\}$ will include its partitions, and $\{\{a, b\}, \{b, c\}\}; \{\{a, c\}, \{b, c\}\}; \ldots; \{\{a\}, \{b\}, \{a, b, c\}\}; \ldots$ for example.} or more accurately, minimal covers\footnote{A minimal cover $M$ of a set $S$ is a smallest set of non-empty subsets of $S$ which cover it. For example $\{\{a, b\}, \{b, c\}\}$ and $\{\{a, b, c\}\}$ are both minimal covers, but sets such as $\{\{a, b\}, \{b, c\}, \{a, c\}\}$ and $\{\{a, b, c\}, \{b\}\}$, although covers, are not minimal covers as they each have proper subsets that are covers of the original set.} of the collection of individuals referred to [Gillon, 1992]. Lasensohn rejects the
existence of intermediate distributive interpretations arguing that a meaning postulate concerning the denotation of “write” can account for the above example. He attempts to show that the intermediate distribution does not occur for other verbs, like “earn” [Lasersohn, 1989]. Gillon constructs an example of intermediate distribution using the verb “earn”, effectively showing that intermediate distribution cannot be accounted for with meaning-postulates for particular verbs [Gillon, 1990].

Consider the representations of reciprocals: an ambiguous conjunction cannot account for the intended reading of:

Despite their current membership of the common market, only
45 years ago, Germany and England and France and Italy were
fighting against each other in one of the worst wars in history.

The alignment of forces cannot be captured by any bracketing of the subject noun phrase [Schwarzschild, 1992].

Link and Lønning have suggested that intermediate distributive readings need not be made explicit in the representations of NL sentences. Thus, we may take some sentences to be ambiguous between distributive and collective readings, and the intermediate readings are subsumed by the collective reading. The collective representation would then be ambivalent as to the possible intermediate distributive readings. This view has support from the fact that it is possible to disambiguate the distributive and collective readings, with words like “together” or “each”, but there is no convenient means of marking a sentence so that a particular intermediate distributive interpretation is favoured. The structured group readings of reciprocal sentences would similarly be subsumed by a collective representation.

Schwarzschild also considers intermediate distribution, and presents arguments against elaborating the structure of the noun phrase in the compositional representation of such sentences. I shall present his ideas in some detail as it indicates why I choose not to represent intermediate distribution explicitly in the theory I develop later. He uses the following example to show that intermediate distributive readings exist, and that the exact nature of the reduction of such readings is context dependent:

“Imagine a situation in which two merchants are attempting to
price some vegetables. The vegetables come in various varieties
and they are piled in baskets. To determine their price, the vege-
tables need to be weighed. Unfortunately, our merchants do not
have the appropriate scale. Their grey retail scale is very fine and

---

13 They presented this view at the Third European Summer School for Language, Logic and Information at Saarbrücken, 1991, and attributed it to Krifka and Schwarzschild. Schwarzschild has suggested that all verbs are read distributively with respect to some context dependent partition, or cover of the noun phrase [Schwarzschild, 1990]. I shall elaborate on this view below.
is only meant to weigh a few vegetables at a time. Their black
scale is coarse, meant to weigh small truckloads. Realising this
one of the merchants truthfully says:

The vegetables are too heavy for the grey scale
and too light for the black scale.

In order to save space in our explanation, let us reword his utter-
ance:

a) The vegetables are too heavy for the grey scale.
b) The vegetables are too light for the black scale.

[The sentence (a)] is false on its distributive reading . . . It is true
on its collective reading but that is not what the merchant in-
tended. [The sentence (b)] is false on the collective reading . . .
It is true on its distributive reading, but again that is not what
the merchant intended to say. The physical arrangement of the
vegetables in the baskets suggests a plurality-cover of the ve-
etables with the cells of the cover corresponding to basketsful of
vegetables. [The original utterance] is true and informative on
the intended intermediate distributive reading reading because
the verb phrase is true of every member of that cover.” [Schwarz-
schild, 1990]

He suggests that we might like to consider all verbs as being considered dis-
tributive with respect to some contextually determined cover [Schwarzschild,
1990]. This is perhaps at odds with the evidence that the fully distributed
form can be easily disambiguated using such terms as “each”.

Schwarzschild presents a strong argument against structured nominals.

He suggests that these readings have been considered important because
there does seem to be some import taken from the choice of nouns. If the
animals are just the cows and the pigs, then:

The cows and the pigs were separated.

seems to suggest something different from:

The young animals and the old animals were separated.

In a null context — where there are no prior statements about the manner in
which the animals were separated — the former suggests separation took
place by species: “the cows were separated from the pigs”, whereas the
latter that the separation was by age [Schwarzschild, 1992]. The question
is whether this aspect of the import of these sentence should be construed
as some inherent part of the meaning of conjoined noun phrases. However,
with some verb phrases such as “filled the barn to capacity” it is intuitively
the case that the sentences:
CHAPTER 2. BACKGROUND

The animals filled the barn to capacity.
The cows and the pigs filled the barn to capacity.
The young animals and the old animals filled the barn to capacity.

have the same import. If the noun phrase is given a different structure in each case, then the verb phrase must have homonyms, or lifted forms, so that when the verb phrase applies to “the animals”, the lifted form applies to any structured representation of them. In sentences where there are both structured and unstructured interpretations of one noun phrase, as in:

The young animals and the old animals arrived from different places and filled the barn to capacity.

As mentioned before, the noun phrase could be given a structured representation, and the unstructured property “filled the barn to capacity” can apply to the structured nominal by using a lifted form. Lifting, or type-shifting, as used by Lønning comes from Partee and Rooth [Partee and Rooth, 1983]. It is effectively a generalisation of meaning postulates that have been used in the literature on plurals [Bartsch, 1973, Scha, 1984, Hoeksema, 1983]. Schwarzschild suggests that the reverse of this phenomenon also exists, so that if a verb phrase applies to a structured nominal, then a lowered homonym of the verb applies to the unstructured nominal. If we have:

The cows and pigs were separated.

then it is the case that:

The animals were separated (by species).

We might prefer to utter the first sentence to indicate directly that the animals were separated by species (according to some Gricean maxim of quantity), but the adjunct “by species” would cancel this maxim.

Schwarzschild considers how these lifting and lowering operations might iterate. Lifting a lifted predicate results in a lifted predicate. Lowering a lowered predicate results in a lowered predicate. Lowering a lifted predicate results in the original predicate, if it is not a higher-order predicate (otherwise it is a null predicate, as is a lifted higher-order predicate according to Schwarzschild’s definitions). The more complex case is the lifting of a lowered predicate: such a predicate may apply to a group whose structure does not reflect the appropriate interpretation. If the animals were separated according to species, then we have:

*unshifted predicate:*
The young animals and old animals were separated.
CHAPTER 2. BACKGROUND

\textit{lowered predicate:}\n\begin{quote}
The animals were separated (by age).
\end{quote}
\textit{lifted lowered predicate:}\n\begin{quote}
The cows and the pigs were separated (by age).
\end{quote}

However, if something is separated by age, it is separated. Thus it is the case that “the cows and pigs were separated”, although the mode of separation is not indicated. Schwarzschild seeks to argue that there might be times were it is appropriate to utter a sentence with the lifted lowered predicate: the different structure of the nominal may be appropriate for some other conversational reason. Taking another example from Schwarzschild, if we have four lawyers, two defence and two prosecution, and one of the defence lawyers is a woman, the rest are men, then we might utter one of the following:

\textit{unshifted predicate:}\n\begin{quote}
The defence lawyers and the prosecution lawyers used to fight each other in court every day.
\end{quote}
\textit{lowered predicate:}\n\begin{quote}
The lawyers used to fight each other in court every day.
\end{quote}
\textit{lifted lowered predicate:}\n\begin{quote}
That woman and those three men used to fight each other in court every day.
\end{quote}

In some circumstances, it might be most appropriate to utter the final, lifted lowered reading, when identifying the individuals concerned [Schwarzschild, 1992]. The lifted lowered predicate also allows partitions of the conjoined noun phrases not obtainable with a structured conjunction. The ambiguity — the distinction in meaning — that structured nominals allow us to represent requires us to make verbs ambiguous in a manner that can only be expressed if structured nominals are adopted. The ambiguity between the various lifted and lowered verbs does not appear to have any motivation other than to recover the meanings otherwise lost by the structured representation of nominals, and that ambiguity is not even expressible unless structured nominals are adopted. There is no means of indicating the desired reading of the verb phrase with regard to the various homonyms. When it is possible to elaborate on the intended interpretation of a verb phrase, by adding adjuncts, such as “separated \textit{by age}” then there is no need for a structured representation of the nominal. This can be taken as evidence that the various homonyms, created by lifting and lowering, do not represent a genuine ambiguity, but are just an artificial requirement of the belief that the static representation of conjoined nominals must be structured.

Schwarzschild presents further evidence that this is not a genuine ambiguity, using examples of verb phrase deletion:
CHAPTER 2. BACKGROUND

In this neighbourhood, the upper class children, and the lower class children do not hate each other, but the adults do.

The various children could not get along with each other nor could the men and women.

In the first sentence, the first verb is understood on its unshifted reading, and the deleted verb, when replaced, is understood on its lowered reading. The opposite applies in the second sentence. According to the use of verb phrase deletion to test ambiguity [Zwicky and Sadock, 1975], it should not be possible for the verb to be given different interpretations in the antecedent and deleted positions.

Considering the structured readings of nominals themselves, if it is assumed that they exist, then all predicates (derived from verb phrases) are potentially ambiguous. There are then no unambiguous predicates which can test the non-coreference of nominals such as “the cows and the pigs” and “the old animals and the young animals”. Schwarzschild suggests that the real difference between the meanings of such nominals is not in their static interpretation, but in their context change potential — that is, their effect upon the discourse. The noun phrase “the cows and the pigs” is given a static interpretation corresponding to “the animals”, but in the discourse itself, two further objects become available for anaphoric reference, namely “the cows” and “the pigs”, in that order. In a null context these additional referents can provide a clue as to the means by which the animals are separated in “the cows and the pigs were separated”. Schwarzschild suggests that this ties in with a dynamic treatment of “respectively” constructions\(^\text{14}\), as in:

George and Mike are running with Dan and Lloyd respectively.

where, according to Schwarzschild, the record of the linear order, required to interpret the sentence, is an aspect of the discourse and not the referents [Schwarzschild, 1992]. The evidence for this is that intersentential inferences are sometimes needed to recover the relevant reduction, as in the following example:

“The first book is 2,000 pages long and it barely fits in the book bag. The second one is only 20 pages long, you can put it in your pocket. I refer to these as the fat book and the skinny book respectively.”

as opposed to:

\(^{14}\)Including “correspondingly”, “analogously”, “equivalently”, and “in that order”.
“I refer to Miller’s books as the fat book and the skinny book respectively.”

which is difficult to interpret in a null context such as when appearing as the first sentence of a discourse [Schwarzschild, 1990].

Reducing Reciprocals

We might well contemplate what the interpretation of reciprocals might be: whether there is any reduction of their meanings to non-reciprocal sentences. Hoeksema and Link (in his earlier work) regard the reciprocal reading as a special kind of collective reading, a unary predication of the sum (that is, conjunction) of terms, or other definite descriptors, and of generalised quantifiers [Hoeksema, 1983, Link, 1984]. Lønning notes [Lønning, 1989] that their approaches seem to work for examples such as:

Mary and John love each other.
Kim, Leslie and Sandy love each other.
The girls love each other.
Some girls love each other.

but fails for conjoined generalised quantifiers:

Every linguist, and every philosopher love each other.

and non-logical quantifiers such as “most”:

Most linguists love one another.

The sentence:

All linguists love one another.

would be represented as predicating “love one another” of the sum of all linguists. We might pose the question “who loves whom?”. Link tries to connect the meaning of the reciprocal with the meaning of a corresponding binary relation [Link, 1984]. We may question what the sufficient and necessary conditions might be, in terms of pairs of linguists, for us to assert “all (of the) linguists love one another”. In many cases, it seems hard to distinguish the stage-level from the generic interpretations of such sentences.

Others have tried to give a representation of reciprocals directly in terms of binary relations, as opposed to unary relations over plural terms [Hintikka, 1974, Barwise, 1979, Westerståhl, 1987]. Thus:

Every linguist and every philosopher love each other.
is interpreted, roughly, as:
\[
\text{Every' (linguist')}_x \\
\text{Every' (philosopher')}_y \rightarrow \text{love-each-other'}(x, y)
\]

This is an example of the use of branching quantifiers, where although the nominal quantifiers have the same scope, they do not outscope each other [Hintikka, 1974]. This approach cannot cope with plural (agreement) noun phrases like:

Some girls love each other.

or conjoined plural (agreement) noun phrases:

All linguists and all philosophers love each other.

and, like the collective interpretation of reciprocals, it does not suggest a suitable representation for sentences with non-logical quantifiers, such as “most”. Note also, that as the verb phrase is represented by a binary predicate, only noun phrases with two conjuncts can be represented. To overcome this, the verb phrase could be taken to be a unary predicate of collections, as in the collective approach to reciprocals:

\[
\text{Every' (linguist')}_x \\
\text{Every' (philosopher')}_y \rightarrow \text{love-each-other'}(x \oplus y)
\]

This is suggested by Lønning. A second suggestion of his is to take the union of the collective and branching quantifier approaches in order to cover most of the data [Lønning, 1989]. Then verb phrase “love each other” would be given two interpretations, one which is a unary relation predicated of collections, the other a binary relation, as used in the branching quantifier approach. Lønning represents the branching quantifiers representation of a reciprocal:

\[
\frac{Q_1 x}{Q_2 y} p\text{-each-other'}(x, y)
\]

as something like:

\[
Q_1 x Q_2 y (p(x, y)) \land Q_2 y Q_1 x (p(y, x))
\]

Combining the two approaches leads to two representations for some of the examples, and still does not account for plurals with the quantifier “most”. There may be an equivalence between the two readings.

Fiengo and Lasnik, and Langendoen explored the possibility of collapsing the truth conditions of reflexives to predications over individuals [Fiengo and Lasnik, 1973, Langendoen, 1987]. Langendoen showed that it is not possible to give an equivalence, and Fiengo and Lasnik present an example
— “the trays are stacked on top of each other” — where it is not possible to present a principled, necessary condition. Langendoen also looked at reducing ordinary collectives to relations between individuals — reducing “the girls and boys looked at each other”, to “the girls looked at the boys and the boys looked at the girls” — and found similar difficulties. Lenning has suggested that no reduction is possible [Lenning, 1989, Lenning, 1987a].

Some of the various reductions in the truth conditions of reflexives that have been proposed are:

(i) Langendoen’s \textit{Strong Reciprocity}: where “the boys love each other” just in case every boy loves every other boy, for example [Langendoen, 1987]\textsuperscript{16}.

(ii) Langendoen’s \textit{Weak Reciprocity for Subsets}: “the prisoners released each other” if each prisoner released others, and was released by others [Langendoen, 1987].\textsuperscript{17} This is semantically very much like Gillon’s adaptation of a proposal by Higginbotham [Gillon, 1992, Higginbotham, 1981].

(iii) Fiengo and Lasnik’s \textit{Strong Reciprocity within Partitions}: some reciprocal sentences are true if there is a partition of the denotation of the subject such that strong reciprocity holds in each partition [Fiengo and Lasnik, 1973]: “the men are hitting each other” would be true if the men were in pairs that were hitting each other.

Langendoen and Fiengo & Lasnik show that the above proposals are too strong. “The boys kissed each other” can be true even if not every boy kissed every other boy (Strong Reciprocity), and if there was one boy who was not kissed then Strong Reciprocity within Partitions fails. Weak Reciprocity could be met if the un kissed boy was in a group that was kissed [Schwarzschild, 1992], but then we may question what it is that the meaning of the reciprocal is being reduced to: does it help us determine when a reciprocal sentence is true? Roberts makes a proposal based upon a suggestion by Emmon Bach in which there is a context sensitive quantifier \textit{ENOUGH} and the reciprocal holds of the sum if \textit{ENOUGH} individuals in the sum have the appropriate relation to \textit{ENOUGH} other individuals in the sum [Roberts, 1987].\textsuperscript{18} This is not too strong for the above example.

\textsuperscript{15}It has been argued that this example presents problems because of the phrase “on top of”, which appears to express strict immediate precedence [Langendoen, 1987, Gillon, 1992]. Interpreting such phrases as strict precedence can cause problems in sentences without a reciprocal pronoun, as in: “Uncle Elephant was wearing everything on top of everything” [Gillon, 1992].

\textsuperscript{16}Set-theoretically, we can define Strong Reciprocity as: if $R(P)(S)$, where $R(P)$ is the reciprocal version of $P$, then for each $x \in S$, for all distinct $x' \in S, xPx'$.

\textsuperscript{17}Weak Reciprocity for Subsets can be defined as follows: if $R(P)(S)$ then for every $x \in S$ there exists $u \subseteq S, x \notin u$, such that $xPu$ or there exists $s \subseteq S, x \in s$ such that $sPu$, and there exists $v \subseteq S, x \notin v$ such that $vP$ or there exists $t \subseteq S, x \in t$ and $vPt$.

\textsuperscript{18}Slightly more formally, according to Roberts proposal, $R(P)(S)$ holds if \textit{ENOUGH} individuals in $S$ are $P$-ed by \textit{ENOUGH} others in $S$, and \textit{ENOUSH} individuals in $S$ $P$. 
CHAPTER 2. BACKGROUND

For some sentences, these proposals are not strong enough. "The prisoners on the two sides of the room could see each other" is judged false if there is an opaque barrier between the two sides of the room. However, Weak Reciprocity is met, as is Strong Reciprocity within Partitions. Roberts' proposal does make the sentence false if there are roughly comparable numbers of prisoners on each side of the partition. This is accidental, however: Schwarzschild notes that it is possible to construe the circumstances such that if there are a significant majority of prisoners on one side, then there might beEnough prisoners on one side to make the sentence (incorrectly) true. He suggests that the crucial point is to determine the relevant (or operative) subsums (of prisoners, in this example) to be used in the notion of Weak Reciprocity for Subsets. The operative subsums are to be determined by the reciprocal predicate, and the context. He combines this notion of operative subsums with Roberts' quantifier Enough [Schwarzschild, 1992]:

If a reciprocal applies truthfully to a noun phrase denoting sum S, then:

(i) There must be two or more "operative" subsums of S, call them $s_1, s_2, \ldots, s_n$.

(ii) In the extension of the main predicate there are two or more pairs of these subsums of the form $(s_i, s_j)$ where $s_i \neq s_j$.

(iii) Each of the operative subsums is an element of such a pair.

(iv) Enough of the members of $S$ are included in some operative subsum.

The idea is that the operative subsums are identified with the cells of a partition of the denotation of the subject. This is very weak. It does not give a procedure for determining the operative subsums. Schwarzschild notes that the relevant information need not be linguistic. This favours the view that the actual noun phrases used in a reciprocal need not compositionally allow the determination of the operative subsums (this would then require a compositional interpretation of non-linguistic evidence).

The weakness of this definition echoes the weakness that seems required in a reductive analysis of the meaning of generic interpretations. It perhaps indicates why a reduction of the meaning of reciprocals seems to give a generic flavour, where apparent counter-examples do not undermine the truth of the sentence.

Following his rejection of the static representation of structured noun phrases, Schwarzschild suggests that some variety of anaphora resolution can be used, though not to treat the operative subsums as pronominal elements that become tied to referents in the discourse. Rather, the discourse provides the properties which partition the subject denotation as appropriate. The sentence:

*Enough other individuals in S.*
CHAPTER 2. BACKGROUND

The animals were separated.

can be treated as elliptic for:

The animals were separated by \( n \).

where \( n \) is a noun denoting a family of properties, such as “age”, “gender” (as opposed to “by female”, “by 10 years old”). Some mechanism allows the appropriate family of properties to fill in the missing adverbial phrase. So, according to Schwarzschild, the difference between:

The cows and the pigs were separated.
The animals were separated.

is that the former, in a null context, provides the source of the missing family of properties indicating the relevant partition [Schwarzschild, 1992].

Transitive Verbs

It is perhaps also legitimate to be cautious as to the problem of reducing the truth conditions of sentences such as “three men carried four pianos” [Roberts, 1987]. Scha analysed such sentences as:

600 Dutch firms have 5000 American computers.

in terms of cumulative — as opposed to collective — readings. The cumulative reading of this sentence entails that there are exactly 600 Dutch firms, and 5000 American computers [Scha, 1981]. Partee showed that these cumulative readings, where neither quantifier is within the scope of the other, are really a sub-class of the collective readings with two or more noun phrases [Roberts, 1987]. Link suggested that we might wish to consider transitive verbs as distributive on one or both arguments [Link, 1991b]. Others have also considered the phenomena of distributivity with two or more arguments of a verb at the same time [Scha, 1984, Scha and Stallard, 1988]. Gil has studied the readings of sentences with transitive verbs (with just two noun phrases) accepted by native speakers. He found four commonly accepted readings: two where the noun phrases were read distributively, with either scope; a strongly symmetric interpretation, akin to Langendoen’s Strong Reciprocity; and a weakly symmetric reading, like Langendoen’s Weak Reciprocity [Gil, 1982]. Roberts criticises Gil’s examples — the verbs used are considered to be strongly reciprocal (like “read”; “see”; and “run”), and the nouns may get an indefinite group-denoting interpretation — and notes that Gil does not consider the group-group interpretation present in Scha’s ‘cumulative’ examples, where there is no reduction of the relationship between the collections representing the subject and object nouns [Roberts, 1987].
It would seem that giving a general compositional or reductive strategy for interpreting sentences with transitive verbs is as problematic as reducing the meanings of reciprocals, with the added complication of quantifier scoping effects. Roberts believes that the natural interpretation of Scha’s ‘cumulative’ sentences is the group-group interpretation, and that readings where the verb is taken to be distributive on one or both arguments are often highly marked. The suggestion is any reduction to quantification elements, with Scha’s cumulative examples, reciprocals, and multiple definite plural noun phrases, are not part of the truth conditions, but are implied by various pragmatic factors (lexical, contextual or general) except in the case of reciprocals, where the reciprocal itself may contribute a distributive element [Roberts, 1987]. Schwarzschild supports this view, and gives some examples where non-linguistic context affects the possible interpretations of sentences with transitive verbs [Schwarzschild, 1990].

Because of the problems of quantifier scoping, and of deciding the acceptable distributive and collective readings for sentences with transitive verbs, I shall avoid them in my formal theory.

Collectives and Intensionality

Landman used some of the ideas for representing structure in nominal expressions to elaborate upon the representation of collective nouns, exemplified by committee-like terms [Landman, 1989]. These are sometimes confusingly referred to as group or group denoting nouns, leading to an ambiguity in the term “group” as used in the literature: the denotation of committee-like nouns should not be confused with the groups that arise from adding structure to conjoined nominals [Lønning, 1989, Page 143], although nouns represented by grouped sums are also sometimes called group denoting. In British English, the collective nouns have plural number agreement in their singular form. They appear to allow several objects to be grouped into one. Examples of groups are “committee”; “bunch”; and perhaps words like “forest”, as in “a committee of men”; “a bunch of flowers”; “a forest of trees”. In American English, the singular group nominals (at least when they appear without an “of” phrase) have singular agreement. Some speakers of English allow either number agreement with the group forming noun or the underlying objects, depending upon which objects seem most relevant, even when the underlying objects are not presented in an explicit “of” phrase.19

19 Some native speakers of Dutch tell me that this is an increasingly common occurrence in that language.

When we say “the committees were founded this year” it is the committees that were founded, not the members. If we take “were founded” to be distributive, then some means must be found to stop distribution to the members. Were we to make the rather naive assumption that the committees
CHAPTER 2. BACKGROUND

were the sum of their members, then “the committees” would have to be, in turn, the sum of these objects grouped in some way to stop distribution, in a manner akin to the intermediate distributive readings. There are cases when we seem to be able to refer to the underlying objects with a group nominal: “there was a bunch of flowers, . . ., the bunch was red”, meaning that the flowers in the bunch were red.\footnote{With a singular group noun, some native speakers take the number agreement of the verb to indicate whether it is the group itself, or the underlying objects that are being referred to. For these speakers, “the bunch was p” implies that the bunch, as a whole, had the property $p$; however, “the bunch were $p$” implies that the individuals within the bunch each had the property $p$. The number agreement of received British English, and the remaining ambiguity with plural groups should make us cautious of trying to obtain some semantic generalisation from this. Lønning argues that for a theory of the semantic representation of plurals, the semantic arguments should have the most weight [Lønning, 1987a]. Syntactic evidence of number agreement should only be taken as secondary evidence.}

A formal reason for supporting the extensional view of groups, due to Lønning, is given if we consider a formal representation using complete atomic Boolean algebras [Lønning, 1989]: if we treat the copula as representing equality, then:

John and Mary are a couple.

cannot be represented with:

$$j' \oplus m' = c'$$

unless, “a couple” actually refers to a sum. Clearly, this can also be construed as a counter-example to taking the copula as equality.

As mentioned in arguments above, doubt has been cast upon the need for an explicit, compositional, reductive representation of intermediate distributive readings. If we dispose of such readings by allowing them to be subsumed by collective readings, then we are left with the problem of representing group readings. However, the use of theories designed originally for intermediate distribution and structured reflexives to explain the semantic behaviour of committee-like nominals originates in a rather extensional view of these objects. In Link’s treatment of intermediate distributive readings and committee-like terms he recognises that the latter should be taken to be intensional [Link, 1984]. As observed in [Bennett, 1977] and by others, two different committees may have exactly the same members, but still be different committees and have different things predicated of them. If the committees necessarily have the same members, then even a possible worlds approach cannot deal with this. A committee must be more than its members. To say that a committee meets is different to saying that the members of a committee meet: when a committee meets, all of the members need not
be present. Two committees may necessarily have the same members, yet the committees remain distinct (different things may be predicated of them).

An additional argument against the identification of groups with sums of individuals is that certain things can be predicated of a collection referred to as a group, which cannot be predicated explicitly of the sum:

The group of boys has five members.
The boys have five members.

Also, group-denoting nouns themselves can be pluralised, and ascribed collective properties [Chierchia, 1982a]:

The groups have 214 members, altogether.

Blau argues initially that the group-denoting NPs have the same denotation as sums when ascribing collective properties, but there is a difference in behaviour with distributive properties [Blau, 1981]. Later he admits that groups and sums must have different denotations even with collective properties. He could not overcome the identification of the following sentences:

John counted the cards.
John counted the decks of cards.

Landman proposes that the identity between sums and groups should be prevented, and can be represented using singleton set formation [Landman, 1987, Landman, 1989]. A sum is thus represented by a set, and a group constituted of that sum is the singleton set of the set. Landman reappraises Link’s work on plurals, and recasts it directly in terms of sets, rather than lattice structures. He initially argues that all definite NPs are ambiguous between sums and groups. The denotation of group denoting nouns contains only singleton sets. Later, to avoid problems with sentences where the NP seems to have both a collective and distributive reading:

The boys shared a pizza and drank a coke (each).

he has all definite plural NPs denoting singleton sets (his groups), and redefines the distribution operator, to allow distribution into groups.

Landman claims that he re-phrases Link’s work in order to give more structure to the semantics, however, Lønning points out that the earlier writings of [Link, 1983, Link, 1984, Link, 1991b] have even more structure in the models [Lønning, 1989].

Landman’s analysis leads to two committees necessarily being the same committee if the members are the same, which is equivalent to an identity between sums and groups. Lønning says that all that can be salvaged from

---

21 As groups admit these apparent counter-examples, they may be akin to generic terms.
22 This may be a poor example; the semantics of “altogether” may undermine its validity.
this is an ambiguous conjunction. This is effectively the same ambiguity as proposed by [Link, 1984] to account for intermediate distribution [Lønning, 1989].

As Lønning writes, it is important in the formal semantics of NL to get the inferences right. However, it can be hard to discern those inferences that are part of the logical behaviour of language, from those that are contingent on extra-linguistic world knowledge. It is difficult to see how far words like “group” and “member” should have their meanings tied up in the logical vocabulary of the semantics. In this respect, a weak theory is better than a stronger one.

It is possible to allow committee-like objects to be represented by terms which, although related to the underlying members, would not have to be equated if they had the same underlying objects. This intensional treatment of groups can be divorced from the theory associated with intermediate distribution. Barker did this explicitly in his account of group readings [Barker, 1992]. He took committee-like objects to be distinct if they could possess distinct properties. A function $f$ would take such a group to the sum of its underlying individuals. When ascribing a property to such a group $x$ in NL, there is then a potential ambiguity as to whether the property was applied to the group $x$ itself, or to the underlying objects $fx$. Similarly Dölling extended mereology to cover group phenomena [Dölling, 1990]. He took the meaning of “the group of children met” to predicate “met” of some group which is constituted by children.

The existence of strong arguments against an identification of groups with sums surely means that, in general the identification should not be made, but that a sentence like:

The group of boys met.

has an ‘extensional’ distributive reading which is equivalent to the explicit distribution in:

The members of the group (of boys) met.

This account of intensionality is perhaps inadequate in the general case, as it fails to account of similar phenomena with nominals not normally regarded as committee-like. These are examined by Landman [Landman, 1989]. We might say the following, without contradiction:

The judge is strict.
John is liberal.
John is the judge.

This makes sense if we take John to be a strict judge or — as Landman would have it — John, as a judge, is strict, though John (as himself) is liberal. There are two approaches to these examples: Landman accounts
for these examples by introducing *individuals under roles*. Individuals can be embellished with the property describing the guise under which they are being considered; the alternative is to use property modifiers, so the first sentence above attributes the property “strict-judge” of John, as opposed to attributing “strict” of John-as-a-judge. Landman rejects the use of property modifiers, but I believe it can be made to account for all of his examples. Although not essential for a theory of plurals, I employ Landman’s ideas when controlling distributive inferences with mass terms (Chapter 8).

A further intensional issue discussed in the literature concerns definite descriptors. If two definite descriptors have the same physical extension (they ‘denote’ the same physical object, in some sense), should they be considered to be logically equivalent — are they *really* the same? As a potential counter-example, even if all window-parts make up the windows, and all the windows are made of window-parts, they may be ascribed different properties in NL: we might say the windows were destroyed without wanting to imply that the window-parts were destroyed. In this case we might like to block distribution in the same manner as with committee-like objects, by having some form of non-associativity. I favour the view that the non-associativity should come about because in the semantics, the windows and the window-parts are *not* equated. These issues are covered later, when I discuss different ontologies and how they affect a treatment of mass terms. There I indicate how it might be possible to allow the extensions of the definite descriptors to be equated, without resulting in contradictions in the semantic representations.

**Distribution, Scoping, and Logical Complexity**

Bennett and Lønning both attempt to demonstrate that the semantics of plurals show that NL is at least second-order [Bennett, 1975, Lønning, 1989]. Lønning also demonstrates that the expected inferences for NL only require a subset of second-order logic which satisfies general models [Orey, 1959] and hence which is effectively first-order.

One requirement for an apparently second-order logic to be effectively first-order is that there is no genuine case of universal quantification over all possible sets, or in this case, all possible collections. Link suggests that there may be such examples [Link, 1987]:

- (All) competing companies have common interests.
- (All) children of the same age stick together.

Taking the first example, it is intended to be read as:

- Each collection of companies, such they are competing, have common interests.

Lønning suggests that such examples can only be given a dispositional, or generic interpretation [Lønning, 1989]. This becomes clearer if we consider replacing “all” with “most”: 
CHAPTER 2. BACKGROUND

Most competing companies have common interests.
Most children of the same age stick together.

where it cannot be the case that most collections of children stick together. Roberts discusses the representation of non-logical quantifiers like “most”; “few”; and “many” in some detail [Roberts, 1987]. If we take the sentence:

Few people agree (with each other) on this issue.

the representation of “agreement” must take a collection as its argument. If this is paraphrased as “there are few groups of people who agree with each other on this issue”, then it would seem that there is quantification over arbitrary sums. This cannot be the correct interpretation. If there are four individuals \(a, b, c, d\) and the first three agree, then there are 4 collections of individuals that all agree: \(a \oplus b\); \(a \oplus c\); \(a \oplus b \oplus c\); \(b \oplus c\), and 7 collections of individuals that do not all agree: \(a \oplus d\); \(b \oplus d\); \(c \oplus d\); \(a \oplus b \oplus d\); \(a \oplus c \oplus d\); \(b \oplus c \oplus d\); \(a \oplus b \oplus c \oplus d\). Thus the sentence is incorrectly taken to be true, according to this analysis. Roberts attributes this argument to Angelika Krazier. For the example in question, Roberts notes that what is at issue is the cardinality of the maximal collection of people who agree. Determiners like “few” then give rise to restricted quantification, not arbitrary quantification, over collections. There are, however, examples that call this analysis into question. The sentence:

Few people are wearing matching sweaters.

would be true on the suggested translation, if there were 100 people, and 50 pairs of matching sweaters, as a maximum of 2 people are wearing a matching pair. Roberts suggests that in this case we require quantification over sums that are pairs [Roberts, 1987]. One conclusion that might be drawn from some of these problems is that in no general formal unpacking is possible. However, it does appear to be the case that such determiners do not give rise to quantification over arbitrary collections. Later, Link seems to support Lønning’s objection, suggesting that the noun phrase “most competing companies” should be read as “in most of the situations where we have a group of competing companies” [Link, 1991a].

Schein explored taking the quantification to be over events, so that:

Few experts (ever) agree.

is paraphrased “whenever there is an event of agreeing, it involves few experts” [Schein, 1986]. Roberts gives examples that show this can be counter-intuitive:

Few good students are unprepared.
seems to be about individuals of a certain character, not events [Roberts, 1987].

Even if universal quantification over collections is disallowed, there is a potential problem with existential quantification over collections (stemming from indefinite noun phrases) if the scope of negation effectively gives rise to universal quantification over arbitrary collections. For the behaviour of NL to be limited to first-order power would require restrictions on the scope of negation in NL sentences involving collectively read noun phrases.

For a second-order representation to be restricted to the logical power of first-order logic, there must be some means of restricting — or securing [Orey, 1959, Lönning, 1989] — variables ranging over collections so that they need only be considered to be instantiated to a subclass of collections (those collections allowed by a general model). Simplifying somewhat, plural noun phrases cannot secure these second-order variables (in Lönning’s representation, a pluralised predicate ranges over all collections) but un-negated verbs can. Negated predicates (representing negated verb phrases) cannot secure variables ranging over collections. This means that for the representation of NL sentence to be restricted to an effectively first-order fragment of logic, if negation is given narrow scope in such sentences as:

Some of the girls did not lift the stone.

then according to Lönning’s thesis, the quantification must be over individual girls not collections of girls. That is, if the representation of NL is restricted to first-order power, then both collective and distributive readings are available with wide scope negation, but only distributive readings are available with narrow scope negation.

Lönning examines other cases of the restrictions on the scope of terms in sentences with collective readings to support his case. Lakoff claims that while the sentence “all the boys carried the couch upstairs” is ambiguous between the distributive and the collective readings, when in an opaque context such as “Sam believed that the boys carried the table upstairs”, then only three readings are available: when “believed” has widest scope then both the distributive and collective readings are available; however, if “all the boys” has widest scope, Lakoff claims that only the distributive reading is allowed [Lakoff, 1970]. Lönning, although admitting that this has been disputed and that Lakoff does not consider all the relevant data [Roberts, 1987], uses this to give intuitive support to the idea that collective readings have narrow scope with respect to negation [Lönning, 1989].

Verkuyl also discusses the relationship between negation and indefinite collectives within the same sentence [Verkuyl, 1988]. He claims that there is no distinction between collective and distributive interpretations in negated sentences (where the negation has narrow scope). Lönning doubts the evidence for this view, believing that the readings do not become equated, but
that the collective reading is impossible. He suggests that this may be because the correct interpretation of indefinite NPs is as indefinite terms bound by the context in some way as in theories of discourse such as those devised by Kamp and Heim [Kamp, 1981, Heim, 1982]. However, Roberts shows that not all scoping phenomena can be treated in the same fashion. For example, the collective terms can outscope other operators, in particular, the distribution operator [Roberts, 1988] as in the sentence given before:

Five insurance associations gave a $25 donation to several charities.

where the subject noun phrase can be read collectively, and outscopes a distributive interpretation of "several charities". This interpretation of the sentence indicates why the distributive/collective distinction cannot be accounted for by quantifier scoping [Roberts, 1987] or by counting the number of events [Lønning, 1989].

Lønning also links his account of logical complexity with the literature on referential readings of indefinites. In this connection, Fodor and Sag present an argument against always giving indefinite terms minimal scope [Fodor and Sag, 1982]. They claim that indefinites are ambiguous between existential quantifiers and referring expressions, where in the latter case, the indefinite has maximal scope in the sentence; and that the indefinite either has narrow or maximally wide scope, never intermediate scope. They found the wide scope reading to be more natural with more specific indefinite descriptions. Lønning says that their examples also show that referential, wide scope reading, is available when a noun phrase is read collectively [Lønning, 1989]. Donnellan has a similar distinction between referential and attributive usage of the definite descriptions, which corresponds to Fodor and Sag’s quantificational and referential distinction [Donellan, 1966]. Lønning notes that the attributive usage corresponds to Russell’s original quantificational interpretation of the definite descriptor where there is a claim to uniqueness. In the case that a speaker intends to refer to a particular object with the referential usage of a definite descriptor (like “the boy”) there is no requirement for uniqueness even within a particular context. With referential usage we have “John went on a trip with his grandparents” where only two grandparents are intended, as opposed to the attributive (quantificational) use in “Every child loves his grandparents”, where all four are intended [Lønning, 1989]. Going back to the previous example, it would suggest that in the sentence:

---

23 See Footnote 4: with this representation, collective readings always have narrow scope, counter to the example.

24 The argument against the ambiguity being concerned with the number of events parallels that against the ambiguity being one of quantifier scoping. On this account, the fully collective reading would be represented by one event; the fully distributive by one event for each individual. In the example, there would be a “giving” event for each charity, yet the subject noun phrase may have a collective interpretation.
CHAPTER 2. BACKGROUND

Some of the girls did not lift the stone.

the noun phrase “some of the girls” can be given a collective interpretation when the negation is given narrow scope, by representing it as a definite reference to a particular collection of girls, rather than quantification over collections of girls.

If this view is accepted, then there is no major difference between indefinite and definite noun phrases: referential indefinites can be used like a (referring) definite descriptor, and the attributive definite descriptor can be treated as a (quantified) indefinite. The only legitimate distinctions would be between introducing new objects, and referring to old objects, as used in theories of discourse [Kamp, 1981, Heim, 1982]. The attributive (quantificational) use introduces an object, the referential use refers to it. Lorenz says that in discourse an object is never introduced with a negated property: this would support his view that when the indefinite has wide scope, and is given an attributive, or quantificational interpretation, then the verb is usually not negated.

The alleged referential/quantificational ambiguity has its roots in work by Strawson [Strawson, 1950, Strawson, 1952]. It has been disputed, as it can be taken as a distinction in intended use, rather than in the semantic representation [Ludlow and Neale, 1991]. Although these issues are important, in my opinion, many of them belong in a theory of discourse. If a theory for the representation of plurals is intimately bound up with a theory of discourse, then many of the issues may be confused. I will not attempt to make my account compatible with any particular theory of discourse representation.

As I adopt a first-order theory of properties to represent the semantics of NL, I need not consider the merits of taking English to be restricted so that, given a second-order representation language, its semantics is effectively restricted to be first-order. The representation I shall use will only allow a first-order formalisation of the semantics of English regardless of whether the interpretation of NL has restrictions on the relative scoping of collectively read noun phrases and negation: all quantification is restricted to denotable terms, there is no quantification over arbitrary terms. Any such scoping restrictions may be a consequence of pragmatic issues in a theory of discourse. They cannot be motivated by consideration of the restriction of the logical power of the representation language to that of a first-order theory when the representation I shall use is already first-order.
2.2 Link’s Treatment of Plurals

In this section I shall examine in detail an existing analysis of the semantics of plurals due to Link as developed in [Link, 1991a]. There are many such analyses, I shall choose this presentation as it is given in terms of axioms within the theory, rather than as constraints on a model of the representation. The details are given as an example of a theory of plurals without the embellishment of property theory. Link’s treatment will be used as a basis for an axiomatisation of plural terms in property theory, and subsequently as a basis for an axiomatisation of a potentially atomless mereology.

The representation language in [Link, 1991a] is equivalent to one of the systems presented in [Lønning, 1989], except that the separation of the axioms of the language from those of its model is cleaner. Lønning is more concerned with issues of the logical complexity of a language for representing plurals with respect to various models. However, certain points in Link’s paper can be clarified by referring to Lønning’s work.

Although the phenomena covered are within extensional interpretations of sentences (as opposed to opaque readings, and perhaps generic and habitual interpretations), and the logics typically used to represent such readings are extensional logics, there is an implicit assumption about some primitive intensional behaviour in the objects of discourse that permits, for example, a human to be treated as more than the sum of its physical parts.

Link takes the view that definite descriptors are terms which may fail to denote, rather than the Russellian view that they are existential quantifiers with a uniqueness condition. This requires either a free or partial logic [van Fraassen, 1966, Scott, 1970]. Link chooses a free logic, where the quantifiers range over defined terms, and the proposition $Et$ holds iff $t$ is defined. His equality relation is strict: it requires the related terms to denote. If a definite descriptor fails to denote, then it cannot appear as an argument of the strict equality. No useful proposition in his language of wff can be constructed involving such a term. If we take the computational view of strictness, then if a definite descriptor fails to denote, any strict operation involving it will be effectively undefined. A proposition containing such an undefined term will be undefined if the logical connectives are strict, else its undefinedness will be contingent upon the exact occurrence with non-strict logical connectives.

Link assumes some primitive sum-formation operator $\oplus$, which creates a new individual out of two terms. An ordering relation can be defined in terms of this operator in the usual way:

**Definition 2.1** If the sum of $t, s$ is $s$, then $t$ is a part of $s$.

$$t \leq s \iff t \oplus s = s$$

Like equality, the ordering relation is strict: it requires both arguments to be defined. Thus the notion of definedness can be given in terms of the ordering relation:
Definition 2.2 \( A \) term is defined if it can be an argument to the ordering relation.

\[ Et \leftrightarrow \exists x(x \leq t) \]

Link excludes a bottom from his theory (there is no term less than all other terms). Thus his test for atomicity does not take account of such an element:

Definition 2.3 An atom is something which denotes, and has no proper parts.

\[ At(t) \leftrightarrow Et \& \forall x(x \leq t \rightarrow x = t) \]

Link has lambda abstraction of variables from propositions in his language. The precise nature of this abstraction is not clear from his axiom:

**Axiom 2.1**

\[ Et \rightarrow (\lambda x(\varphi[x])t \leftrightarrow \varphi[t]) \]

If only a single abstraction can be performed just on propositions, and they can only appear in the context of an application, then it does not seem to serve any great purpose. If, alternatively, it can appear in isolation, then what is the status of such an object in the language? Lønning provides us with an elaboration of this abstraction: the denotation of \( \lambda x(\varphi[x]) \) is that set of terms \( t \) for which \( \varphi[t] \) holds [Lønning, 1989]. Predication is modelled as set membership. This makes it explicit that we are dealing with a purely extensional framework.

The next three axioms make the summation operator symmetric, associative and idempotent.

**Axiom 2.2**

\[ t \oplus s = s \oplus t \]

**Axiom 2.3**

\[ t \oplus (s \oplus r) = (t \oplus s) \oplus r \]

**Axiom 2.4**

\[ t \oplus t = t \]

As mentioned above, Link prefers there not to be a bottom element, an element which is a part of everything:

**Axiom 2.5**

\[ \sim \exists x \forall y(x \leq y) \]
CHAPTER 2. BACKGROUND

It does seem to be the case that in English, there is no term which refers to such an object. However, even though there may be no such term in natural language, it is perhaps questionable whether we can then infer that there is no such object. Here, there is perhaps a confusion between an axiomatisation of referring expressions in natural language, and axioms governing the language chosen for representing natural language terms. Lønning draws a distinction between mathematical objects, and the use of those objects as denotations. It will turn out that for the models of our chosen representation language it is convenient for there to be a bottom element. Thus the absence of a bottom in the representation language would be inconvenient. No problem, however, is created if we just say that there is no plural or singular term in English that denotes such an object.

The next two axioms are related. Axiom 2.6' requires that all (plural) objects have atomic parts. Axiom 2.6 is stronger, it requires that all (plural) objects are have atomic parts, and that two different terms must have different atoms as parts, unless one of those terms is a part of the other:

Axiom 2.6

\[ \forall x \forall y (x \not< y \rightarrow \exists u (At(u) \& u \leq x \& u \not< y)) \]

Axiom 2.6'

\[ \forall x \exists u (At(u) \& u \leq x) \]

If there is some object that has a property, then it must be the case that the definite descriptor of that property is defined. That is, a definite descriptor is defined if there is some object which has the property in question:

Axiom 2.7

\[ \exists x (P x) \rightarrow E (\sigma x P x) \]

Link calls this definable completeness: if there is an object with property \( P \), then the sum of objects with that property, \( \sigma x P x \), exists. Completeness here is used in the lattice-theoretic sense, as opposed to logical completeness with respect to some model for the language.

This notion of definable completeness is used by Link and Lønning because it reduces the logical complexity of the language. The account of definable completeness in [Lønning, 1989] relates this lattice-theoretic completeness to the notion of logical completeness: if there is a set in the model corresponding to the denotation of a term in the language, then the sum (or least upper-bound) of that set is part of the model, and it is denoted by the supremum term in the language. It is shown that without this restriction to definable completeness, then the language cannot be proven to be logically complete: if only those terms which can be denoted by terms in the language are present in the model, then completeness is guaranteed.
CHAPTER 2. BACKGROUND

The definite descriptor of plural terms must have some particular properties, namely, if the definite descriptor denotes then all objects falling under the description must be part of the definite descriptor:

**Axiom 2.8** The definite descriptor $\sigma$ gives rise to an upper bound of all objects having the property in question.

$$\exists x (Px) \rightarrow \forall y (Py \rightarrow y \leq \sigma x Px)$$

and the definite descriptor is the least such object:

**Axiom 2.9** The definite descriptor is the least upper bound of the objects having the property in question.

$$\exists x (Px) \rightarrow \forall y (\forall x (Px \rightarrow x \leq y) \rightarrow \sigma x Px \leq y)$$

The next axiom says that if $^*P$ holds of an object, then that object is the least upper bound of some objects which have the property $P$:

**Axiom 2.10**

$$^*Pt \leftrightarrow t = \sigma x (Px \& x \leq t)$$

We can view $^*$ as an operator which turns a property into a collective property. Note, the axiom does not say that all parts of that object have the property $P$.

It must be ensured that if the least upper bound of two collections of atoms is the same, then they must be the same collections of atoms:

**Axiom 2.11** The atoms are supremum-prime.

$$\forall u (At(u) \& u \leq \sigma x Px \rightarrow \exists z (Pz \& u \leq z))$$

A distribution operator can be defined.

**Definition 2.4**

$$^D Pt =_{df} \forall u (At(u) \& u \leq t \rightarrow Pu)$$

With Lønning’s interpretation of the $^*$ operator in his language, the semantics of $^*$ and $^D$ become equivalent.

To show that this system is sound, a model which satisfies these axioms can be given. Link shows that his axioms are complete and sound with respect to definably complete atomic Boolean algebras with the bottom element removed. Singular terms denote atoms, plurals denote sums of atoms, and the definite descriptor/supremum operator denotes the least upper bound of a set which is denoted by a term in the language. Conjunction between nouns can be interpreted as join in the algebra.
Chapter 2. Background

Lønning shows that a representation with only definable completeness — such as the current theory — will be logically incomplete with respect to lattice-theoretic complete models. This parallels monadic second-order logic, which is logically complete with respect to general models, but not standard models [Lønning, 1989; Orey, 1959]. Thus definably complete models become the general models of these logics of plurality. However, if we are only interested in soundness, then a complete atomic Boolean algebra is adequate.

As an indication of how this theory might be used in the formal semantics of NL, given the following sentences:

(i) The boy cried.
(ii) Every man laughed.
(iii) John and Mary met.
(iv) The men died.
(v) Some women met.
(vi) All (of the) boys met.

the first five might be represented as follows:

(i) cried'($\sigma x$boy'$_x$)
(ii) $\forall x$(man'$_x$) $\rightarrow$ laughed'$_x$(x))
(iii) met'$_x$(j' $\oplus$ m')
(iv) died'$_x$(\sigma xDman'$_x$)
(v) $\exists x$(Dwoman'$_x$ $\&$ met'$_x$(x))

Considering example (vi), it is not so clear what interpretation could be given to a sentence with a noun occurring with the determiner “all”, as in “all boys met”. Such sentences appear strongly to favour a generic interpretation. It is hard to give it a stage level representation in some sense equivalent to “all of the $p$?”, it appears to differ little from that of “the $p$?”. What we might say is that “all (of the)" is distinguished from “the” in that it does not impart existential commitment. Thus sentence (vi) can then be represented as:

$$E(\sigma x$boys'$_x$) $\rightarrow$ met'$_x$(\sigma x$boys'$_x$)

which can be true, even if there are no boys. If we require there to be some boys, this could be changed to:

$$E(\sigma x$boys'$_x$) $\&$ met'$_x$(\sigma x$boys'$_x$)

which is false, if there are no boys. The third option is to treat it as a definite descriptor — so the truth value is not determined if there are no boys — embellished with a notion of involvement. Link offers a tentative proposal for this notion by giving some axioms concerning what it means for an individual to be involved in a particular predication. He notes that this can only be given a partial formal characterisation due to its pragmatic nature [Link, 1983].
CHAPTER 2. BACKGROUND

The NL quantifiers “all”, “some” and “the” give rise to collections in their representations, the quantifier “every” gives rise to singular individuals in its representation. This tallies with evidence from sentences involving count terms and quantifiers. Noun phrases with the quantifier “every” admit of singular anaphoric reference, in addition to singular verb agreement:

[Every woman]; thinks that [she]; will be a good mother.
* [Every woman]; think that [they]; will be a good mother.

unlike noun-phrases with the quantifier “all”, “some” and “the”, for example:

[All women]; think that [they]; will be good mothers.
* [All women]; think that [she]; will be a good mother.

(Note that the sentence “[all women]; think that [they]; will be a good mother” is just about acceptable, but has different entailments from “[all women]; think that [they]; will be good mothers”.)

Here “every” gives rise to universal quantification in the semantics. Syntactic agreement constraints require that “every” combines with singular count nouns, so the universal quantification is effectively restricted to singular individuals, which presumably can bind only singular anaphora. The quantifiers “all”, “some” and “the” are treated here as producing some collection, which similarly bind plural anaphora only.

Singular nouns denote simple predicates, whereas plural nouns can denote the singular predicate with the distributivity operator $^D$. Verbs that always distribute, like “die”, belong to a class of predicates which, if they apply to a term, also apply to the atoms of that term. Predicates that are ambiguous with respect to distribution can optionally have the distributive operator $^D$. This approach assumes that there is an ambiguity in the verb phrase between collective and purely distributive verbs, and that the collective reading subsumes possible intermediate distributive readings. This seems to be the simplest option if no account is taken of the discourse context of the utterance. Alternatively, we could follow Schwarzschild and have the various indeterminacies of the utterance reside in an operator that distributes to contextually determined covers of the noun phrase [Schwarzschild, 1990, Schwarzschild, 1992].

If we considered examples involving negation, such as:

Some women did not lift the stone.

this could give rise to a proposition that cannot be given an effectively first-order treatment [Lønning, 1989]:

$$\exists x (^D \text{woman}'(x) \& \sim \text{l-t-s}'(x))$$

as neither $^D \text{woman}'(x)$ nor $\sim \text{l-t-s}'(x)$ can secure the variable ranging over collections $x$ [Lønning, 1989, Orey, 1959]. This representation can only be
limited to first-order power if quantification is restricted to denotable collections. Notice that if “lift the stone” is given a distributive interpretation:

$$\exists x (D\text{woman}'(x) \& \sim D\text{t-s'}(x))$$

then this would be equivalent to Lønning’s suggested interpretation:

$$\exists x (\text{woman}'(x) \& \sim \text{t-s'}(x))$$

where quantification is restricted to individuals. The collective reading may seem undesirable as it is too easily satisfied: if there are ten women, and seven of them lifted the stone, then there would be 1013 collections of women that did not (collectively) lift the stone [Lønning, 1989].

Regarding groups, we could take committees and the like as atomic objects, and adding an operator which takes a group and returns the members of that group [Barker, 1992]. This allows groups to block distribution and thus obtain the correct reading of “the committees were founded this year”; and allows committees, for example, to be distinct, even though they have the same members.

I shall use the theory presented here as a basis for an axiomatisation of plural terms in property theory. Later, I will weaken the axioms to allow the domain to be potentially atomless.
2.3 Mass Terms

In this section I shall discuss the different formal apparatus that has been used to represent NL mass terms. Much of the essence of this review is taken from Bunt’s book on mass terms [Bunt, 1985]. The various ontological possibilities, and how they can be affected by the choice of semantic theory, are discussed in the next section.

Sentences with mass terms seem to be able to support entailments similar to those that occur with plurals. There are cumulative inferences:

\[
\text{Water is liquid} \\
\text{Oil is liquid} \\
\therefore \text{Water and oil are liquid}
\]

and apparently distributive inferences:

\[
\text{Water is liquid} \\
\therefore \text{Hot water is liquid}
\]

As with plurals, when discussing some of the examples, a problem is constantly presented by generic readings. Again, this work is primarily concerned with non-generic readings: readings that do not permit apparent counter-examples.

Quine defines mass terms as those objects that are cumulative, and do not “divide their reference” [Quine, 1960]. This means that mass terms lack a built-in mode of dividing their reference: they do not indicate, by themselves, what their minimal parts might be, nor what non-minimal parts are appropriate. This is related to distribution, as there are no prima facie relevant parts, to which we can restrict the distribution of a property. This can be taken as an argument against the use of an extensional mereology: there is little point in having a part-whole structure if the part-of relation cannot capture the relevant mode of division.

Quine distinguishes the use of mass terms before and after the copula. He claims that the use of a mass term after the copula results in the ascription of a property. The sentence:

This puddle is water.

would then result in the property is a bit of water being ascribed to the puddle. Before the copula, he treats mass terms on a par with proper names:

Water is a liquid.

attributes the property of being a liquid, to that object named “water”. In the light of more recent work on generics, we may consent to this if the sentence is given its natural generic interpretation. However, Quine apparently is not attempting to give an account of generics. He does not take the object
named by “water” as some intensional kind [Carlson, 1977], but as the single sprawling object that is the water of the universe.

Bunt discusses why Quine does not have the mass term denote such an object after the copula, and interpret the copula as part of. Thus the sentence “this puddle is water” could be represented, as:

\[ \text{this-puddle} \leq \text{water} \]

rather than something like:

\[ \text{water}(\text{this-puddle}) \]

The former gives a representation of “water” as a term rather than a property, corresponding to its representation in “water is a liquid”:

\[ \text{is-a-liquid}(\text{water}) \]

As Bunt himself notes, the reason why Quine does not choose this representation for mass terms after the copula, is that the interpretation of the copula as a part-of relation gives rise to distributive inferences, since the part-of relation is transitive. If “this puddle is water” is true, then, on the suggested interpretation, every part of this puddle would also be water (including the non-watery parts, such as the physical atoms). Quine would equate the truth conditions of \( x \leq \text{water} \) with those of \( \text{water}(x) \) only when considering those \( x \) that are not too small to count as water. This problem arises if we adopt a highly extensional ontology for mass terms, and have no formal method of only examining the relevant parts.

Quine’s proposal has been extensively criticised. Burge notes that Quine has no treatment of mass terms that occur neither before nor after the copula [Burge, 1972], as in:

\[ \text{Phil threw snow on Bill.} \]

Further, Quine does not establish a logical connection between the predicative and non-predicative uses of mass terms [Moravcsik, 1973], thus he is unable to account for inferences of the form:

\[ \begin{align*}
\text{This puddle is water} \\
\text{Water is wet}
\end{align*} \quad \therefore \quad \text{This puddle is wet} \]

nor that “water is water” [Pelletier, 1974].

Moravcsik gives two proposals for the treatment of mass terms [Moravcsik, 1973]. In one, he proposes that the object that is referred to by a mass term after the copula is the “whole” consisting of objects that are not too small to count as the substance in question. According to this “\( x \) is water” would be represented as:

\[ x \leq \text{water}^{\prime} \]

nor that “water is water” [Pelletier, 1974].
CHAPTER 2. BACKGROUND

This has the effect of restricting the distributive inferences that follow from using a part-of relation to interpret mass terms occurring after the copula: undesirable parts cannot be distributed to. As all its parts are water, then it is the case that:

\[ \text{water}' \leq \text{water}' \_p \]

Bunt notes that this proposal solves the wet puddle puzzle. It does not solve the “water is water” puzzle, however, as mass terms before the copula are still interpreted à la Quine. This statement would result in the assertion that all the parts of water are part of the “whole” consisting of parts not too small to count as water:

\[ \text{water}' \leq \text{water}' \_p \]

Unless we choose an ontology in which there cannot be such parts, then this results in a contradiction: the “whole” of parts not too small to count as water would be the same thing as the water itself. For this account to work, it is not possible for “wholes” to form objects that are sums of their parts, and have a uniform, transitive part-of relation. Otherwise the enterprise collapses at the outset since a part of a minimal part of water would not be water, but via transitivity, it would be a part of that “whole” of exclusively watery parts [Bunt, 1985].

In a second theory, Moravcsik follows the logic of this argument and explicitly proposes distinct part-of relations for each substance. The expression:

\[ x \leq_{SP(m)} m \]

means that \( x \) is a part of the mass term \( m \) and is large enough to count as a bit of \( m \). Moravcsik does not elaborate on how the various orderings are related: no transitivity between the different orderings is defined, for example. As it stands, this account cannot deal with the wet puddle puzzle, which is typically solved by using the transitivity of a part-of relation [Bunt, 1985].

Parsons presents an approach to the semantics of mass terms, where there is no ambiguity between substances before and after the copula, and where there is only one ordering relation amongst the substances [Parsons, 1970]. His theory is presented with a particular ontology in mind, where there are objects; bits of matter; and substances. For Parsons, mass terms denote substances. He has two relations, one, \( \mathcal{P} \), between objects and substances, and another, \( \mathcal{P} \), between matter and substance. The expression:

\[ \text{my-ring} \mathcal{P} \text{gold}' \]

\(^{25}\)Size is not all that matters: the ordering relationship should also prevent, as an example, some water and a hydrogen atom being taken as a part of water in the sense of \( \leq_{SP(m)} \). Generally, if \( x \leq_{SP(m)} m \) then \( x \) must not include a non-\( m \) part of \( m, y \), which is not a part of something large enough to be a part of \( m \) in \( x \). As an axiom:

\( (x \leq_{SP(m)} m \land y \leq x) \rightarrow \exists z(y \leq z \land z \leq_{SP(m)} x) \). We could strengthen this to a biconditional:

\( x \leq_{SP(m)} m \rightarrow \forall y(y \leq x \rightarrow \exists z(y \leq z \land z \leq_{SP(m)} x)) \).
intuitively means “my ring is gold”, and:

\[ m \triangleright \text{gold}\]

means that \( m \) is the matter that constitutes the gold. By distinguishing between substance and matter (and substance and object), the problem of minimal parts is avoided. The distinction between objects and matter avoids problems concerning the metaphysics of change. These ontological issues are discussed more fully in the next section. Bunt argues that Parson’s notion of substance is not well-defined. He states that substances defined in terms of some property, “small” for example, when fused, may not have that property. Thus we may have \( m_1 \) and \( m_2 \) defined in terms of being small, yet \( m_1 \) fused with \( m_2 \) need not be small [Bunt, 1985]. This is an argument for allowing the representation of the definite descriptor to combine only with cumulative properties (ignoring singular count nouns), as does seem to be the case in NL. This could be accounted for by representing the definite descriptor with an operator that takes the maximal element of a property [Scha, 1981]. With inhomogeneous notions such as “small” (assuming that it can have a predicative role), there will be no well-defined maximal element.

Further theories of mass terms can be taken to be essentially set-theoretic. Bunt divides the proponents of these theories into two camps, those that take the mass nouns to denote sets [Cartwright, 1965, Cartwright, 1970, Clarke, 1970, Grandy, 1973], and those that take them to denote predicates [Burge, 1972, Grandy, 1975, Pelletier, 1974]. He puts ter Meulen [ter Meulen, 1980] in both these camps [Bunt, 1985]. In an extensional framework, these notions are related. Burge would represent, for example, the word “snow” with a predicate that holds of any bit of snow. He takes “the snow” in:

The snow in the garden is two metres high.

to refer to some entity [Burge, 1972]. Bunt criticises the theory for being too weak, it does not, in itself, give a treatment of the “water is water” puzzle, for example.

There are set-based approaches which reduce the meanings of mass terms to those of count terms, by using some means of individuation. Then “furniture” can be interpreted as “pieces of furniture”; “water” as perhaps “puddle of water” [Strawson, 1959, Clarke, 1970]. The individuating term is context-dependent. However, any set-theoretic treatment of these individuated terms will be highly extensional, so “the ice”, if interpreted as “the pieces of ice” will denote different objects if the actual make-up of the ice changes [Cartwright, 1965, Pelletier, 1974]. Bunt suggests that any move to avoid this problem by resorting to an artificial individuating standard faces the problem of what that individuating standard might be, particularly with abstract mass terms, such as “time”.

Cartwright and Grandy have proposed that we take a mass term \( m \) to denote the set of all objects, or quantities, that are \( m \) [Cartwright, 1965,
CHAPTER 2. BACKGROUND

Grandy, 1973]. This notion of 'quantity' is perhaps the artificial individuating standard to which Bunt alludes. Montague takes such quantities to be minimal (atomic) individuals [Montague, 1973]. Pelletier and ter Meulen take this as a starting point for their intensional theories [Pelletier, 1974, ter Meulen, 1980]. Ter Meulen, however, takes 'quantity' to be an homogeneous notion, it can refer to the sums and parts of other quantities, not just some minimal elements. The intensional theories make a Quinean-like distinction between predicative and nominal uses of mass terms, except that, unlike Quine's account, the relationship between the two uses is given: the nominal use of a mass term is represented by the intension of the predicative use, in the possible worlds sense of intension [Montague, 1973, ter Meulen, 1980, ter Meulen, 1981, Pelletier, 1974]. Bunt believes that this distinction is of general applicability [Bunt, 1985]. To take an example from Bunt, the sentences:

Gold is a metallic element.
My tooth is filled with gold.

show the nominal and predicative uses of the mass term "gold", just as:

Diamonds are hard carbonites.
The ring is ornamented with diamonds.

show the same uses of the count term "diamonds". I suspect that these intensional accounts of mass terms are really striving to capture the notions of generic nominals. It should then come as no surprise that they can be used for both mass and count generics.

Bunt offers some criticisms of these theories. Ter Meulen's notion of 'quantity' causes some problems with quantifiers such as "much" or "most". Taking the quantifier "most" as an example, we might wish to elucidate its semantics by means of taking a bipartition of the denotation of the subject noun phrase, the largest partition having the property in question. In ter Meulen's theory, there is no means of measuring the largest partition: assuming that they are infinitely sub-divisible, there would be $\omega^2$ quantities in each partition.

The interpretation of the definite descriptor similarly has problems, as "the gold" intuitively refers to one thing. The only things that can be referred to in these theories are quantities, yet on ter Meulen's construal, such a quantity of gold will contain other quantities [Grandy, 1975]. Further, such quantities can be quantified over in turn as in "all the gold". The most natural solution, to treat the definite descriptor as denoting some maximal quantity, is not available to these set-theoretic approaches, as they have no means of constructing such an object: they have no formal notion of mereological fusion, or part [Bunt, 1985]. For similar reasons, they are presented with problems when trying to model the behaviour of non-homogeneous properties, such as in the sentence "the gold weighs seven ounces". If we were to
say that the sum of the weights of the quantities of gold is seven ounces, we would be in error, as there is no means of preventing overlapping quantities from counting more than once [Bunt, 1985].

Bunt’s representation is based around ensemble theory [Bunt, 1979, Bunt, 1985]. It is intended to capture the concept part of. He strives to avoid an essentially atomic theory — a theory that requires there to be individuals. He distinguishes between continuous and discrete ensembles, where the former are atomless, and the latter have minimal parts. He proves that his discrete ensembles behave like sets: for every atomic ensemble, there is a term (itself an ensemble) which is its unique element. This is like the member-singleton relation in set theory. In the continuous ensembles there is an empty ensemble that is a part of all other ensembles. This is unlike mereology, where there is no such term [Leonard and Goodman, 1940]. The set-like, discrete ensembles can be used to model the count nouns and verb-phrases and the continuous ensembles can be used to model (continuous) substances.

Bunt seeks to avoid the problems of minimal-parts and the distribution of properties to inappropriate parts by adopting what he terms the homogeneous reference hypothesis. This asserts that:

“Mass nouns refer to entities as having part-whole structure without singling out any particular parts and without making any commitments concerning the existence of minimal parts.”

This he takes to be a linguistic point: it does not matter what our theory of the world may be like, language behaves (according to Bunt) as if this hypothesis holds. This is distinct from Cheng’s view of distributive reference:

“Any part of the whole of the mass object w is w.”

where this is not a condition on the way a mass term refers, but on the referent of such a term, that is, it is an ontological commitment [Cheng, 1973].

What is the significance of Bunt’s hypothesis? We may be tempted to take it as asserting that within the theory, substances are independent of matter, in that they do not have parts that do not have the relevant property, perhaps be adopting something similar to Moravcsik’s proposal for a restricted notion of the part of relation [Moravcsik, 1973]. For example, by taking all parts of “the wood” to be “wood”, we never reach parts too small physically to count as “wood” (or else linguistic intuitions provide no insight into the limiting case). As mentioned above, Bunt rejects Moravcsik’s proposal because, as expressed, it is not coherent. Alternatively, the principle of homogeneous reference could be interpreted as an ontological commitment external to the theory, so that the denotations of “the wood” and “the sawdust” are really different things, even when all the wood in the world is turned into sawdust. This would suggest an intensional view of substances, as adopted by Parsons,
thus avoiding Quine's problems with the transitivity of the part-of relation. Bunt does not take this intensional view, however. He criticises Parsons' ontology as being too complex. Further, he asserts that there is no difference between using the terms "wood" and "sawdust" when the fusions of their physical extensions are equated, apparently ignoring the fact that were all wood turned into sawdust, "the sawdust was made yesterday" would not entail "the wood was made yesterday". Bunt would also reject the proposal to take the principle of homogeneous reference as an ontological commitment external to the theory, as this is against his assertion that ensemble theory is ontologically neutral. In addition, it does not account for instances where we might want to say that one substance is truly part of another. For example, it is very tempting to take "mud" to be a part of "muddy water".

Link also has an account of mass terms [Link, 1983]. This is a lattice-theoretic treatment, which he tries to integrate with a theory of plurals. Essentially, any fusion of substance in the lattice $D$, used to represent mass terms, can appear as an atom in the lattice $E$, used to represent count terms. The elements of the lattice $D$ are then a subset of the atoms of the lattice $E$. This is intended to correspond to the notion of being able to "quantise" substances, turning them into count terms, with phrases such as "a quantity of $m". He avoids some problems in the count domain by giving the representation extra structure: in his account, "windows" and "window parts" are not directly related, thus there is no equation of "the windows" with "the window parts". He defines a function $h$, a complete join-homomorphism\footnote{This means that $h(\bigcup_\mathcal{E} X) = \bigcup_\mathcal{D} X$ where $\bigcup_\mathcal{E}, \bigcup_\mathcal{D}$ are the join operators for the lattices $\mathcal{E}, \mathcal{D}$ respectively.}, which maps every count term onto a mass term, and is the identity function for mass terms. This corresponds to the notion of an object being composed of substances. Thus, count terms may be distinguished:

\[
[\text{the windows}] \neq [\text{the window parts}]
\]

even if they are composed of the same stuff:

\[
h([\text{the windows}]) = h([\text{the window parts}])
\]

This allows distribution to be blocked correctly in the count domain (ignoring group readings). However, we may still question whether this is adequate for mass terms: are mass terms (substances) and count terms (individuals and collections of individuals) all that there are, or are these too realised by something else, such as physical matter? If they are all that there is, then they are presumably taken to be the real extensions of things. They cannot themselves have additional structure: in addition to mud being part of muddy water, atoms of hydrogen must also be part of water. Homogeneous properties predicated of muddy water (or water) must then also be predicated
of the mud in the muddy water (or the hydrogen in the water). This leads to the familiar problem sentence (attributed to Emmon Bach):

The snow making up the snowman is quite new but the H₂O making it up is very old.

The problem could be addressed by adopting a richer ontology, more like that proposed by Parsons: rather than have count objects constituted by substance, both count objects and substances could be constituted by physical matter [Parsons, 1970]. I discuss these issues in more detail later.

At first glance, Link’s theory can account for the fact that mass terms can only be counted if they are first quantised. However, any arbitrary fusion of substance can be taken to be a “quantity”. Thus “the three quantities of water” can refer to any bit(s) of water: it is always possible to divide a quantity of water into three quantities. It seems to me that “quantity of water” has more intensionality than it is attributed in Link’s theory. We can refer to “the quantity” itself, rather than “the beer” in “the quantity of beer that I drank”. The syntactic structures which quantise substances seem very much like group referring noun-phrases, which are widely accepted as possessing some variety of intensionality [Link, 1983, Landman, 1989, Barker, 1992].

Lönning has a theory for mass terms where he tries to account for the semantics of sentences involving mass terms, negation, and non-logical determiners like “most”; “much”; “little”; “less than two kilogrammes of” [Lönning, 1987b]. He notes that the semantics will be seriously in error if we adopt the set theoretic approach — as in [Montague, 1973, ter Meulen, 1980] — and take “not p” to refer to the complement of p, and use the cardinality of sets to determine the truth of sentences involving non-logical determiners. If “some water disappeared”, modelling negation as complement can entail “the water did not disappear”. If mass terms, modelled as sets of quantities, are potentially infinitely divisible, then the cardinality of quantities will be the same. Lönning makes mass terms denote elements of a Boolean algebra, which, together with generalised quantifier-like interpretations of determiners, allows him to address the representation of non-logical determiners and negation.²⁷

²⁷Roeper has a not-dissimilar theory [Roeper, 1983], except that the logical connectives are non-standard, and the relations he covers are more general than those that can be represented by Lönning’s theory. Lönning, in a talk at the Third European Summer School in Language, Logic and Information, at Saarbrücken, August 1991, expressed doubts as to whether these more general relations appear in English. He can represent relations R provided they satisfy the constraint: if (a, b) ∈ R and (c, d) ∈ R then (a, d) ∈ R: all relations between mass term generalised quantifiers effectively relate the generalised quantifiers independently to some other fixed term. English appears to have this behaviour, taking an example from Lönning: “all water is denser than some alcohol” can be paraphrased as “there is a degree such that all water is denser than it and some alcohol is not denser than it”. All of Roeper’s examples have this behaviour.
Lønning adopts an *homogeneous constraint* as an heuristic principle, similar in motivation to Bunt’s notion of homogeneous reference. He recognises that mass terms may not actually refer homogeneously, but that they behave as if they do. He expresses his constraint as:

Mass noun phrases combine only with *homogeneous expressions* to form sentences.

This can be viewed as a test for mass nouns. Lønning takes δ to be an homogeneous expression if the following distributive and cumulative inferences hold:

\[
\begin{array}{c|c}
\text{Cumulative:} & \text{Distributive:} \\
\hline
\beta \alpha \delta & \alpha \delta \\
\text{not-} \beta \alpha \delta & \text{There is some } \beta \alpha \\
\end{array}
\]

where β is a predicative or intersective adjective.\(^{28}\) The adjective β must itself be an homogeneous expression for it to combine with a mass term. Note that the expression of distribution used by Lønning effectively restricts distribution into a noun phrase, involving the term α, to those subquantities that are also α. We might like to use inference rules similar to this in order to control distribution into mass terms: if we have “much water δ”, then the subquantities of “much water” that are themselves water, have the property δ.

Lønning thinks that noun phrases involving definite descriptors should be distinguished from mass noun phrases as they can combine with inhomogeneous expressions. The sentence:

\[
\text{Much water weighed two grammes.}
\]

is ill-formed, according to Lønning, because “weighed two grammes” is inhomogeneous, yet:

\[
\text{The water that John drank weighed two grammes.}
\]

is a sentence, thus “the water that John drank” should not be taken to be a mass noun phrase. As with plurals, we might still argue that the distributive behaviour is given by the verb phrase, not just the noun phrase, as definite descriptors of otherwise mass terms can combine with homogeneous properties and allow distributive (and cumulative) inference. In Lønning’s theory, homogeneous expressions denote elements of the Boolean algebra: they will hold of a mass term if they dominate it in the algebra. Inhomogeneous expressions denote subsets of the Boolean algebra. As an example, the definite descriptor “the p” will denote the subset consisting of the supremum

\(^{28}\) Lønning assumes an extensional treatment of adjectives, where βα is to be taken to be the intersection of those things that are β and α.
of $p$. Mass noun phrases, like “much water”, denote subsets of the algebra. Thus mass noun-phrases with non-logical determiners are treated as properties that will hold of the elements that satisfy them. Definite noun phrases denote elements. An homogeneous property holds of such an element if it dominates the definite descriptor in the Boolean algebra.

It may be interesting to see if there are examples of non-cumulative verbs that combine with mass terms, like those given by Schwarzschild, where some implicit quantifier seems to be involved [Schwarzschild, 1990].

I shall use insights from Lenning’s proof-theoretic notion of distribution. Taken together with Landman’s notion of intensional individuals [Landman, 1989], a theory can be developed along the lines rejected by Roepen [Roepen, 1983] where:

All dirty water is water.

gives rise to the expression:

$$\forall x (\text{dirty-water}'(x) \rightarrow \text{water}'(x))$$

so that only those terms that have the relevant property (being dirty water) are effectively distributed to.
2.4 Ontology

Here, I shall present some ontological considerations relevant to the formalisation of nominal expressions. I shall use the terms thing, object or individual to refer to the extensions of non-abstract count nouns; substance corresponds to the extensions of non-abstract mass nominals; manifestation, physical mereology or physical matter refer to the actual stuff, out there in the world. The considerations that follow discuss whether any of these three — objects, substances, and physical matter — can be equated.

Let us initially consider whether two different things (objects) have the same physical manifestation. Carter suggests that different objects might have the same manifestation, provided that they are of different kinds [Carter, 1990]. Presumably, in the case of plurals, a deck (of cards) may be of a different kind to some cards. One criterion for judging whether objects are of the same kind might be to see whether they have the same elements, or nice parts, where nice parts of a mereological fusion of cats, are cats, as opposed to bits of cats [Lewis, 1991]. Plural objects with different nice parts may then have the same physical manifestation. It does not necessarily follow that objects with the same elements, or nice parts, must be equated, as I shall argue later when discussing whether substance should be distinguished from physical matter.

Should objects and physical matter be equated? As objects outlast a particular physical make-up, the answer is surely ‘no’. If we consider the metaphysics of change, as it applies to the count terms, then we can see the problems in taking count terms to be the sums of their physical parts. A cat may lose a whisker, but it remains the same cat. A person, over the period of a lifetime, has few, if any, physical parts in common with the person they were at birth, yet they are the same person. These are the familiar arguments from Aristotle [Burge, 1975].

Should substances (denoted by mass terms) be distinguished from physical matter? Parsons suggests that they should [Parsons, 1970]. He cites the argument, for which he credits Quine, that:

“even if all and only furniture was composed of wood, it would not follow that wood = furniture, since parts of the chairs might be wood without being furniture.”

Bunt disputes this example, suggesting that the problem lies in taking “furniture” to be a mass term with minimal parts. He cites the same example, replacing the individuating term “furniture” with “sawdust”, so that all the wood in the world is made into sawdust. He then says that “wood” and “sawdust” can legitimately be equated, in non-intensional contexts [Bunt, 1985]. However, as mentioned earlier, this cannot be the case, as we might truthfully say “this sawdust was made yesterday”, without commitment to
“this wood was made yesterday”. If this is ruled inadmissible in an extensional theory, as it is some ‘intensional context’, then we might question what a ‘non-intensional context’ could be. Quine’s argument in this case would presumably be that the sawdust and wood are still distinguishable because they have different nice parts: there are parts of the sawdust that are not sawdust, but are wood. However, it is conceivable that some sawdust might be sufficiently finely ground for its minimal nice parts to constitute the minimal nice parts of wood, but we might still dissent from the view that the sawdust is the very same thing as the wood. It might be argued that the fusions can be distinguished because the properties are genuinely different — even on a possible worlds analysis — as the extensions of the properties vary differently with time: though they may now have the same extension, they are different properties because they had different extensions. Despite this, we must still favour an intensional ontology: if we point at the wood/sawdust and say “this sawdust was made yesterday” then, with an extensional ontology, the stuff that “was made yesterday” is the extension of both “the wood” as it is now and “the sawdust” as it is now. No matter that the properties once had different extensions, the stuff being predicated of will be the same, unless we adopt the view that “the wood” and “the sawdust” denote different things now, and hence have an intensional ontology.29

Bunt interprets Parsons [Parsons, 1975] as saying that in the case of the wood-furniture example, it is not the role of a semantic theory to make such distinctions. One may then question what the role of a semantic theory is. I believe that Parsons’ point was that it is not the role of a semantic theory to dictate essentially philosophical choices. This is an argument for philosophical neutrality, whenever possible, not for getting the semantics wrong.

Consideration of change provides an additional argument in favour of the absence of identity between substance and physical matter. If we have some water and part of it evaporates, we might say it is the same water that we started with. It is possible to say “this is the water I pumped from the well yesterday. I spilt some of it”. The continuation suggests that physically, “this water” and “the water I pumped from the well” must be different, as I spilt some of the latter. We can also say “I spilt some of this water”. So, “this water” at some time previously had a greater physical extension than it does now, yet to be able to refer to it as “this water” seems to imply there is some individual representing “this water” which can survive physical mereological change, just as objects in the count domain can. Thus, although these are perhaps not conclusive arguments, care should be taken in assuming that the semantics of natural language should take theories of physics into account. The mereology used for representing mass terms might not comply with the

29 An intensional ontology creates a problem for the semantics of demonstratives: how do we know what is being indicated when different objects can have the same physical extension? Landman suggests that we never really indicate a thing in itself with a demonstrative [Landman, 1986b].
behaviour of the physical extension. That is to say, some dirty water may be the same dirty water tomorrow, even if the dirt is in dynamic equilibrium with some sludge at the bottom, and even if some water evaporates, and it still has parts that are dirt and water. Thus we can have a part-whole structure which survives physical mereological changes. A similar argument is presented by Cartwright [Cartwright, 1965]. This more intensional view of the part-whole representation may allow a count individual to be equated with the sum of its parts, as those parts survive physical mereological change.

Should substances be distinguished from objects? This question has at least two interpretations:

(i) Are substances of the same order as objects — both manifest by physical stuff, neither subsuming the other?

(ii) Are the fusions of substances to be equated with particular objects (at a certain time) — when we have a gold statue, is the statue (at a particular time) the same thing as the gold [Grandy, 1975]?

Considering the first interpretation, Laycock seems to suggest that substances are of the same order as objects [Laycock, 1972, Laycock, 1975], as both substances and objects seem capable of surviving physical change [Cartwright, 1965]. As to the second interpretation, if we take the view that stage level [Carlson, 1977] count nominals are akin to mass terms [Chierchia, 1982a, Chierchia, 1982b] then it might be tempting to equate particular objects with fusions of substances. It could be worth examining examples of stage level predication, and seeing whether they lack, or do not require, the relevant notion of intensionality. There are problems with equating stage level predication with predication of either substance, or physical manifestation. If it is agreed that stage level predication is akin to predication of (a representation of) some physical manifestation, and present progressives are taken to exemplify such readings, then the intensionality of the so-called group readings would collapse: if committee A has the same members as committee B, then “committee A is meeting” would imply “committee B is meeting”, and perhaps further, that all the members are in attendance. But neither of these conclusions are necessarily the case. Even if we use some ideas from Landman [Landman, 1989], and take the stage level committees as their physical manifestations under the role of being a certain committee, there are still problems: “committee A is being disbanded at this moment” appears to be a stage level predication, yet it is surely the committee itself that is being disbanded, not the members under the role of being committee A members.
Chapter 3

My Approach

The Issues Addressed

The most significant issues in the literature on plurals which I address concern the treatment of ‘non-denoting’ descriptors (in a manner compatible with a theory of presupposition); representing Landman’s individuals under guises, or roles [Landman, 1989], without creating a universe of individuals under all possible roles; and controlling distributive inferences with mass terms, using insights from Landman, without making the assumption that the extensions of mass terms are homogeneous (either directly, or by way of atomicity). This more general treatment of distribution permits a uniform representation of mass and count nominals.

These issues may not seem to be central issues in linguistic semantics: they do not necessarily add significantly to the NL coverage of formal semantics. Perhaps the major point of the thesis is a formalist one, concerning the appropriate approach to producing formal semantic theories: the thesis presents an essentially axiomatic formal semantics of plurals and mass terms, as opposed to the more usual model-theoretic treatment. It is the axiomatic method that leads to any novelty in the coverage of non-denoting terms, as it allows the theory to be incomplete in a useful manner: an axiomatic system can remain silent on issues which are irrelevant for the purposes of the theory. When the objective of an axiomatic system is to capture certain intuitions, if these become unclear on issues that are independent of the core points we wish to capture, then the axioms can remain silent (incomplete) with respect to these more debatable points. In contrast, a model-theoretic treatment usually demands a complete theory, where all the issues are decided. This distinction is elaborated below.

Axiomatics

In a model-theoretic formal semantics of natural language, NL expressions are mapped into a representation language, and this is interpreted in some
set-theoretic model (see [Dowty et al., 1981] for example). This has the
disadvantage of not allowing the representation to be used directly in some
proof theory.

When examining new phenomena in NL, the model-theoretic approach is
to look for some mathematical structure, or system which appears to display
the appropriate behaviour. This structure is then used as a model for the
formal semantics of NL. The representation is made to adopt the behaviour
of this model by asserting that all the entailments of the model are legitimate
inferences in the representation (completeness), and all inferences in the rep-
resentation are supported by the model (soundness). By insisting that the ent-
ailments of the representation are precisely those of the model, there is a
danger that artifacts of the model may lead to too strong a logic, for example,
if the model employs a type hierarchy to avoid self application paradoxes,
as in Montague semantics, then the constraints of the strongly typed regime
are inherited by the semantic representation, even if not justified by (or even
counter to) considerations of NL itself. Demanding completeness requires all
issues to be decided one way or another, even when our intuitions provide
no useful insight.

It seems better first to establish the behaviour of NL — in this case, NL
plurals and mass terms — and then axiomatise this in a suitable framework.
These axioms together with a suitable inference rule in the logic, such as
modus ponens, allow the representation to be used directly for performing
proofs. A model that satisfies these axioms can be found later, merely to
show that the theory is consistent. The framework I have chosen is PT, an
axiomatic property theory [Turner, 1992] based on Frege Structures [Aczel,
1980]. This is a weak, highly intensional theory. The fine grained intension-
ality is suitable for capturing the intensional aspects of NL. The weakness
of the theory means that no implicit assumptions are made concerning valid
NL inferences. The theory can be strengthened with axioms that satisfy our
intuitions, as required. Only those aspects of behaviour of which we are cer-
tain need be axiomatised. Any term whose status is in doubt can be left
unanalysed. It is this approach that is adopted in representing non-denoting
terms, and predication of non-denoting terms.

Outline of Thesis

After presenting the property theoretic framework (Chapter 4), I initially
axiomatise just the plurals (Chapter 5). This gives a clear indication of how
the axiomatic method can be used in NL representation — with particular
regard for non-denoting terms — without complicating the picture with some
of the more unusual constructions used in the final theory. The formalisation
closely follows Link’s axioms [Link, 1991a] given in §2.2. I show how an
axiomatic approach to non-denoting definite descriptors might be embodied
into a formal theory of presupposition.
CHAPTER 3. MY APPROACH

Then I consider a slightly weaker axiomatisation of a part-whole structure (Chapter 6) which does not make any assumptions about atomicity (minimal parts). I shall use this as a basis for the extensions of nominal expressions in the final theory (Chapter 8), which will allow a unified treatment of mass terms and plurals.

This final theory controls distributive inferences with mass terms, without assuming atomicity or homogeneous extensions, by using a development of Landman's ideas concerning individuals appearing under some guise, or role [Landman, 1989] (Chapter 7). As this copes with distribution in the case where there are no atoms, it can also cope with distributive inferences with count terms, where there appear to be atomic individuals. The theory can be embellished with a notion of singular and plural, and offers a treatment of count and mass terms within a unified framework, as well as giving a property modifier formalisation of Landman’s roles. This theory does not cover distributive inferences into conjoined noun phrases in order to avoid some combinatorial complexity in the presentation of its central function: that of how to control distribution into mass terms.\(^1\)

Non-denoting Terms

My theory shows how the treatment of non-denoting definite descriptors can be given an adequate treatment using property-theoretic semantics. The principle points are whether such descriptors can denote some bottom element in a semantic theory, and what predication of such descriptors results in. In the mass domain, Parsons argues that non-existent substances may be different, yet if we have an extensional treatment, and say they both denote bottom, then we are perhaps forced to equate them [Parsons, 1970, Parsons, 1975]. If the existence of a bottom element in the theory is rejected, then some other treatment must be given. A free logic treatment would allow us to say that such descriptors denote elsewhere [van Fraassen, 1966, Link, 1991a].

The result of predication of such terms is undefined. Predication of non-denoting terms can be viewed as a variety of category mistake. Property theory has been axiomatised to come to terms with much more awkward examples of category mistakes, namely the logical paradoxes which result from unconstrained self predication. The property theory I shall use, PT, avoids the paradoxes by having axioms too weak to prove propositionhood of such sentences, and only those things which are propositions may have truth values (PT rejects bivalence). A similar position is open to us with non-denoting terms. We may have a class of denotable objects. If definite descriptors are represented as suprema of denotables, then we can require

\(^1\)Distribution across conjunction could be added either as an extension to this theory, or as part of a theory of ellipsis. If the grammar sketched in Appendix B — which gives a syntactic treatment of ellipsis — is adopted then, in combination with the final theory, distributive inferences with conjoined plurals and mass terms are covered.
(in the axioms) that such suprema are themselves denotable only if they are non-empty. This achieves the effect of ‘denoting elsewhere’ in free logic. The undefinedness of predication of such terms can be obtained by requiring that all natural language derived properties are properties of denotables, and can only provably form propositions with terms that are provably denotable.

This is compatible with a theory of helpful answers [De Roeck et al., 1991], where an answer is obtained by attempting to construct a proof of the putative proposition underlying a question. If the proof fails, a helpful answer can be constructed from information concerning where the failure occurred. On answering questions with non-denoting definite descriptions, the same method can be used to indicate why an answer cannot be given. So with the question:

Is the present king of France bald?

the helpful answer:

There is no present king of France.

can be constructed from the failure to prove that the underlying assertion is true or false, indicating that the question wrongly presupposes the existence of a present king of France. The treatment of non-denoting definite descriptors is, in effect, a limited treatment of the more general notion of presupposition.

Distribution with Mass Terms

Initially it may seem attractive to take a mereology [Leonard and Goodman, 1940], devoid of intensional burdens, as the correct treatment of mass terms, as in [Grandy, 1975, Lønning, 1987b, Moravcsik, 1973] for example. If we follow Bunt’s homogeneous reference hypothesis, repeated here:

“A mass noun refers in such a way that no particular articulation of the referent into parts is presupposed, nor the existence of minimal parts.” [Bunt, 1985]

then it would seem that the representation theory should not presume the existence of atoms. This is not the same as presuming that there are no atoms: mass terms might just be silent on this issue as it is not a relevant concern of the use of such terms. For example, if we wish to remain ontologically neutral, then it may be the case that the denotations of mass terms are equated with physical substances, which are atomic according to modern physics. In PT we can model this by constructing axioms that are too weak to decide this issue: atoms may exist but the axioms will be too weak to prove their existence.
CHAPTER 3. MY APPROACH

We may wish there to be distributive properties. However, we surely require a means of blocking distribution at some appropriate level: from “dirty water is water” we do not want to be able to prove that all parts of dirty water, without restriction, are water, as we could then infer “some dirt is water”.

In the plural domain, distribution can be blocked by the use of atoms, or minimal parts, so that only nice parts are recovered. In the mass domain it seems harder to adopt primitive atoms, as it is clear that which ‘atoms’ should be used depends upon the subject noun phrase. What is a nice part of “dirty water”? It must be different from a nice part of water, or dirt. But does dirty water have water and dirt, as mereological parts, or are these nice parts like those often used in the plural domain, independent, in this case, from parts of dirt and water?

As it is the restriction of inferences in which we are interested, perhaps we should re-examine a rather simple-minded approach to distribution, in which the distributive inference, which enables us to infer that any (claret) part of claret is wine from:

All claret is wine.

is achieved via some rule of the form:

\[ \forall x (\text{claret}'(x) \rightarrow \text{wine}'(x)) \]

where \( x \) ranges over denotables. This effectively says that claret is, in some sense, part of wine. Note that I am not proposing this as the representation of this sentence, merely a formula derivable from the axioms which enables the distributive inference to take place, unlike in Link’s theory of mass terms [Link, 1983].

This treatment does, on the face of it, have some nice properties. If we say that:

All dirty water is water.

then only those denotables of which we would predicate “dirty water”, are to be counted as “water”. This apparently avoids any question as to the relevant parts of dirty water: we do not even contemplate how to avoid predicating “water” of the dirt in dirty water. Dirty water is “part of”, or a subclass of, water, yet the dirt itself is not.

This approach has been discredited [Roepet, 1983] because of examples such as:

\[ \text{If I were examining only mass terms, then I might take this to be the representation of the sentence, as the distributive inference can always occur with mass terms in conjunction with the quantifiers “all” and “some”. However, this would imply that distributive behaviour obtains as a result of the representation of the noun phrase, counter to the evidence from the plural domain.} \]

\[ \text{Hans Kamp has pointed out to me that this may not be adequate: taking the sentence “all water is liquid”, the solution to controlling distribution that I propose would allow} \]
CHAPTER 3. MY APPROACH

All phosphorus is red or white.

There are some quantities of phosphorus that will be both red and white, and so it would be erroneous to say the sentence entails:

\[ \forall x (\text{phosphorus}(x) \rightarrow \text{white}(x) \lor \text{red}(x)) \]

An adequate account of distributive behaviour is tied up in the reduction of "red or white". We could reinterpret this as:

quantity such that it is red or it is white, or part of it is red, and the remainder is white

In the second part of this definition, there is a problem in restricting the parts that are considered: we want to take parts that avoid splitting any parts which count as phosphorus into parts that are not phosphorus. For non-abstract terms, we might adopt the physical atoms as minimal parts, and allow distributive properties to apply to these minimal parts. This leads us into trouble as a minimal part of dirt water, for example, will not be a physical atom: depending upon the nature of the subject noun used to refer to some denotable, the relevant minimal part may vary.

To be able to consider the appropriate nice parts, we must be aware of the predicate used to select the quantities. However, in the standard, compositional semantics of the above propositions, the subject is lost when considering the semantics of "red or white". Intuitively, we need an ability to represent denotable terms under different guises. One possibility would be to pair a denotable with the property used to conjure it up. Then all the relevant nice parts could be constructed. The phrase "the furniture" could be represented as something like "the fusion of that which is furniture, as furniture" and "the wood" as "the fusion of that which is wood, as wood". No matter whether the two fusions are equated, the appropriate nice parts could always be found. This would avoid the need to make the representations of substances homogeneous, so muddy water could have parts that are mud, for example.

Landman has considered the problem of representing individuals under roles, within a theory of plurals [Landman, 1989]. It will be fruitful to examine his theory, and reappraise it within PT. To summarise briefly, his representation involves intensional objects corresponding to terms under all their possible guises. Thus he may have "John, as himself", "John, as a judge" and "John, as a cleaner". John, as a cleaner, might be on strike, even the inference that anything that is water is liquid. Yet whilst a single water molecule is conceivably "water", it would seem possible that "being liquid" is a property of a body of water (or collections of bodies of water) not of individual molecules (or collections of discrete molecules). If these views of the meaning of "water" and "being liquid" are accepted, then my approach to distribution is too strong. I offer a defence in Chapter 8.
though John, as a judge, is not. Landman argues for modifying the individuals, rather than the properties, for example, he would have “John-as-a-judge is on strike” as opposed to “John is a striking-judge”. I show that it is possible to treat his examples with property modifiers, in particular with the semantic analogue of subsective adjectives [Kamp, 1975], without requiring a universe of intensional individuals.

Landman does not use these individuals under roles to control distribution, but a representation enriched with guises does allow such a possibility. We could represent “liquid” with a distributive property. The sentence “the dirty water is liquid” would then allow us to infer that any part of the dirty water \textit{which is itself dirty water} is liquid. Potentially equal fusions of distinct properties can also be distinguished. If all the wood is turned into sawdust, and if we take the fusions of wood and sawdust to be identical, the sentence “the sawdust was made yesterday” can be interpreted as “the sawdust was sawdust-made-yesterday”. The undesired entailment of “the wood was made yesterday” no longer arises.

This means of axiomatically controlling distribution is implicitly suggested by Lønning in his test for homogeneity [Lønning, 1987b] given in §2.3. If adjectives are represented with property modifiers, then we can construe his test as the following inference rule:

\[
\begin{align*}
\forall x(\beta \alpha x \rightarrow \alpha x) \\
\forall x(\alpha x \rightarrow \beta x) \\
\exists x(\beta \alpha x) \\
\vdash \forall x(\beta \alpha x \rightarrow \beta x)
\end{align*}
\]

For example:

[assuming that all hot coffee is coffee]
all the coffee is liquid
there is some hot coffee
\[\vdash \text{all the hot coffee is liquid}\]

thus, generalising this, ‘being liquid’ distributes to all the parts of the coffee that are coffee. If the coffee contains sugar, this does not allow us to infer that the sugar is liquid, even if it is ‘part of’ the coffee. All that is required are axioms that, when together with \textit{modus ponens}, support the above inferences. This not only allows us to consider distribution with mass terms, but will also enable us to distinguish predications of equated suprema, when the suprema are presented in terms of different properties, and can provide an account of plural distribution within a unified theory.

\section*{A Unification}

With plural terms, it is often taken that certain objects, or individuals, which can be referred to by singular count nouns, are atomic. Distributive properties then distribute to these individuals, and no further.
CHAPTER 3. MY APPROACH

If we were to equate stage level individuals [Carlson, 1977] with fusions of substance [Chierchia, 1982b], then Landman’s roles would provide a means of distinguishing the fusion as substance or as individual. Thus, John as a stage level individual would be the physical manifestation of John, as John. The role would also indicate the appropriate level of “atomicity” to be used in the case of distributive properties.

If we allowed dirty water to denote a mereology that has parts consisting of dirt and water, and if we can give a semantics whereby the relevant nice parts, namely those that are dirty water could be recovered, then it would also be formally possible to treat the stage level individuals in the count domain as mereological entities, with parts. That is, singular count terms could, in principle, denote entities equal to the mereological sum of their parts. If we chose an ontology where stage level individuals were equated with fusions of substance then this theory would allow a unification of the stage level count and mass domains without semantic error, because it would be possible to distinguish predication of stage level individuals from predication of sums of parts. Such a simplification, however, should not be taken to be desirable just because it is possible.

Rather than force the issue, I shall adopt a representation which allows nice parts (either count or mass) to be equated to the sum of their parts — so that a sum of nice parts would have some not very nice parts — but does not force it. If we take the example of the new gold ring, made from old gold, then I do not force the object which is the gold, and that which is the ring to be equated. The representation will, however, allow them to be equated, if so desired, without resulting in a contradiction. Similarly, the theory will allow “the Evening Star” and “the Morning Star” to have the same extension (Venus), but still enable the descriptors to be distinguished, giving the potential for useful formal interpretations of sentences like:

The Morning Star and the Evening Star were once thought to be different.

It might become apparent that many of the formal devices employed become redundant should a highly intensional ontology be adopted, where all nominals necessarily have different extensions if they are ascribed different properties in natural language. However, to adopt a particular ontology for no other reason than to make some presupposed formal representation obtain the correct inferences is to allow a perhaps inappropriate choice of formal semantics to dictate the choice of ontology. There do seem to be instances where the requirements of a semantic theory and an ontology appear inextricably linked. Committees, for example, must surely have some independence from their members, even with stage level predication. But there are cases where semantic issues need have no bearing on ontological considerations, such as the question of whether the substance phosphorus has parts which
are not phosphorus. I shall strive to take a neutral, or weak view on this issue throughout the thesis.

Other Issues

The basic theory of plurals I develop and the more complex final theory — which unifies the treatment of plurals and mass terms — are not concerned with representing the various possible intermediate distributive readings. I formally only elaborate upon the fully distributive, and the fully collective readings. The collective readings are assumed to subsume any intermediate distributive readings. I find Schwarzschild’s arguments — detailed in §2.1 — against structured-group representations to be convincing. The operator that models conjunction is thus associative, and the operator that models the definite determiner also does not add structure to the representation of nominal terms. For simplicity, however, I do not adopt his proposed context sensitive distribution operator which produces appropriate intermediate interpretations, nor do I offer a reductive interpretation of reciprocals.

The final theory in Chapter 8, and the grammar in Appendix B, are only intended to cover sentences with intransitive verbs. Adding coverage of transitive verbs would add the complications of quantifier scoping ambiguity. Further, there appears to be little empirical data on the acceptable combinations of distributivity and collectivity in sentences with more than one noun phrase.

The arguments for committee-like objects to be intensional, and distinct from their members is strong. To account for the different interpretation of sentences with these terms, it would be fairly straightforward to add ideas from Barker’s theory [Barker, 1992]. To avoid over-complicating the theory, I do not do so.

The theory that I shall use for representing the semantics of NL is essentially first order. This means that considerations of how the scoping of accepted readings of sentences affects the logical power of English cannot be considered within this theory. The problem of undecidable semantics arises when the semantic theory allows arbitrary fusions of (potentially infinite) collections which cannot be denoted by NL nominals. In the theory to be adopted, only those collections that are in the extension of a denotable property may be considered. Thus, the issue of undecidability does not arise.
Chapter 4

Propositions, Properties, and Truth

In this chapter, I present a theory of propositions, properties and truth (PT) due to Turner and Aczel [Turner, 1992, Aczel, 1980]. Before giving the general framework and the formal details of PT, I shall try and motivate why such a theory is useful for the semantics of natural language. The remainder of the thesis is concerned with strengthening PT in various ways in order to represent plurals and mass terms. In the appendix, a model of PT with plural terms is given to demonstrate the theory’s soundness.

Many workers in this field use an extensional logic as a basis for their representations ([Link, 1991a, Lønning, 1989] for example). In general, such extensional representations are inadequate for the semantics of NL (see [Chierchia, 1982a, Chierchia, 1984, Chierchia and Turner, 1988]). Intensionality is apparent in the case of opaque predication. The notion of possible worlds is often introduced into the representation language in order to account for this phenomenon. In such approaches, a proposition is a set of possible worlds, and a property is a function from individuals to propositions. Possible worlds can be used to analyse modal notions such as possibility and necessity. Given an accessibility relation, possibly p is the set of worlds that have an accessible world in p. The possible worlds analysis has been extended to model knowledge and belief [Hintikka, 1962, Kripke, 1963]. This set-theoretic account fails on certain instances of propositional attitudes: an agent may know a certain proposition p, and there may be some proposition q, denoting the same set of possible worlds, yet we might not want to conclude that the agent knows q. The possible worlds account of knowledge and belief forces us to this unwanted conclusion. Such propositions are typified by those involving mathematical truths.

A further problem with possible worlds models is that they are typically strongly typed. Properties and relations can only hold of objects of a specific type. To generalise semantic notions across different types, type lifting, or type shifting has to be employed [Partee and Rooth, 1983] within the seman-
CHAPTER 4. PROPOSITIONS, PROPERTIES, AND TRUTH

The strong typing bars self-predication and so avoids the paradoxes involving self-reference, such as the Liar: “this sentence is false”. However, the strong typing also disallows unproblematic cases of self-predication, such as “this sentence is six words long”, and prevents the expression of universal properties.

These problems can be overcome if we treat propositions, and properties, as primitives. A property like “red” is not the set of red things, it is just itself, the property of being red. Similarly, the proposition, “2 + 2 = 4” is not merely a truth value, or a set of possible worlds, but it is a basic object, different from “$e^{i\pi} + 1 = 0$”, even though, from the laws of mathematics, these propositions must always be true together.

Such a language, with a highly intensional notion of properties and propositions which avoids the paradoxes without banning self-predication through strong typing, is exemplified by PT, Turner’s axiomatisation of Aczel’s Frege Structures [Turner, 1992, Aczel, 1980].

General Framework

Conceptually, PT can be split into two components, or levels. The first is a language of terms, which consists of the untyped $\lambda$-calculus, embellished with logical constants. A restricted class of these terms will correspond to propositions. When combined appropriately using the logical constants, other propositions result. As an example, given the propositions $t, s$, the conjunction of these, $t \land s$, is also a proposition, where $\land$ is a logical constant.

Some of the propositions will, further, be true propositions. When combining propositions with the logical constants, the truth of the resultant proposition will depend upon the truth of the constituent propositions. Considering the previous example, if $t, s$ are both true propositions, then $t \land s$ will also be a true proposition.

There may be terms that form propositions when applied to another term. These terms are the properties. The act of predication is modelled by $\lambda$-application.

The essential point to note is that this is a highly intensional theory as the notion of equality is that of the $\lambda$-calculus: propositions are not to be equated just because they are always true together; similarly, properties are not to be equated just because they hold of the same terms (i.e. form true propositions with the same terms).

There are problems with the theory so far: the logical constants have no proof theory; and the notions of being a proposition, or a true proposition, cannot be expressed within this language of terms. That is, although we can consider terms as propositions, or true propositions, and comprehend how the proposition-hood and truth of a term depends upon the proposition-hood and truth of its constituent terms, we cannot express these notions formally within the language of terms: some meta-language is required. This is the
purpose of the second component of PT: the language of well formed formulae (wff). This is a first order language where the terms (the objects which can be quantified over) are those of the \( \lambda \)-calculus extended with logical constants, as discussed above. The language of wff has two predicates, \( P \) for ‘is a proposition’, and \( T \) for ‘is a true proposition’. Clearly, this gives the formal means for axiomatising the behaviour of propositions and true propositions. For example, the informal discussion concerning the behaviour of the logical constant \( \land \) can be formalised as follows:

“given the propositions \( t, s \), the conjunction of these \( t \land s \) is also a proposition”:

\[
P(t) \land P(s) \rightarrow P(t \land s)
\]

“if \( t, s \) are true propositions, then \( t \land s \) will also be a true proposition”:

\[
P(t) \land P(s) \rightarrow (T(t \land s) \leftrightarrow (T(t) \land T(s)))
\]

Axioms concerning \( T \) must be restricted so that only terms that are propositions are considered.

The distinction between wff and terms can be taken to be akin to that between extension and intension in Montague semantics [Dowty et al., 1981]. In that theory, however, intensions are derived from extensions.\(^1\) As a consequence, the equality of intensions is that of the extensions, so propositions will be equated if they are always true together, and properties will be equated if they hold of the same objects. This is in contrast to PT, where the intensions are basic. Propositions in the language of terms may have the same truth conditions when \( T \) is applied, but this does not force them to be the same proposition, so we might have:

\[
T(s) \leftrightarrow T(t)
\]

but that does not mean that the terms are equal:

\[
s = t
\]

Similarly, in the language of wff, properties may hold of the same terms, yet they may be distinct. The \( \lambda \)-equality of terms is thus weaker than the notion of logical equivalence obtained when considering truth conditions in the meta-language.

A useful parallel can be drawn between PT and Frege’s notions of sense and reference. We can take a proposition in the language of terms as corresponding to the sense of a statement. The referents of statements can be

\(^1\)In Montague’s intensional logic IL, an intension is given by applying the operator \( ^\wedge \) to an extension.
thought of as truth values, or truth conditions, which are obtained in the
language of wff by applying T to the proposition.

It is possible to give a proof theory for the language of terms without
recourse to a formal meta-theory (see [Hindley and Seldin, 1986] for example).
This might not be so useful for the semantics of NL, as the resultant theory
would not explicitly capture the extensional level.

**Formal Theory**

The following presents a formalisation of the languages of terms and wff,
together with the axioms that provide the closure conditions for P and T.

**Language of terms**

Basic Vocabulary:

- **Individual variables:** \( x, y, z, \ldots \)
- **Individual constants:** \( c, d, e, \ldots \)
- **Logical constants:** \( \forall, \land, \neg, \Rightarrow, \Xi, \Theta \)

Inductive Definition of Terms:

1. Every variable or constant is a term.
2. If \( t \) is a term and \( x \) is a variable then \( \lambda x.t \) is a term.
3. If \( t \) and \( t' \) are terms then \( t(t') \) is a term.

**The Language of Wff**

Inductive Definition of Wff:

1. If \( t \) and \( s \) are terms then \( s = t, P(t), T(t) \) are atomic wff.
2. If \( \varphi \) and \( \varphi' \) are wff then \( \varphi \land \varphi', \varphi \lor \varphi', \varphi \rightarrow \varphi', \neg \varphi \) are wff.
3. If \( \varphi \) is a wff and \( x \) a variable then \( \exists x \varphi \) and \( \forall x \varphi \) are wff.

The theory is governed by the following axioms:

**Axioms of The \( \lambda \beta \)-Calculus**

\[
\lambda x.t = \lambda y.t[y/x] \quad \text{y not free in } t \\
(\lambda x.t)t' = t[t'/x]
\]

This defines the equivalence of terms.

The closure conditions for propositionhood are given by the following
axioms:
CHAPTER 4. PROPOSITIONS, PROPERTIES, AND TRUTH

Axioms of Propositions

(i) \[ P(t) \land P(s) \rightarrow P(t \land s) \]
(ii) \[ P(t) \land P(s) \rightarrow P(t \lor s) \]
(iii) \[ P(t) \land P(s) \rightarrow P(t \Rightarrow s) \]
(iv) \[ P(t) \rightarrow P(\neg t) \]
(v) \[ \forall x P(t) \rightarrow P(\Theta \lambda x. t) \]
(vi) \[ \forall x P(t) \rightarrow P(\Xi \lambda x. t) \]
(vii) \[ P(s \approx t) \]

Truth conditions can be given for those terms that are propositions:

Axioms of Truth

(i) \[ P(t) \land P(s) \rightarrow (T(t \land s) \leftrightarrow T(t) \land T(s)) \]
(ii) \[ P(t) \land P(s) \rightarrow (T(t \lor s) \leftrightarrow T(t) \lor T(s)) \]
(iii) \[ P(t) \land P(s) \rightarrow (T(t \Rightarrow s) \leftrightarrow T(t) \rightarrow T(s)) \]
(iv) \[ P(t) \rightarrow (T(\neg t) \leftrightarrow \neg T(t)) \]
(v) \[ \forall x P(t) \rightarrow (T(\Theta \lambda x. t) \leftrightarrow \forall x T(t)) \]
(vi) \[ \forall x P(t) \rightarrow (T(\Xi \lambda x. t) \leftrightarrow \exists x T(t)) \]
(vii) \[ T(t \approx s) \leftrightarrow t = s \]
(viii) \[ T(t) \rightarrow P(t) \]

The last axiom states that only propositions may have truth conditions.

Note that the quantified propositions \( \Theta \lambda x. t \), \( \Xi \lambda x. t \) can be written as \( \Theta x(t) \), \( \Xi x(t) \), where the \( \lambda \)-abstraction is implicit.

The notions of \( n \)-place relations can be defined recursively:

(i) \[ Rel_0(t) \leftrightarrow P(t) \]
(ii) \[ Rel_n (\lambda x.t) \leftrightarrow Rel_{n-1}(t) \]

We can write \( Rel(t) \) as \( Pty(t) \).

As an illustration of how this theory addresses the paradoxes that can arise when self-reference is permitted, consider a predicate \( R \) whose extension is those predicates that do not apply to themselves. If such a predicate exists, then we can derive the paradoxical proposition:

\[ RR \leftrightarrow \sim RR \]

In PT we can define a term corresponding to \( R \) as follows:

\[ R =_{df} \lambda p. \neg pp \]

Trivially we can prove — from the definition of \( Pty \) and the Axioms of Propositions (iv) — that if we have a property \( p \), then \( Rp \) is a proposition, that is:

\[ Pty(p) \rightarrow P(Rp) \]
Thus, the Axioms of Truth can be applied to $Rp$ if $p$ is a property. The paradoxical proposition originally arose when considering $RR$. For it to occur in PT, the Axioms of Truth must apply to this term. For these axioms to apply, we must show $RR$ is a proposition. This in turn requires that $R$ is a property. This cannot be proven, as $R$ does not form a proposition with arbitrary terms, only with properties. Thus the paradoxical proposition does not arise. This exemplifies how PT allows unproblematic instances of self predication, whilst avoiding the category mistake that gives rise to paradoxical propositions.

In the next Chapter, I shall embody a part-whole structure within the language of terms, following the essence of Link’s theory [Link, 1991a] given in §2.2.
Chapter 5

Plural Terms in PT

If we ignore the so-called group readings and intermediate collective readings, following Link, then the question arises as to which of Link’s intuitions — formalised in §2.2 — do we accept in this property theoretic framework.

To represent sentences involving mass terms and plurals in PT, we can add terms and axioms for a Boolean structure to the language of terms, with summation and a supremum operator. However, there is no restriction on term formation in PT. Would we want, require, or allow, a Boolean/mereological structure involving all terms? It seems appropriate to have such a structure only amongst those terms which represent denoting natural language nominals. This can be achieved by restricting the scope of the relevant axioms. To have a Boolean structure on all terms is an unmotivated and unnecessary strengthening of the theory.\(^1\)

Link uses a free logic as he wants all his terms to correspond to plural objects: he does not want any term to correspond to non-denoting definite descriptors [Link, 1991a]. If he dealt with ‘non-denoting’ terms by making them denote a bottom element, and bottom was in the language of terms, then this would be like saying that there is a term which does not exist, but which is a part of all other terms. The only other option — in order to maintain a one-to-one correspondence between natural language denotable objects, and terms in the representation language — is to say that non-denoting terms genuinely do not denote in the theory, and there is no bottom element. Both of these requirements are met by a free logic with strict equality. In PT, however, we already have a range of terms which cannot correspond to objects referred to in natural language: the language of terms includes objects formed by arbitrary \(\lambda\)-abstraction and application, of which

\(^1\)An additional reason for having only a subclass of terms with a Boolean structure becomes apparent when considering the model theory of PT. A model of PT can be constructed using a lattice theoretic model for the \(\lambda\)-terms [Turner, 1990]. It is convenient to use the ordering and join in this lattice to model the summation and supremum operators. If, however, we give a Boolean structure to all terms, it is not possible to do this: no model of the lambda calculus could then be constructed.
only some will be of interest.\footnote{It is the objective of \textit{Frege Structures}, axiomatised by PT, to carve out those terms of logical interest, and give them structure. Similarly, when PT is used for the representation of NL, the formal theory delineates those terms of NL interest, and gives them structure. There is no reason for all terms to correspond to NL denotable objects.} We also have a theory designed to cope with the paradoxes. Given this, it seems prudent to deal with the intuitively much simpler kind of category mistake that appears to occur with non-denoting noun phrases within the existing theory, rather than produce a free version of PT.

Concerning a bottom element, it would seem that in natural language there is no object which we consider to be a part of all other objects. However, we must be careful of restricting the representation language — the formal theory — because of a restriction in the language being represented, as that would be to assume that all legitimate terms in the formal theory must play a part in the representation of NL. As mentioned before, we already have terms which do not correspond to objects referred to by natural language. Further, it will be possible to construct a bottom element using the supremum operator and inequality.\footnote{The supremum of the empty property: $\sigma x (x \not\in x)$.} To disallow such a term would require a free logic and a notion of strict equality. In the model of PT, the existence of a bottom element is mathematically convenient. The question should not be whether such a term is permissible in the representation language, but how — or whether — it features in the semantics of NL.

If the theory contains a bottom element, we may wonder whether any natural language term can denote it. We might view denotable objects as pieces of information. When such objects are conjoined, then the information in them is combined. The ordering relation is then an information ordering. Bottom is merely an absence of information. We could argue that non-denoting definite descriptors should be represented by the bottom element. This has the consequence that all non-denoting terms are equated, which can be seen as undesirable [Parsons, 1970, Parsons, 1975]. I believe that non-denoting definite descriptors do, indirectly, convey information. Namely, they are informative with regard to the preconceptions of a person who creates an utterance which involves them. They give an indication of the relevant presuppositions required to make sense of the utterance. The existential presuppositions depend upon the non-denoting descriptor. The representation language should not then force a collapse of all non-denoting terms.

Regarding non-denoting nominals, as the part-whole structure is intended to represent denotable terms, it seems clear that they should not, or at least, that we should not be able to prove that they are represented by such terms. When quantifying over objects, in propositions representing natural language sentences, it must be the case that we are only interested in existent, denotable terms. If everyone went to the party, we should not be able to conclude that the present king of France went to the party. It also follows
that we should not be able to prove that bottom is a denotable, as it does not represent a denotable term.

With a sentence involving non-denoting terms, we could take the Russel-lan view, and say that predication of a non-existent object gives rise to falsity. Lønning mentions the possibility of assigning the truth value \textit{false} to any application of a natural language derived property to a non-denoting term. Thus both “the present king of France is bald”, and “the present king of France is not bald” come out false [Lønning, 1989]. As mentioned in §2.1, this results in a contradiction, unless we rephrase \textit{false} here as \textit{not true}, and reject bivalence. The existence of an intermediate truth value is implicit in this account, as in free logic. Blau explores the possibility of assigning an explicit \textit{unknown} truth value to such sentences [Blau, 1981]. In PT we already reject bivalence: not all terms correspond to true, or false propositions. Unlike three-valued accounts of propositions, this is not at the expense of classical theorems, such as the law of excluded middle: if a term corresponds to a proposition, then it must be true or false, but truth values are only held by propositions. To prevent truth or falsity being assigned to sentences involving non-denoting terms, we can claim that they are not propositions (though they have the form of propositions), or, more weakly, we cannot prove that they are propositions. This theory, especially the weaker view, allows for a certain degree of presupposition modelling. If we model question answering as a proof production process [De Roeck \textit{et al.}, 1991], then when attempting to prove “the present king of France is bald” the proof will fail, because we cannot even prove that this is a proposition. The reason for its failure is that there is no present king of France.\footnote{Note that this does not deal with generics, such as “the unicorn has a horn”, which are not within the remit of the formal theory under discussion. If a treatment of generics is added, the arguments concerning “the present king of France” will still hold on both generic, and non-generic readings, as there is no \textit{kind} [Carlson, 1977] corresponding to that noun phrase. Mythological terms, as in “Pegasus is a horse” may be more problematic.} So, the helpful answer to “is the present king of France bald?” would be “there is no present king of France”.\footnote{This may be too weak: as has been indicated by Strawson, there are sentences involving non-denoting terms, to which some find it acceptable to ascribe a truth value. This is typified by passive constructions such as “The exhibition was visited by the present king of France” which some take to be a false proposition, unlike “The present king of France visited the exhibition” [Strawson, 1964]. One argument is that the existence presupposition only occurs for definite descriptors which are in the foreground, or active, in some sense. The theory below does not preclude a strengthening of the axioms to allow propositionhood to be proved in such cases, should a suitable formal theory of the foreground/background distinction be forthcoming. A simple-minded response could be to give a Russelian, quantificational representation of passive definite descriptors, or to allow the passive form of “visit” to form propositions with terms that are not provably denotable. In §10.2 a tentative treatment is offered for some of these examples.}

As can be seen, adopting the various views advocated model the effect of a free logic, without raising the question of what it means not to denote within
a theory. Natural language predication of non-denoting definite descriptors is consigned to the same dustbin as other unhelpful constructions, such as the liar sentence. It just becomes another case of a sortal category mistake. Axiomatic property theory avoids the paradoxes by having axioms that are too weak to prove that they are propositions. It is possible to adopt the same, weak, approach with regard to whether non-denoting definite descriptors denote bottom, and to the result of predication of a non-denoting term.

Lønning, in his consideration of the logical complexity of a representation with plural terms, notes that there may be some objects which cannot be denoted by natural language nominals, yet which cannot be excluded by consideration of the representation language in isolation. A formal theory of nominals (plurals or mass terms) can be said to be complete (in the lattice-theoretic sense) if the fusions of arbitrary collections of denotables — as opposed to denotable collections of denotables — may be considered as new terms. This gives rise to a second-order logic which cannot be given a general model [Orey, 1959, Lønning, 1989]. However, it is not clear whether natural language gives rise to (either allows or requires) quantification over objects — arbitrary sums — that cannot be referred directly, in natural language. Link requires that only those collections can be considered as a sum where there is a (natural language) definite descriptor of that collection. This is called definable completeness [Lønning, 1989, Link, 1991a]. Lønning examines other options. In the theory presented below, quantification is over the terms in the representation language. This automatically leads to definable completeness of the language of representation, with respect to its model. Further, to make sense of the semantics of nominals in this theory, the quantifiers contained in the representation of natural language sentences are effectively constrained to range over only those terms which represent natural language nominals.

**Language of plural terms**

To the basic vocabulary of terms is added:

- Predicative constant: \( \pi \)
- Operative constants: \( \oplus, \sigma \)

It is intended that \( \pi \) will hold of plural terms. The term \( \oplus \) is the summation operator. The term-forming interpretation of definite descriptors makes use of \( \sigma \). The term \( \sigma \lambda x. \varphi \) will be used to represent that object which is the sum of objects \( y \) such that \( T(\varphi[y/x]) \) holds (that is, the sum of terms that have the property \( \lambda x. \varphi \)). As with the quantifiers, this will be shortened to \( \sigma x \varphi \).

The language of wff is unchanged, and the theory is governed by the axioms of the \( \lambda \beta \)-calculus as before, together with the previous axioms for propositions and truth.
CHAPTER 5. PLURAL TERMS IN PT

Further Axioms For TP+II

The summation operator $\oplus$ is intended to be the semantic analogue of NL conjunction, in particular, nominal conjunction. To mirror the effect that in many cases NL conjunction is symmetric, that is:

John and Mary met.

has the same import as:

Mary and John met.

the operator $\oplus$ is also taken to be symmetric.

Axiom 5.1 The summation operator is symmetric (cf. Axiom 2.2).

$$a \oplus b = b \oplus a$$

In general, however, NL conjunction is not symmetric: consider ‘respectively’ sentences, and the temporal ordering often implicit in sentential conjunction. The semantics I give is intended to capture the static meaning of NL. Following Schwarzschild, I take any ordering effects to be part of the dynamic meaning, to be accounted for in some theory of discourse [Schwarzschild, 1990, Schwarzschild, 1992].

If ordering is to be ignored, then:

John and Mary and Harry

refers to the same complex term as:

Harry and Mary and John

The next axiom, taken together with the symmetry of $\oplus$, allows any permutation of summed terms to be equated by making $\oplus$ associative.

Axiom 5.2 The summation operator is associative (cf. Axiom 2.3).

$$a \oplus (b \oplus c) = (a \oplus b) \oplus c$$

As mentioned in §2.1, non-associative summation has been used to capture intermediate distribution, but such a treatment of these readings is inadequate.

The operator $\oplus$ can also be made idempotent. This is perhaps harder to exemplify in NL. The idea is that:

John and John

just refers to the same individual as:

John
Clearly, we never say “John and John”, when both terms refer to the same individual, as it is redundant. Perhaps we might refer to:

John and the judge

when John is the same person as the judge. Then, if we know that John is the judge, the sentence:

John and the judge are 42.

has the same import as:

John is 42.

**Axiom 5.3** The summation operator is idempotent (cf. Axiom 2.4).

\[ a \oplus a = a \]

In Chapter 7, I will discuss cases where we might not wish to equate the representations of “John” and “the judge”. This will not undermine idempotence.

Definitions in the theory do not strengthen it, they merely provide a clearer expression of other axioms. This definition provides the notion of one term being part of another:

**Definition 5.1** (cf. Definition 2.1)

\[ t \leq s =_{de} t \oplus s = s \]

Again to improve expressibility, this can be internalised as follows:

**Definition 5.2**

\[ t \triangleleft s =_{de} t \oplus s \approx s \]

This definition gives us the following theorems:

\[
\begin{align*}
P(t \triangleleft s) \\
T(t \triangleleft s) & \iff t \leq s
\end{align*}
\]

The predicative constant \( \pi \) is intended to be a property that holds of just all the (provably) denotable terms.

**Axiom 5.4** The term \( \pi \) is a property.

\[ P(\pi t) \]

For simplifying the expression of axioms, a predicate \( \Pi \) is defined in terms of \( \pi \). It will be used, later, in axioms that restrict supremum formation to just plural terms:
CHAPTER 5. PLURAL TERMS IN PT

Definition 5.3

$$\Pi t =_{df} T(\pi t)$$

We might like to think of $\Pi$ taking the role of $E$ (cf. Definition 2.2) in Link’s theory.

If we take two count nominals and conjoin them, the result is a new count nominal. Reflecting this in the representation, we can have the following closure condition on denotable plurals:

**Axiom 5.5** The sum of two plurals is itself a plural.

$$\Pi a \& \Pi b \rightarrow \Pi(a \oplus b)$$

Note that we cannot have $\Pi(a \oplus b) \rightarrow \Pi a \& \Pi b$ as it cannot be guaranteed that the unrestricted parts of a plural are plural.

We want to be able to form the supremum of plurals, that is, fusions of classes of plurals. The term $\sigma$ is intended to do this, yet it is not restricted to forming suprema of plurals, although we are only interested in its behaviour with plurals. The following axiom is indifferent to whether there is a bottom element, and to whether such an element is a plural. It is also indifferent as to what a ‘non-denoting’ definite descriptor denotes. It is important to realise, however, that this axiom is deliberately too weak to justify saying that a non-denoting definite descriptor is a plural. If a property has a mixed extension, this axiom remains silent as to whether the supremum of such a property is a plural.

**Axiom 5.6** If some property has an extension, and it is contained in the domain of plurals, then its supremum is also a plural.

$$\forall p((\text{Pty}(p) \& \forall x(T(xp) \rightarrow \Pi(x)) \& \exists x(T(xp))) \rightarrow \Pi(\sigma xp))$$

This corresponds to some notion of completeness in the domain of plurals. As PT naturally has a Fregean interpretation of a set as a collection of objects that form the extension of some property, then this completeness is only of the definable sets: this axiom thus expresses the notion of definable completeness in Axiom 2.7 with bounded quantification.

Our approach to the so called ‘non-denoting’ definite descriptors is to not have an approach. In particular, the result of attributing a NL property to a non-existent term should not be a statement that is true, or false, it should be a term that cannot be proven to have either truth value. As propositions have truth values, this suggests that the result of such an attribution should not be a proposition. This effect can be achieved by having the denotation of ‘properties’ that arise in the representation of (the plural part of) NL restricted to a class of terms that only form propositions when attributed to terms that are provably plural. I shall adopt a weaker view which defines a class of terms that cannot be proven to produce propositions when attributed to non-plural terms (or rather, terms that are not provably plural):
Chapter 5. Plural Terms in PT

Definition 5.4 A term is a plural property if it forms propositions when applied to a plural term.

\[ \text{P}t \gamma (p) =_{df} \forall x (\Pi x \rightarrow P(px)) \]

This is consistent with a proof-theoretic treatment of helpful answers [De Roeck et al., 1991]: when a proof of the sentence “the present king of France is bald” fails, the weaker theory will result in an unclosed branch corresponding to the fact that “the present king of France” is not a denoting plural. In conjunction with the weak closure condition on plurals, given in the previous axiom, it will also be possible to indicate that it is not possible to prove that “the present king of France” is a denoting plural because there is no present king of France. This corresponds to the feeling that the sentence “the present king of France is bald” sounds odd precisely because there is no present king of France.

Corresponding to singular count nominals, a notion of atomicity in the plural terms can be defined. These are plural terms that have no proper parts.

Definition 5.5 Atoms are the least plurals which are not part of all plurals (cf. Definition 2.3).

\[ It =_{df} \forall x (\Pi x \& x \leq t \rightarrow x = t) \& \sim \forall x (t \leq x) \]

Atomicity can be internalised.

Definition 5.6 (Internal atomicity.)

\[ t =_{df} \lambda t (\Theta x (\exists x \land x \ll t) \land \neg \Theta x (t \ll x)) \]

This definition gives the following theorems:

\[ \text{P}(it) \]
\[ \text{T}(it) \leftrightarrow \text{It} \]

We can define forms of the quantifiers, and supremum operator, restricted to denoting count terms.

Definition 5.7

\[ \forall \Pi x \varphi =_{df} \forall x (\Pi x \rightarrow \varphi) \]

Definition 5.8

\[ \exists \Pi x \varphi =_{df} \exists x (\Pi x \& \varphi) \]

Using these quantifiers, we have a logic where free variables range over all terms, and quantified variables range over denoting count terms, that is, if we restrict ourselves to bounded quantification, then we effectively have a free logic. These quantifiers can be given internal analogues.
Definition 5.9
\[ \Theta_x(t) =_{df} \Theta x(\pi x \Rightarrow t) \]

Definition 5.10
\[ \Xi_x(t) =_{df} \Xi x(\pi x \land t) \]

A restricted supremum operator can also be defined.

Definition 5.11
\[ \sigma_x(px) =_{df} \sigma x(\pi x \land px) \]

The next axioms help characterise I1 (and hence \( \pi \)) and \( \sigma \).

If the nominals “the workers” and “the bosses” are not the same, it is surely because they are constituted by different individuals. This idea is generalised in the following axiom:

Axiom 5.7 Different plural terms have different atoms as parts (cf. Axiom 2.6).
\[ \forall x y (x \leq y \rightarrow \exists u (u \leq x \land u \neq y)) \]

Later, when looking at Landman’s intensional individuals, I will discuss how nominals may have the same extension, yet have different import, so “the workers” and “the bosses” may be ascribed different properties, even if constituted of the same individuals.

The operator \( \sigma \) is intended to capture the meaning of the definite determiner “the”. I take “the men” to refer to the collection of all men (should it exist). Any individual man must be part of this collection.

Axiom 5.8 All plural terms having a property must be part of the supremum of plural terms having that property (cf. Axiom 2.8).
\[ \forall p \forall y ((Pty \land T(px)) \rightarrow y \leq \sigma x px) \]

As any individual man is part of the collection of humans (“the humans”), it follows that the collection consisting of all men (“the men”) is also part of the collection of humans.

Axiom 5.9 If all plural terms which have a property \( p \) are part of some other plural term \( y \), then the supremum of terms having \( p \) must also be part of \( y \) (cf. Axiom 2.9).
\[ \forall p \forall y ((Pty \land \forall x (T(px) \rightarrow x \leq y)) \rightarrow \sigma x px \leq y) \]

The final axiom of the current theory is intended to capture the notion that a collection of individuals is dependent upon the individuals concerned: should we take different individuals, then we would have a different collection. There is no sense in which:
CHAPTER 5. PLURAL TERMS IN PT

John and Mary and Harry

can be the same term as:

John and Mary

This restriction can be achieved by noting that “Harry” is also part of “Mary and Harry”, which in turn is part of the larger term. Thus “Harry” takes part in the generation of the larger term, and is not part of it contingently. Generalising, when an individual is part of a collection of terms having some property, it must either be because it has that property, or is part of a term that does.

**Axiom 5.10** Different atoms have different suprema (cf. Axiom 2.11).

\[ \forall p \forall u ((\text{Pt} \_y(p) \& \text{If} \_u \& u \leq \sigma _x pxr) \rightarrow \exists \_y z (\text{T}(pz) \& u \leq z)) \]

Next, some definitions are given to add a collective operator * and a distributive operator D. To remain compatible with the account of non-plural-denoting terms, no requirement is placed on the truth values, or proposition- hood, of collective or distributive predication outside the domain of plurals.

**Definition 5.12** (cf. Axiom 2.10).

\[ * =_{df} \lambda p \lambda t (t \approx \sigma x (px \land x \ll t)) \]

**Definition 5.13** (cf. Definition 2.4)

\[ D =_{df} \lambda p \lambda x (\Theta u (iu \land u \ll t \Rightarrow pu)) \]

From these definitions we can infer the following:

\[ \begin{align*}
\text{Pt} \_y(p) & \rightarrow \text{Pt} \_y(*p) \& \text{Pt} \_y(Dp) \\
\text{Pt} \_y(p) \& \text{If} & \rightarrow (T(Dp) \iff T(pt)) \\
\text{Pt} \_y(p) \& \text{If} & \rightarrow (T(*p) \iff T(pt)) \\
\text{Pt} \_y(p) & \rightarrow (T(Dp) \iff \forall u (iu \land u \ll t \rightarrow T(pu))) \\
\text{Pt} \_y(p) & \rightarrow (T(*p) \iff t = \sigma x (\lambda x (px \land px \land x \ll t)))
\end{align*} \]

The element \textit{top} — the plural of which all others are a part — can be defined as follows:

**Definition 5.14** The top element \( T \) of the plurals is the supremum of all plural terms.

\[ T =_{df} \sigma _x x \approx x \]

The complement of a plural term — the supremum of plural terms not part of a given plural term — can be defined as:
Definition 5.15 The complement $\bar{s}$ of the supremum $s$ of plurals having property $p$ is the supremum of plurals not having property $p$.

$$\bar{s} = \text{def } \sigma_x (\neg \varphi)$$

Note that as this is definitional, a bottom element $\perp$ — the term that is a part of all plural terms — can be defined as $\top$, but it cannot be proven that $\neg \perp$.

To implement the formal semantics of natural language in a Montague style using PT we can follow the account due to Turner [Turner, 1992]. The predicates $P, T$ can be viewed as characterising types in the language of terms. These types can be enriched with $Det, Quant$ corresponding to those terms which represent determiners and quantifiers respectively, in natural language [Turner, 1992]. These are defined by:

$$Det(f) = \text{def } \forall x (Pty(x) \rightarrow Quant(f(x)))$$

$$Quant(f) = \text{def } \forall x (Pty(x) \rightarrow P(f(x)))$$

and the function space predicates can be constructed using:

$$(P \rightarrow Q)(f) = \text{def } \forall x (P(x) \rightarrow Q(f(x)))$$

Thus those objects which give propositions when predicated of quantifiers are of type $Quant \Rightarrow P$.

To consider intensionality with plural terms, we must consider more carefully the nature of $Pty_{\Pi}$. Presumably we at least want such objects to be able to form propositions with certain intensional quantifiers. We can produce a restricted set of sorts, constrained to plurals.

Definition 5.16

$$Det_{\Pi}(f) = \text{def } \forall x (Pty_{\Pi}(x) \rightarrow Quant_{\Pi}(f(x)))$$

Definition 5.17

$$Quant_{\Pi}(f) = \text{def } \forall x (Pty_{\Pi}(x) \rightarrow P(f(x)))$$

The function space predicates can be defined, as before, with:

Definition 5.18

$$(P \rightarrow Q)_{\Pi}(f) = \text{def } \forall x (P(x) \rightarrow Q(f(x)))$$

A model of PT extended with plurals can be constructed by taking a lattice-theoretic model of lambda terms [Scott, 1970], and giving this a Frege structure, as detailed in [Aczel, 1980]. A sublattice of the model for the lambda terms can be strengthened to a complete Boolean algebra. Being a plural corresponds, in the model, to being in this Boolean algebra. The
summation and supremum operators can be interpreted as the summation and supremum operators of the lattice. More details are given in Appendix A.

We now have incorporated Link's plural logic into PT. This completes the analysis of extensional, non-group denoting plurals in a property theoretic setting.

I have avoided the issue of why some things should count as atoms. So that distribution works correctly, we have tacitly assumed that the representations of singular nominals are atomic, so a human, for example, must be atomic, as should individual windows and window parts. Yet it is true that a hand must be, in some sense, part of a human, and that a window part is part of a window. These senses of being "part of" are not representable in the theory as it stands. One way out of this problem is to explore the mereological (part-whole) behaviour of individuals: a hand is part of a human in the sense that the physical manifestation of the hand is part of the physical manifestation of the human. As noted in §2.3, mereological (part-whole) structures, have been suggested as a means of representing sentences involving mass terms, ([Bunt, 1985, Lønning, 1987b] for example) although it turns out that there are problems with this concerning distributive properties, and the inability to distinguish predication of the suprema of different properties when the suprema are equated.

The next chapter is concerned with axiomatising a domain of potentially atomless natural language denotable terms. This may not be directly suitable for a treatment of mass terms. Essentially, distributive inferences require the representation of mass terms to be homogeneous. For many simple mass terms this might be appropriate, but for essentially composite terms — like muddy water, for example — it is not obvious that the extension should be homogeneous.

Later, I shall reconsider the nature of the representation of mass and plural terms, so that natural language properties typically have extensions in the domain of denotable terms, and where distribution is restricted by the nature of the referring expression rather than by some independent structural restriction of the extension, such as the adoption of primitive atoms. This does not stop denotables being treated as atomic with respect to some property: a human is an atomic part of a collection of humans, in the sense that a human has no proper parts that are human. This enterprise will allow a unified treatment of plurals and mass terms.

Presupposition

I offer a tentative formal treatment of the existence presuppositions that occur with definite descriptors (so that "the present queen of France is bald" can only be proven to be a proposition if there is a present queen of France), which might conceivably be extendible to other presuppositions that occur with universally quantified sentences: "all Mary’s lovers are French"; fac-
CHAPTER 5. PLURAL TERMS IN PT

sives: “Joan regrets getting her PhD in linguistics”; and cleft constructions: “It was Lee who got a perfect score on the semantics quiz” [Chierchia and McConnell-Ginet., 1990]. I shall not attempt to do so, but will examine further some of the formal — as opposed to linguistic — aspects of presupposition with definite descriptors. My treatment, as presented so far, cannot be taken as a complete formalisation of such presuppositions. In particular, it is not possible to formalise “a presupposes b” within the theory as it stands. Generally, we might take the utterance of a to presuppose b when the following holds:

\[ P(a) \rightarrow T(b) \]

Thus, asserting a to be a legitimate proposition, whether true or false, implies that the presupposition b is true. For example, to assert that “the moon is ten billion years old”, whether true or false, implies the truth of the presupposition “there is a moon”.

This is not very helpful for the proof theory, where we want to be able to prove P(a) if we can prove T(b). If we cannot prove T(b), we cannot prove P(a). This indicates why the above axioms concerning the definite descriptor (interpreted as a supremum operator) essentially embody the form:

\[ T(b) \rightarrow P(a) \]

where b is a proposition concerning the existence of an object, and a is a proposition containing a definite reference to that object. This gives:

\[ \Gamma, T(b) \vdash P(a) \]

Then P(a) \rightarrow T(b) becomes a theorem of the representation if a presupposes b. To be able to prove that a presupposes b we could have the axiom:

\[ (P(a) \rightarrow T(b)) \rightarrow (a \text{ presupposes } b) \]

However, because of the nature of material implication, this would make a presuppose b for all b, if a is not (provably) a proposition. Even if a were a proposition, then it would presuppose all truths. The essential conditional nature of presupposition, not captured here, is that if a is not a proposition, then a proof of the truth of a — perhaps amongst other things — would have allowed a proof of its propositionhood, and if a is a proposition, the absence of a proof of the truth of b would have prevented this being proven. Material implication does not have sufficient intensionality, or content-sensitivity, to describe the desired behaviour as it is not an assertion about, or a relation between, its two arguments [Halmos, 1963].

Outside the theory, we might have a sentence S (at least one not constructed from sentential clauses) to presuppose p (which I shall write as \( S \succ p \)) under the following circumstances:

\[ S \succ p \text{ whenever } \Gamma \vdash P([S]) \text{ implies } \Gamma \vdash T(p) \]
Presumably, only propositions can be presupposed.

In their consideration of presupposition, Chierchia and McConnell-Ginet formalise the notion of expressions filtering the presuppositions of constituent clauses [Chierchia and McConnell-Ginet, 1990, Chapter 6]. The following discussion uses their examples. They suggest that when two sentences are conjoined, the conjunction has the presuppositions of the two conjuncts, except when the first conjunct — perhaps together with the current common ground: those propositions which are accepted for the current conversation — implies a presupposition of the second, in which case that presupposition is filtered, or removed, from the presuppositions of the conjunction. They argue that with the sentence:

Keith has three children and all of Keith’s children are asleep.

the second clause has the presupposition:

Keith has some children.

but as this is implied by the first clause, this presupposition is not inherited by the sentence. This is contrasted with:

All of Keith’s children are asleep and Keith has three children.

where they suggest that the presupposition of the first clause is inherited by the complete sentence.

Implication between sentences behaves in a similar fashion:

If John were here, what Linda lost could be recovered.
If Linda lost something, what she lost would not be valuable.

In both cases, the consequent has the presupposition:

Linda lost something.

Yet the second example does not inherit this presupposition, as it is implied by the antecedent.

Sentential disjunction is different. The order of the disjuncts is not important. If the truth of one of the disjuncts is incompatible with a presupposition of the other disjunct, that presupposition is not inherited by the disjunction. Chierchia and McConnell-Ginet exemplify this with:

\footnote{They give the common ground of a context \( c \) as \( \text{comgrd}(c) \): a set of propositions. They have propositions as sets of possible worlds. This allows \( c^* = \text{def} \bigcap \text{comgrd}(c) \) to be the set of possible worlds where the common ground holds. This is a possible worlds representation of a partial state of information [Groenendijk and Stokhof, 1984]. In property theoretic treatment of propositions, a partial state of information that satisfies the common ground could be just the theoretical closure \( Th \) of \( \text{comgrd}(c) \), that is \( c^* = \text{def} Th(\text{comgrd}(c)) \).}
Either no one has solved the projection problem, or it was Linda who did it.

The second disjunct has the presupposition:

Someone solved the projection problem.

But this is not compatible with the first disjunct, and is not inherited by the disjunction.

We can summarise Chierchia and McConnell-Ginet’s suggested behaviour for iterated constructions of sentences — using conjunction, implication and disjunction between sentences — as something like:

(i) If $S$ is an atomic sentence, then $S \vDash p$ iff $P(p)$ and whenever we can prove that the interpretation of $S$ is a proposition, we can prove that $p$ is true.

(ii) If $S \vDash p$, then $S$ and $S' \vDash p$ and if $S$, then $S' \vDash p$.

(iii) If $S' \vDash p$ and $T(p)$ is not a theorem of $S$ together with the common ground then $S$ and $S' \vDash p$ and if $S$, then $S' \vDash p$.

(iv) If $S \vDash p$ and and $\sim T(p)$ is not a theorem of $S'$ with the common ground, then $S$ or $S' \vDash p$ and $S'$ or $S \vDash p$.

(v) If $S \vDash p$, then not $S \vDash p$.

This differs from their characterisation as it considers the filtering effect of a on each presupposition of b (and vice versa), rather than the entire set of the presuppositions of b.

Chierchia and McConnell-Ginet seek to explain this behaviour in terms of the context change potential of sentences. This forms part of a compositional treatment of dynamic effects in NL. They allow sentences to be added to the common ground if their presuppositions are theorems of the common ground. As an example, the context change potential of a sentence $S$ is $\|S\|$, a partial function between partial states of information: if the presuppositions of S are in $c^a$, then $\|S\|(c^a)$ will be the new context. With conjoined sentences, the filtering of the presuppositions of the second conjunct by the first conjunct is ‘explained’ by the rule:

$$\|S \land S'\|(c^a) = \|S'\|(|\|S\|(c^a))$$

I shall not elaborate any further on a formalisation of context in terms of property-theoretic semantics and its use in explaining the filtering of presuppositions, but I will discuss how the examples of intra-sentential filtering can alter the choice of appropriate axioms for $P$ and $T$.

It should be noted that the behaviour of presuppositions with disjunctions of sentences are problematic for the axiomatisation of propositionhood $P$ and truth $T$ used so far. If we take a disjunction of sentences $S$, $S'$ to be represented by:

$$S \lor S'$$
then the above suggests that this does not have a presupposition \( p \) of \( S \) (or \( S' \)) if the truth of \( S' \) (\( S \) respectively) is incompatible with it. If \( S \) is true, then we would want \( S \lor S' \) to be true. The relevant axiom of truth:

\[
P(S) \land P(S') \rightarrow (T(S) \lor T(S') \leftrightarrow T(S \lor S'))
\]

does not always allow this inference: assuming that if \( S \) is true, then a presupposition \( p \) of \( S' \) is false, so (according to my construal of presuppositions for atomic sentences) \( S' \) cannot be a proposition, and the axiom of truth does not apply. This example illustrates that a stronger, lazy disjunction is called for, so that:

\[
P(S \lor S') \rightarrow (T(S) \lor T(S') \leftrightarrow T(S \lor S'))
\]

where we want \( P(S \lor S') \) to hold provided one of the disjuncts is a proposition. The axiom for propositionhood of disjoined terms is now too weak: we want to be able to assert the truth of a term that cannot be proven to be a proposition using the original axiom of propositionhood for internal disjunction:

\[
P(S) \land P(S') \rightarrow P(S \lor S')
\]

Care must be taken in strengthening this axiom: if we take it to be:

\[
P(S) \lor P(S') \rightarrow P(S \lor S')
\]

then the logic becomes inconsistent. Consider the case where \( S \) is a false proposition, and \( S' \) is a paradoxical statement, such as Russell’s Liar. According to the proposed rule, \( S \lor S' \) will also be a proposition. However, as \( S \) is false, the truth value of \( S \lor S' \) will depend upon the truth value of a paradox.\(^7\) If one of the disjuncts is false, then the other disjunct must be a proposition for the disjunction consistently to be a proposition. This is expressed in the following axioms for the propositionhood of disjunctions:

\[
P(a) \land (\neg T(a) \rightarrow P(b)) \rightarrow P(a \lor b)
\]

\[
P(b) \land (\neg T(b) \rightarrow P(a)) \rightarrow P(a \lor b)
\]

If we take implication between sentences to be represented by material implication, then a similar argument applies. The sentence:

If there is a present king of France, then he will be bald.

should not inherit the presupposition that there is a present king of France. Representing this as:

\[
S \Rightarrow S'
\]

\(^7\)With a model of property theory based upon the notion of stability [Gupta, 1982, Herzberger, 1982], the truth value of \( S \lor S' \) would be unstable, and so it cannot be a proposition.
then, reinterpreting Chierchia and McConnell-Ginet’s constraints on presuppositions, we do not want the propositionhood of this sentence to depend upon the existence of the present king of France. The relevant axiom:

\[ P(S) \& P(S') \rightarrow P(S \Rightarrow S') \]

requires that \( S' \) (“the present king of France is bald”) be a proposition. This cannot be proven, as there is no present king of France, and so \( S \Rightarrow S' \) cannot be proven to be a proposition. The current axioms for propositionhood, together with my construal of presupposition for atomic sentences, forces the presuppositions of \( S' \) to be inherited by \( S \Rightarrow S' \) always, in opposition to the desired filtering of presuppositions by \( S \). This can be rectified by replacing the above axiom with:

\[ P(S) \& (T(S) \rightarrow P(S')) \rightarrow P(S \rightarrow S') \]

Thus the conditional is a proposition even if the consequent is only a proposition when the antecedent is true. If \( S \) is false, then we would like \( S \Rightarrow S' \) to be true, even if \( S' \) is not provably a proposition (because its presuppositions are false, for example). The relevant axiom of truth should then be strengthened from:

\[ P(S) \& P(S') \rightarrow (T(S \Rightarrow S') \leftrightarrow (T(S) \rightarrow T(S')) \]

to:

\[ P(S) \& (T(S) \rightarrow P(S')) \rightarrow (T(S \Rightarrow S') \leftrightarrow (T(S) \rightarrow T(S')) \]

These stronger axioms for implication are those given in [Turner, 1990, Chapter 5]. As indicated in that work, they allow Martin-Löf’s notions of dependent product and dependent sum, as they occur in his type theory [Martin-Löf, 1982], to be defined within PT. Martin-Löf’s type theory has been used to model dynamic effects in NL, as exemplified by anaphoric resolution [Ranta, 1991]. Thus, the considerations of presupposition, given here, provide independent motivation for a strengthening of the axioms of PT, which then enables it to capture a theory that can treat dynamic effects in NL. However, I shall not seek to cover NL dynamic phenomena in my semantics.
Chapter 6

Mereology in PT

This section is concerned with axiomatising a mereological (part-whole) structure over the natural language denotable terms. This theory is intended to model part-whole relations, without commitment to material extensionality: it is not necessarily a physical mereology, counter to the nominalistic bent of earlier presentations of mereology [Leonard and Goodman, 1940]. It is intended to give a general treatment of part-whole relations that seem to be required for treatments of both mass terms and plurals.

Most of the axioms presented below are much like those in the previous chapter, where I motivate the axioms by reference to NL examples. Although the structure axiomatised here perhaps could be used to represent NL mass nominals directly, I shall not do so, and I will not use NL examples as motivation. Rather, I shall take this structure to represent the underlying denotation of nominals in the final theory, given in Chapter 8: elements of the structure will form the extensions of nominal expressions, although the representations of nominals will not be equated with them. It is these elements I refer to as natural language denotable terms.

To represent mass terms directly in this theory, and to allow distributive inferences, would require some means of stopping distribution of properties to inappropriate parts. This suggests a commitment to homogeneous extensions for mass terms, as adopted by Bunt [Bunt, 1985]. Although simple mass terms might seem to behave homogeneously, it is not clear that they should be given homogeneous extensions in general, especially when considering composite mass terms, like muddy water. Alternatively, some notion of atomicity could be used to block distribution. This would still have to conform to an homogeneous principle: with muddy water, each atom would have to be of muddy water, not mud or water. To allow composite mass terms to be constituted of their component substances, the final theory will control distributive inferences using the guise, or role, under which a term is referred to, to restrict the parts to which a property can distribute. Chapter 7 explores the notion of guises/roles.

The structure of the domain of denotables should be the weakest re-
required to capture the appropriate part-whole relationships in the extensions of nominal expressions. Concerning this, the literature on mass terms raises the question of the existence of primitive atoms within a semantic theory (as opposed to atoms used as a merely formal device to block distribution) [Bunt, 1985]. The answer is supposed to be that which would be given by some unsophisticated naive physics account, rather than that provided by current science. That is, although theories in modern physics suggest that the physical manifestations of substances are atomic (they have minimal parts), it need not be the case that the formal semantics of NL has to embody this. In particular, it is clear that people can communicate, and use mass terms, without being aware of the theory of physical atoms. What is at issue, is whether the everyday usage of NL mass terms in itself requires, or presupposes some atomic level in their semantics. These atoms need not correspond to physical atoms. We could be suspicious of whether such an issue has any meaning to an unsophisticated person. Should such atoms exist, we might question their countability, and ponder whether the part-whole ordering is dense. These questions may have no definitive answer. I shall axiomatise the domain so that the issue of atomicity is undecided, just as it seems to be in the literature. I shall take the view that there may be atoms, but their existence will not be provable in the formal system.

As in the theory of plurals of the previous chapter, it is not clear that there is some denotable bottom element that is a part of all denotable terms. As before, it will be possible to define a bottom element, but it will not be possible to demonstrate that it is denotable.

In summary, the following theory axiomatises a mereology as a Boolean algebraic-like structure, which will not be provably atomic, and will not provably contain a bottom element. The theory is intended to capture the structure that appears to underlie the referents of NL nominal expressions.

Language of mereological terms
To the basic vocabulary of terms is added:

\[
\begin{align*}
\text{Predicative constant:} & \quad \delta \\
\text{Operative constants:} & \quad \oplus, \otimes, \sigma
\end{align*}
\]

It is intended that \( \delta \) will hold of natural language denotable terms. As before, \( \sigma \) is the supremum operator, and the term \( \oplus \) is the summation operator. Within the domain of denotables, \( \otimes \) will be its dual.

The language of wff remains unchanged. The theory is governed by the axioms of the \( \lambda \beta \)-calculus as before, together with the previous axioms for propositions P and truth T.

**Further Axioms For a Mereology in PT**
The operator \( \oplus \) in this theory is used to indicate a fusion of terms. The order in which the terms are considered is irrelevant, and so the operator is both symmetric and associative.
Axiom 6.1 Symmetry (cf. Axiom 5.1).

\[ a \oplus b = b \oplus a \]

Axiom 6.2 Associativity (cf. Axiom 5.2).

\[ a \oplus (b \oplus c) = (a \oplus b) \oplus c \]

Combining a term with itself will result in nothing new.

Axiom 6.3 Idempotence (cf. Axiom 5.3).

\[ a \oplus a = a \]

As in the previous theory, an ordering can be defined in terms of \( \oplus \), which simplifies the expression of some of the axioms.

Definition 6.1 Definable ordering (cf. Definition 5.1).

\[ t \leq s \overset{\text{def}}{=} t \oplus s = s \]

An internal form of the ordering can be defined.

Definition 6.2 Internal ordering (cf. Definition 5.2).

\[ t \ll s \overset{\text{def}}{=} t \oplus s \approx s \]

The terms that are of interest are those which correspond to the extensions of referring descriptions. The predicative constant \( \delta \) is a property which holds of any such NI denotable term.

Axiom 6.4 Internal notion of denotable (cf. Definition 5.4).

\[ P(\delta t) \]

As syntactic sugar, the predicate \( \Delta \) can be defined in the language of wff which holds of NI denotables.

Definition 6.3 A term is denotable, as an external wff, iff it is true that it is denotable as an internal proposition (cf. Definition 5.3).

\[ \Delta t \overset{\text{def}}{=} T(\delta t) \]

The fusion of two denotables will itself be denotable.

Axiom 6.5 The sum of denotables is a denotable (cf. Axiom 5.5).

\[ \Delta a \& \Delta b \rightarrow \Delta(a \oplus b) \]
As with the previous theory which adds plurals to PT, we want to be able to form the supremum of classes of denotables. Again, the term $\sigma$ is intended to do this.

The previous closure condition on denotables can be generalised so that the fusion (supremum) of a class of denotables is itself denotable. The axiom requires that the class be non-empty, in order to prevent a proof of the denotability of a bottom element (an element that is part of all denotables).

**Axiom 6.6** If some property has an extension, and it is contained in the domain of denotables, then its supremum is also in the domain of denotables (cf. Axiom 5.6).

$$\forall p((Pty(p) \land \forall x(T(px) \rightarrow \Delta(x)) \land \exists x(Tpx)) \rightarrow \Delta(\sigma px))$$

Typically, the ‘properties’ of interest in the representation of NL need only form propositions with denotable terms. This restricted class of properties is useful in weakening the theory so that ‘non-denoting’ definite descriptors need not have NL denotable extensions. By restricting NL derived properties to this class, it is not possible to prove that applying such a ‘property’ to a non-denotable (or rather, not provably denotable) term results in a proposition.

**Definition 6.4** A property of denotables produces a proposition, given some denotable (cf. Definition 5.4).

$$Pty_\Delta(p) =_{df} \forall x(\Delta(x) \rightarrow P(x))$$

We can define forms of the quantifiers, and supremum operator, restricted to denotable terms.

**Definition 6.5** (cf. Definition 5.7).

$$\forall_\Delta \varphi =_{df} \forall x(\Delta x \rightarrow \varphi)$$

**Definition 6.6** (cf. Definition 5.8).

$$\exists_\Delta \varphi =_{df} \exists x(\Delta x \& \varphi)$$

Restricting ourselves to these quantifiers, we again achieve the effect of a free logic, where free variables range over all terms, and quantified variables range over denoting terms. As before, these quantifiers can be given internal analogues, and a supremum operator which is restricted to denotables can be defined.

**Definition 6.7** (cf. Definition 5.9).

$$\Theta_{\delta} x(t) =_{df} \Theta x(\delta x \Rightarrow t)$$
Definition 6.8 (cf. Definition 5.10).
\[ \Xi_x(t) =_{df} \Xi x(\delta x \land t) \]

Definition 6.9 (cf. Definition 5.11).
\[ \sigma_x(px) =_{df} \sigma x(\delta x \land px) \]

If two denotables are not equal, then there is a denotable that is part of one, but not part of the other. The antecedent can be strengthened, as two denotables may be unequal, yet one be wholly contained in the other.

Axiom 6.7 Different denotables have different parts (cf. Axiom 5.7).
\[ \forall \Delta xy(x \not\leq y \rightarrow \exists \Delta u(u \leq x \land u \not\leq y)) \]

The term \( \sigma_x px \) is intended to form the fusion of all denotables having the property \( p \). Any denotable with the property \( p \) should be part of this term.

Axiom 6.8 All stuff having a particular property must be a part of the supremum of stuff having that property (cf. Axiom 5.8).
\[ \forall p \forall \Delta y( (\text{Pt}_\Delta(p) \& T(y)) \rightarrow y \leq \sigma_x px) \]

I shall illustrate the need for the following axiom by way of example. If we consider the denotable substance “cold mud”, any bit of cold mud will be part of the fusion of mud. If we take the fusion of bits of cold mud, it will also be part of the fusion of mud.

Axiom 6.9 If all denotables having property \( p \) are part of some other denotable \( y \), then the supremum of denotables having \( p \) must also be a part of \( y \) (cf. Axiom 5.9).
\[ \forall p \forall \Delta y( (\text{Pt}_\Delta(p) \& \forall \Delta x(T(px) \rightarrow x \leq y)) \rightarrow \sigma_x px \leq y) \]

The intuition behind the next axiom is akin to that of Axiom 5.10 where different collections of minimal parts have different suprema. There is a complication, as there is no reference to atoms in this theory: we cannot consider a unique way of dividing a denotable into its minimal parts; different collections of denotables may have the same supremum if they cover the same denotables; further, we cannot guarantee to pick denotables that are part of a denotable with the required property. However, in this final case, the denotable will have a part with the required property.
**Axiom 6.10** Different portions of denotables have different supremaums: something that is part of the supremum of a property must be part of something with that property, or have a part with that property (cf. Axiom 5.10).  

\[ \forall p \forall u ((Pty(p) \land u \leq \sigma_\epsilon x p r x) \rightarrow (\exists z (p z \land u \leq z) \lor \exists z (p z \land z \leq u))) \]

Obviously, this is not much use if we have uncountable parts, but that would always cause problems.

As before, it is possible to define an operator * which turns a property into a cumulative property [Lönning, 1989, Link, 1991a].

**Definition 6.10** (cf. Definition 5.12)  

\[ * =_{def} \lambda p \lambda t (t \approx \sigma x (p x \land x \ll t)) \]

There is no obvious role for a distributive operator [Lönning, 1989, Link, 1991a] in this theory, as it stands, as there are no atoms to distribute to. Later, in Chapter 8, I will indicate how distributive behaviour can be regained.

As in the previous chapter, the denotable of which all other denotables are a part (top) can be given as:

**Definition 6.11** The top element \( \top \) of the denotables is the supremum of all denotable terms (cf. Definition 5.14).

\[ \top =_{def} \sigma_\epsilon x (x \approx x) \]

Complement can be defined as:

**Definition 6.12** The complement \( \bar{\varphi} \) of the supremum \( \varphi \) of denotables having property \( p \) is the supremum of denotables not having property \( p \) (cf. Definition 5.15).

\[ \sigma_\epsilon x \bar{\varphi} =_{def} \sigma_\epsilon x (\neg \varphi) \]

Again, a bottom element \( \bot \) can be defined as \( \overline{\top} \), but it can not be proven that \( \Delta \bot \).

Amongst the denotables, \( \otimes \) can be the dual of \( \oplus \). This will be used in the final theory to model disjunction.

**Axiom 6.11** Within the domain of denotables, \( \otimes \) is the dual of \( \oplus \).

\[ \Delta a \land \Delta b \rightarrow \overline{x \oplus y} = \overline{x} \otimes \overline{y} \]

I give the duality in terms of an axiom, although it is possible to make both \( \otimes \) and \( \oplus \) definitional — noting that any denotable \( a \) can be given as the supremum \( \sigma_\epsilon x (x \approx a) \) — using the following equalities:

\[ \sigma_\epsilon x p r x \oplus \sigma_\epsilon x q x x = \sigma_\epsilon x (p r \lor q x) \]
\[ \sigma_\epsilon x p r x \otimes \sigma_\epsilon x q x x = \sigma_\epsilon x (p r \land q x) \]
CHAPTER 6. MEREEOLOGY IN PT

These equalities are a consequence of the axioms given. If they are taken as
definitional, then the axioms governing $\oplus$ (and $\otimes$) are satisfied.

To implement the formal semantics of natural language in a Montague
style using PT we can follow the account in the previous chapter and define
a restricted set of types, constrained to denotables.

**Definition 6.13** (cf. Definition 5.16)

$$\text{Det}_\Delta(f) =_{df} \forall x(\text{Pty}_\Delta(x) \rightarrow \text{Quant}_\Delta(fx))$$

**Definition 6.14** (cf. Definition 5.17)

$$\text{Quant}_\Delta(f) =_{df} \forall x(\text{Pty}_\Delta(x) \rightarrow P(fx))$$

The function space predicates can be defined as before.

**Definition 6.15** (cf. Definition 5.18)

$$(P \rightarrow Q)_\Delta(f) =_{df} \forall x(P(x) \rightarrow Q(fx))$$

To have a conventional lattice-theoretic treatment of plurals would just re-
quire the addition that the denotable terms are founded on atoms. However,
it is possible to treat count terms without recourse to atoms. For example,
we could say that a singular property — corresponding to a singular count
term — can hold of a term only if it does not hold of any proper part of that
term. A plural property — corresponding to a plural count term — would
then be the collective form of a singular property.

These axioms are no stronger than those given in the theory of plurals in
PT. As this is the case, a model of PT extended with plural terms will also
be a model of this theory of potentially atomless terms. Although the model
may possess atoms, this has no effect on a proof of the consistency of this
logic.

In the next chapter, I shall explore insights from Landman’s work on
representing individuals under roles, and show how they can be reinterpreted
using property modifiers in PT. In Chapter 8, I show how these roles can be
used to obtain the correct distributive behaviour for mass terms without
assuming homogeneous extensions.
Chapter 7

Intensional Subjects

From now on, I shall assume that we do not want to rule out the possibility that objects may be equated with a fusion of substance(s), or that ‘different’ objects, and substances — objects referred to by the suprema of different properties — may have the same extension, in particular, that compound substances like “muddy water” may have the same extension as their constituent substances, “mud and water”.

One means of testing whether the semantics can effectively distinguish the extensions of nominals, when those extensions are equated, is to examine the inferences which can be made. In particular, distributive inferences should differ according to the guise under which the extension is considered: different nice parts should be recoverable. Landman has considered the representation of individuals under roles, within a theory of plurals [Landman, 1989]. I shall present his theory, and show that his formalisation is either inconsistent or trivial. This does not mean that his intuitions should be discarded. I then give a naive translation of his theory into PT, and show that it suffers the same difficulty. I reformulate the theory in a consistent manner which seems to satisfy Landman’s intuitions. As a consequence of the weak typing of PT, it transpires that many of Landman’s axioms become theorems of this reappraisal. Following this, I show that Landman’s example sentences, used to motivate the need for representing individuals under roles (where the roles modify individuals), can be treated by using property modifiers (where roles effectively modify verb phrases). This is, in some senses, a simpler theory, as property modifiers seem to be required for an adequate treatment of adjectives and adverbs. The weak typing of PT means that the same term can act as both a property and a property modifier, without recourse to type-shifting mechanisms.

In the final theory, given in Chapter 8, I shall use property modifiers to recover the nice parts of the extensions of mass terms, and so control distributive inferences, without having to assume that the extensions are atomic, or homogeneous.

Typical of the phenomena Landman tries to address is the apparent failure
of substitution in subject position nominals. In the literature on plurals, this phenomena is exemplified by:

The members of committee A are the members of committee B.
Committee A had a meeting.
Committee B did not have a meeting.

Possible worlds treatments of this are inadequate as the two committees may necessarily have the same members, so there is no contextual means of distinguishing between them. A more intensional approach is to take the committees as basic, primitive objects, and have some function that can be used to recover their members [Barker, 1992].

Initially, Landman takes nouns to be ambiguous between their sum and group reading, citing as motivation Link’s treatment of groups [Link, 1984] which has been used to represent the intermediate distributive readings as discussed in detail by Lønning and Schwarzschild [Lønning, 1989, Schwarzschild, 1990, Schwarzschild, 1992]. Landman takes the ambiguity to be in the definite descriptor, rather than the conjunction, so “the boys” would be ambiguous between a sum $\sigma x \text{boy}'x$, and group reading $\langle \sigma x \text{boy}'x \rangle$ or, in his notation, $\uparrow \sigma x \text{boy}'x$. As Lønning notes, this overgenerates. Landman models these groups with singleton sets. However, in the second part of the paper, he models all (collectively read) sums as denoting singleton sets. To account for sentences with conjoined distributive and collective verbs, such as:

The boys met at school and were wearing their gold earings.

he effectively has to modify the operator $\mathbf{D}$ — with the aid of type-shifting — to allow distribution into groups. According to Lønning there is effectively no difference between Landman’s group representation of collectives and the more conventional sum representation. Lønning suggests that all that can be recovered from this part of Landman’s account is the notion of an ambiguous conjunction [Lønning, 1989].

Initially, Landman uses groups for the representation of committee-like objects [Landman, 1989]. Presumably, this is principally to model the effect that such objects block distribution. On its own, this does not account for the intensionality of committee-like objects, which is part of the objective of Landman’s theory of intensional individuals, to be discussed here. Even if we adopt Barker’s theory for the intensional behaviour of collective nouns as accounts of apparent subject position intensionality, such a treatment does not attempt to unify the apparent intensional behaviour of collective nouns with the singular examples of subject position opacity evident in the following sentences:
CHAPTER 7. INTENSIONAL SUBJECTS

The judge is strict.
John is liberal.
John is the judge.

It seems possible to assert all these sentences together, without there necessarily being a contradiction. I take this as evidence that what is really at issue here is some kind of property modification, so the above sentences can be summarised by "John is liberal, whilst being a strict judge". Landman takes this as evidence for intensional individuals [Landman, 1989] — individuals restricted under different roles, or guises — so that "John, as himself, is liberal, and, as a judge, is strict". If Landman's view is adopted, then Barker's theory could be modified to have extra terms representing individuals under their different roles. Landman's theory allows these extra terms to be constructed from the the bare individuals and the properties that they have, rather than taking them to be primitive terms.

It is Landman's intention that these intensional representations do not add structure to the sums: he does not use individuals under guises to block distribution. He maintains, instead, the group forming operator † for this purpose. This takes a sum of individuals and produces a term into which properties may not distribute.

Landman says that there is a variable degree as to how much connotation is added. The notion of "chairman" seems more intensional than "blue-eyed boy": failure of substitution is more likely to occur with "chairman" as more connotation is added. Landman claims that this is a function of the intensionality of the corresponding properties.

Using another example from [Landman, 1989]:

The judge earns exactly £20000.
The cleaner earns exactly £2000.
John is both judge and cleaner.
John earns exactly £22000.

If we follow Landman's suggestion, then the individuals referred to in these sentences are really individuals under certain roles. An individual may hold different properties under different roles — the properties an individual has under one role may contradict those it has under another role. This may be made explicit in English by such sentences as:

John, as a judge, earns £20000.

Landman dismisses the possibility that the predicate is not "earning", but "earning-as-a-judge" because of comparative sentences:

The judge and the cleaner have different incomes.

As we shall see, the argument that comparatives cannot be represented with property (or predicate) modifiers relies on a particular assumption about the representation of comparatives. I shall show that it is possible to have an acceptable representation of this comparative using property modifiers.
Landman’s Formal Theory of Roles

In Landman’s formal theory, if \( t \) is a term and \( P \) is a predicate, then \( t \downarrow P \) is an expression denoting \( t \) restricted to its aspect of having \( P \). The expression \( t \downarrow P \) is not a new term. Landman cannot deal with it this way: the logic he adopts is based upon Thomason’s model of propositional attitudes [Thomason, 1980], a typed intensional logic with a poor language of terms.

Properties are of type \( \langle \epsilon, p \rangle \). The denotation of NPs are second order properties of type \( \langle \langle \epsilon, p \rangle, p \rangle \). John in all his aspects, \( \lambda P.Pj \), is such a second order property. John, himself, is of type \( \epsilon \). John, as a \( P \), is \( j \downarrow P \), where \( j \) is of type \( \epsilon \), \( P \) of type \( \langle \epsilon, p \rangle \) and \( j \downarrow P \) of type \( \langle \langle \epsilon, p \rangle, p \rangle \).

The proposition “John, as a judge, is on strike” would be represented as:

\[
(j \downarrow J)(S)
\]

Landman has the following axioms to control the behaviour of these new objects:

Landman’s Axioms

(i) John as a judge, is John: \( (j \downarrow J) \lambda x(x = j) \).
(ii) John as a judge is a judge: \( (j \downarrow J)(J) \).
(iii) John as John is John: \( (j \downarrow \lambda x(x = j)) = \lambda P(Pj) \).
(iv) \( (j \downarrow J)(P) \& (j \downarrow J)(Q) \rightarrow (j \downarrow J)(P \& Q) \).
(v) \( (j \downarrow J)(P) \& P \vdash Q \rightarrow (j \downarrow J)(Q) \).
(vi) \( \exists P((j \downarrow J)(\lambda x(Px \& \sim Px))) \).
(vii) \( \forall P((j \downarrow J)(\lambda x(Px v \sim Px))) \).
(viii) \( (j \downarrow J)(P) \rightarrow J(j) \).

Axioms (iv), (v) make restricted terms filters of properties. Axiom (vi) additionally makes them proper filters, and Axiom (vii), ultra-filters. In his initial system, only individuals can be restricted.

In Landman’s theory, \( \downarrow (j \downarrow J) \) (the group formed by \( j \downarrow J \) under the role \( J \)) is a legitimate restricted term, but, without further additions to the theory, \( (j \downarrow b) \downarrow J \) is not. As his theory of plurals is strongly typed, he cannot just add these restricted terms as terms which can be summed. He therefore has to define a new plural domain for these restricted terms, and add additional constraints like: \( (j \downarrow b) \downarrow P = (j \downarrow P) \downarrow (b \downarrow P) \). For his ordinary plurals he uses the power set model of lattice theory for his plurals. This has the unfortunate consequence that the atoms must be singleton sets, so he needs lifted forms of all his properties such that \( P^*(\{a\}) \leftrightarrow P(a) \).

It should be noted that Landman’s axioms, when taken at face-value, give rise to a trivial logic at best. Assuming that given an individual \( i \) and a predicate \( p \), the restricted term \( i \downarrow p \) will exist, then a problem arises when considering Axiom (ii) with Axiom (viii).
CHAPTER 7. INTENSIONAL SUBJECTS

Theorem 7.1 Landman’s Axioms allow all individuals to have all predicates hold of them.

Proof: Consider an individual $i$ and property $p$. From Axiom (ii): $(i \geq p)p$. This satisfies the antecedent of Axiom (viii), so with *modus ponens*: $p(i)$. □

Further, if there is any predicate that has a complement — that is, anything other than a universal predicate — then the logic is inconsistent. It must be assumed that an individual can be restricted by some predicate only if that predicate holds of that individual. Although not captured by his axioms, this seems to be his intention, as is apparent when he makes some comments on the relatedness of the guise under which an individual is viewed, to properties predicated of this restricted term.\footnote{For example, being well paid, and a judge, both relate to jobs. Landman contends that this gives rise to greater intensionality in the subject of propositions like “John, as a judge, is well paid”. He believes that some aspects, or guises, are naturally more intentional than others, judge is more intensional than drunk, which is in turn more intensional than man with a big nose.}

Of particular relevance to the above theorem, he says that some objects have fewer aspects than others, *wads* having fewer than *humans*, for example. This would suggest that Landman only wants to consider the expression $(i \geq p)$ if $p(i)$. It is not clear that this restriction can be expressed within his formal system.

Restricted Terms in PT

Landman suggests that a more elegant expression of plurals with restricted terms may follow if a type free intensional theory is used, such as Turner’s axiomatic PT. Later, I will promote the use of property modifiers to account for Landman’s examples. Here, I shall show a more direct translation of Landman’s ideas into PT. The basic vocabulary of the language of terms is extended with:

$$\oplus, \sigma, \text{Ed}, \hat{\cdot}, \vec{\cdot}$$

The theory is governed by the axioms of the $\lambda\beta$-Calculus as before. The language of wff is as in Chapter 4, with the further axioms for $\oplus, \sigma$ from Chapter 5.

Initially, I will give three suppositions that naively seem to capture Landman’s theory. These suppositions taken together are, however, inconsistent, as is shown using an argument that exactly parallels that used to show the original theory is inconsistent, or trivial. I then offer a definition of what it is to be a restricted term (or, an individual under some role), and give axioms that, I believe, capture Landman’s intuitions. I demonstrate how some of Landman’s axioms, required because of the strong typing of his base logic, are theorems under this proposal.

First of all, Landman has “John as a judge” being equal to “John”. In PT this is of little use: to say that “John as a judge” is “John” amounts to
loosing the distinction between terms and restricted terms. It is better in this formalism to say that the underlying individual of “John as a judge” is “John”.

**Supposition 7.1** Ext obtains the underlying individual of a restricted term, (cf. Landman’s Axiom (i)).

\[ \text{Ext}(j; J) = j \]

Landman then has an axiom that says “John as a judge is a judge”. In PT we first must preface this with the restriction that “being a judge” is a property.

**Supposition 7.2** (cf. Landman’s Axiom (ii)).

\[ \text{Pty}(J) \rightarrow T(J(j; J)) \]

We must be careful here, however. Under what circumstances can we consider an individual under a particular role? Let us consider a translation of Landman’s Axiom (viii) in conjunction with this:

**Supposition 7.3** If a property holds of John as a judge, then John is a judge (cf. Landman’s Axiom (viii)).

\[ \exists P(T(P(j; J))) \rightarrow T(Jj) \]

Now, if any individual can be considered under the role of having any property, then Supposition 7.2 says an individual under the aspect of a property will have that property. As a property now holds of the restricted individual, Supposition 7.3’s antecedent is true, so the property will hold of the unrestricted individual. From this we can conclude that all properties hold of all individuals (cf. Theorem 7.1). Formally:

**Theorem 7.2** Supposition 7.2 with Supposition 7.3 is inconsistent.

**Proof:** By showing that all individuals have all properties. Assume that Pty(j). From Supposition 7.2 and *modus ponens* we can show: T(J(j; J)). With Supposition 7.3 and *modus ponens* this gives: T(Jj). Using implication introduction, discharging the assumption: Pty(J) → T(Jj). Two applications of universal introduction obtains the result that all individuals have all properties: ∀j∀P(J(j; J)) → T(Jj)). The term λx(x ≠ x) is a property. From the last result, all individuals will have this property: ∀j(T(λx(x ≠ x)j)). From the axioms of T: ∀j(j ≠ j). □

It is clear from this that we must limit those objects considered to be restricted terms. Presumably a term may only be considered under the restriction of having a property, if it has that property. It is not desirable to restrict the formation of terms themselves, as this completely changes the
CHAPTER 7. INTENSIONAL SUBJECTS

character of the theory. Those terms considered as restricted terms must instead be limited, just as those terms considered as propositions and properties are limited, without restricting the language of terms. We can have a new predicate $\mathcal{R}T$ which holds of restricted terms. It must satisfy the following:

$$\mathcal{R}T(j ; J) \leftrightarrow Pty(J) \& T(J_j)$$

Via the definition of $Pty$, this is provably equivalent to:

$$\mathcal{R}T(j ; J) \leftrightarrow T(J_j)$$

This is satisfied if we define $\mathcal{R}T$ with:

**Definition 7.1**

$$\mathcal{R}T(x) =_{df} \exists p(Pty(p) \& (i ; p) = x \& T(p))$$

The intuition behind Landman’s Axiom (viii) is embodied in this definition: if we can consider the restricted term of “John under the aspect of being a judge”, then John is a judge (ignoring implicit counter-factuals).

An internal notion of $\mathcal{R}T$, $\hat{\mathcal{R}}$, can be added to the language of terms, with the following truth conditions:

**Axiom 7.1** ‘Internalised’ $\mathcal{R}T$.

$$\forall p(Pty(p) \rightarrow T(\hat{\mathcal{R}}(t ; p)) \leftrightarrow \mathcal{R}T(t ; p))$$

Supposition 7.2 still seems too strong: in combination with the definition of $\mathcal{R}T$ we can prove that any individual can be considered as a restricted individual under some property, itself restricted under the same property:

**Theorem 7.3** Supposition 7.2 allows us to infer $\mathcal{R}T((j ; J) ; J)$ for any individual $j$, with any property $J$.

**Proof**: Assume $Pty(J)$. From Supposition 7.2 and *modus ponens*: $T(J(j ; J))$. From this, the assumption, and the definition of $\mathcal{R}T$, we derive: $\mathcal{R}T((j ; J) ; J)$. Using implication introduction and discharging the assumption:

$$Pty(J) \rightarrow \mathcal{R}T((j ; J) ; J).$$

After two applications of universal introduction:

$$\forall J \forall j(Pty(J) \rightarrow \mathcal{R}T((j ; J) ; J)).$$

It is not clear that this matches any intuition. A consistent characterisation of Landman’s Axiom (ii) in PT which appears to capture the appropriate intuitions is:
Axiom 7.2

\[ \text{RT}(j; J) \rightarrow T(J(j; J)) \]

It is an open choice whether Supposition 7.1 should also be guarded to give the weaker axiom:

Axiom 7.3

\[ \text{RT}(j; J) \rightarrow \text{Ext}(j; J) = j \]

Landman has “John as John” being the individual sublimation of John. I think this is because he has restricted individuals as determiners (the same type as the individual sublimation), not terms. It seems more intuitive to have “John as John” being “John” himself, rather than the individual sublimation of John, especially when restricted terms are taken as terms.

Axiom 7.4 (cf. Landman’s Axiom (iii))

\[ j; (\lambda x(x \approx j)) = j \]

We naturally have roles as proper filters of properties as a consequence, that is:

(i) The property whose extension is the intersection of the extension of two properties holds of a term when the two properties hold of the term.

(ii) If the extension of a property is contained within that of another, then when the first property holds of a restricted term, so does the second.

(iii) There is no property that both holds, and does not hold of a restricted term.

Because of the strong typing of Landman’s base theory, he has to add these requirements as additional axioms. The weak typing of PT allows them to be theorems, as is demonstrated below.

Theorem 7.4 The property whose extension is the intersection of the extension of two properties holds of a term when the two properties hold of the term (cf. Landman’s Axiom (iv)).

\[ T(P(j; J)) & T(Q(j; J)) \rightarrow T(\lambda x(Px \land Qx)(j; J)) \]

Proof: From axioms of T. \[ \square \]

Theorem 7.5 If the extension of a property is contained within that of another, then when the first property holds of a restricted term, so does the second (cf. Landman’s Axiom (v)).

\[ T(P(j; J)) & \forall x(T(Px) \rightarrow T(Qx)) \rightarrow T(Q(j; J)) \]
CHAPTER 7. INTENSIONAL SUBJECTS

Proof: From application of modus ponens. □

Theorem 7.6 There is no property that both holds, and does not hold of a restricted term (cf. Landman’s Axiom (vi)).

\[ \sim \exists P(T(\lambda x(Px \land \neg Px)(j \circ J))) \]

Proof: From \( \beta \)-reduction, the axioms of T, and first order logic. □

Do we have roles as ultra-filters? This follows from the law of excluded middle:

Theorem 7.7 (cf. Landman’s Axiom (vii))

\[ \forall P(Pty(P) \rightarrow T(P(j \circ J) \lor \neg P(j \circ J))) \]

Proof: Using the law of excluded middle. Assume: Pty(P). From the definition of Pty: \( P(P(j \circ J)) \). From the axioms of P: \( P(\neg P(j \circ J)) \). Taking these last two results, and the axioms of P:

\[ P(P(j \circ J) \lor \neg P(j \circ J)). \]

From the axioms of T:

\[ T(P(j \circ J) \lor \neg P(j \circ J)) \leftrightarrow T(P(j \circ J)) \lor \sim T(P(j \circ J)). \]

Using the law of excluded middle, and modus ponens

\[ T(P(j \circ J) \lor \neg P(j \circ J)). \]

With implication introduction, discharging the assumption:

\[ Pty(P) \rightarrow T(P(j \circ J) \lor \neg P(j \circ J)). \]

Finally, using universal introduction:

\[ \forall P(Pty(P) \rightarrow T(P(j \circ J) \lor \neg P(j \circ J))). \]

□

Further Points

All of Landman’s other restricted individual determiners can be expressed as terms in this extended PT without further complications such as type shifting, or additional part-whole structures: Landman originally gives a power set structure to the unadulterated individuals, and only atomic individuals (or “grouped” collections) can be restricted. In order to represent restricted sums, and sums of restricted individuals:

\[ (j \oplus b) \circ J \]

\[ (j \circ J) \oplus (b \circ J) \]
he has to build an additional plural structure for restricted terms. The PT version of restricted individuals I have given does not hinder the formation of such terms, as it has only weak typing. Much of the effort that Landman expends on his theory seems directed at overcoming the type system of his basic representation language.

When considering the plural structure for restricted terms, Landman effectively stipulates that:

\((j \uplus b)^* P = (j^* P) \uplus (b^* P)\)

which can be generalised to:

\(\sigma x(Dq(x^* P)) = \sigma y(y \approx (x^* p) \land qx)\)

This means that a term, under the role of having a distributive property, cannot be used to block distribution. Landman prefers to keep the intensional notion of roles separate from group formation, which is why he keeps the “group” forming/distribution blocking operator \(\uparrow\) in his theory. Without the extra plural domain, Landman can have determiners of the form:

\(\uparrow (j \uplus b)^* J\)

In the PT theory presented so far, however, there is no intension neutral group forming operator \(\uparrow\). The main function of \(\uparrow\) was to block distribution into committee-like objects, and to represent intermediate distribution. I have no intention of representing intermediate distribution, and feel that Barker’s account of committee-like objects is to be preferred [Barker, 1992]. It would be possible to block distribution with the restricted individual. However, Landman does not do this. His operator \(\uparrow\) blocks distribution, and \(\;\); does not. Even nouns like “committee” he gives a distributive interpretation. For him, “committee” is a property which holds of the individuals in the committee. He represents “the committee” as ‘the sum/group of the people that are members of the committee, as members of the committee’:\(^2\)

\(\sigma x(\text{committee}A'(x); \lambda x(\text{committee}A'(x)))\).

The committee, in this sense, is the fusion of the members of the committee, under a certain role. A fundamental problem with Landman’s account arises when we consider the metaphysics of change: by making the members of a committee part of the term that represents a committee, the committee itself, becomes dependent upon the current members. If the members change, then the committee is different. To ensure that a committee can survive changes in its make-up, it is necessary to give it an existence independent of its members. For example, we might wish to say “the committee was founded

\(^2\)In the manuscript for [Landman, 1989].
in 1920” without wishing to imply that the members (as members) were founded in 1920. A possible way out, for Landman, is to use the apparent connection with more standard intensional notions, where, for example, “the committee” may be taken to be ambiguous, either referring to the members fulfilling that role, or the role of “being the committee” itself, as in the approach explored by [Barker, 1992]. It is possible to give a function that maps a committee onto its current members, but without making the value of that function part of the term representing the committee.

A problem we might discuss, which Landman does not contemplate because of his typed theory, is the limits to be placed on the kinds of objects which can be restricted. Further he does not address how the appropriate aspects are to be found compositionally. Concerning the first problem, one issue is whether restricted terms should themselves be restricted:

\[(j \upharpoonright J) \upharpoonright H\]

This cannot just be made equivalent to:

\[j \upharpoonright \lambda x(\text{J}x \land Hx)\]

because the existence of the former implies \(T(H(j \upharpoonright J))\), whilst the latter would imply \(T(Hj)\).

Taken in combination with a theory of conjoined terms, we may enquire whether sums of restricted terms themselves can be restricted:

\[((j \upharpoonright J) \oplus (b \upharpoonright H)) \upharpoonright S\]

It is partly to avoid the mental gymnastics involved in trying to examine intuitions on these matters that I adopt a property modifier approach, where the appropriate behaviour can be obtained by considering theories of NL modifiers, such as adjectives.

Property Modifiers

It seems more natural to say “John is a strict judge” than “John, as a judge, is strict”. The sentence “the judge is strict” could well be taken to be elliptic for “the judge is a strict judge”. Even when the “as a...” construction is natural, the use of commas, when it is adjacent to the nominal, suggests some movement has occurred:

John, as a judge, earns £2000.
John earns £2000 as a judge.

\(^3\)This appears to parallel the arguments with singular terms, such as “the prime-minister” as an individual, as opposed to “the prime-minister” as some role, or position. Note that [Barker, 1992] does not explore the intensionality of singular, non-collective, terms.
This, perhaps, gives some support to the policy of treating roles as property modification, although syntactic arguments like this are potentially hazardous.

We can attempt a property modifier approach, and see if it can be made to cope with Landman’s comparative examples. The property modifiers will correspond, in some sense, to the *subjective adjectives* [Kamp, 1975] even though they may be derived from the translation of NL verbs. These have also been called *restrictive* or *affirmative* adjectives [Hoepelman, 1983]. The term they modify shall usually be derived from the subject nominal. In this chapter, I shall offer just enough of a formal exposition to indicate how this approach works in principle. The next chapter gives a more detailed account, with definitions and axioms to govern the behaviour of property modifiers.

It is easy to think of “strict judge” as corresponding to some new property:

\[ ((\text{strict}')(\text{judge}') ) \]

where judge’ corresponds to one of Landman’s roles, and the property modifier strict’ corresponds to a predicate in Landman’s theory. That which Landman calls a *role* will be just a property (of stuff) modified by a property modifier. A property modifier \( \mathfrak{P} \mathfrak{M} \) will have the following behaviour:

\[
\begin{align*}
\mathfrak{P} \mathfrak{M}(p) & \rightarrow \forall q( \text{Pty}_\Delta(q) \rightarrow \text{Pty}_\Delta(pq)) \\
\mathfrak{P} \mathfrak{M}(p) & \rightarrow \forall q(\sim \text{Pty}(q) \rightarrow P(pq))
\end{align*}
\]

I can indicate how some sentences might be represented, by intuiting their semantics. These are only intended to give a guide as to the kind of representations that should appear in the truth conditions of sentences.\(^4\)

\[
\begin{align*}
[\text{The judge is strict}] & \simeq (\text{strict}'\text{judge}') (\sigma x \text{judge}'x) \\
[\text{John is liberal}] & \simeq \text{liberal}'\text{j}' \\
[\text{John is the judge}] & \simeq \text{j}' \approx (\sigma x \text{judge}'x)
\end{align*}
\]

The example Landman uses to justify the use of intensional individuals is:

The judge and the cleaner have different incomes.

where John is both judge and cleaner. Presumably he sees this as a problematic example for a property modifier approach as there only appears to be one property — “have different incomes” — which would need to be modified by two arguments. That is, on a property modifier approach, the interpretation of the verb would have to provide for as many modifiers as there are constituent noun phrases in the subject. By having the roles modify the

\(^4\)The theory in Chapter 8 expresses the truth conditions using terms similar to these, but does not require that they are produced directly by a grammar. Indeed, the grammar in Appendix B does not produce these representations.
CHAPTER 7. INTENSIONAL SUBJECTS

interpretation of the noun phrases, then, trivially, there are always the same number of objects to be modified as there are modifiers. However, the appearance of only one verb phrase in the NL sentence does not preclude there being more than one appropriate property in the semantic representation of the sentence. Landman's objection must rest on the assumption that "have different incomes" is represented by some single, irreducible property in the semantics. In particular, his objection must assume that there can be no useful reduction of the phrase modified by the adverb "different". If, in the truth conditions of the sentence, we can provide a property to be modified for each noun phrase, then an interpretation can be given to the sentence using property modifiers. We could paraphrase the sentence as:

The income from cleaning earned by the person who is the cleaner is different to the income from judging earned by the person who is the judge.

with the outcome that:

John's income as a judge is different to his income as a cleaner.

The truth conditions of this paraphrase can be represented as:

\[
\begin{align*}
T(\exists ab\text{income}'a \land \text{income}'b \land \\
\quad \text{different}'(a \oplus b))
\end{align*}
\]

The property different' takes a sum for its argument to ensure a consistent interpretation of conjunction, and to allow arbitrary numbers of terms to be "different". Some meaning postulate such as:

\[
T(\text{different}'(a \oplus b)) \leftrightarrow a \neq b
\]

would have the desired outcome. This could be generalised to arbitrary nominal conjunctions. Although it is not clear to me that we can, or should, attempt to completely decompose the meaning of "different" in the general case.

A hard problem is not just to show that we can give expressions that capture the meaning of the sentence, but also to show that the expression can be derived compositionally. The difficulty is to have the roles, introduced by the noun phrases, modify the appropriate predications. The final theory in Chapter 8, in conjunction with a grammar like that outlined in Appendix B, seeks to ease this difficulty by exploiting the distinction in PT between representations and truth conditions. The NL sentences do not have to be translated directly into terms involving the intensional analogues of the logical connectives (the logical constants). Only the truth conditions of the
representations of sentences need be given in terms of the truth conditions of logical connectives. This has two advantages: NL constituents do not have to be translated into complex λ-abstractions that require meaning postulates or lemmas to unpack; and secondly, the truth conditions of problematic sentences need not be given, they are not forced upon us by the truth conditions of the logical connectives. The example sentence might be represented as:

\[(\text{different}'(\text{incomes'}))(\text{the'judge'} \oplus \text{the'cleaner'})\]

The truth conditions of this term would be given by:

\[
\exists ab(T(((\text{income} a)\text{judge'})\sigma x\text{judge}'x) \& \\
T(((\text{income} b)\text{cleaner'})\sigma x\text{cleaner}'x) \& \\
T(\text{different}'(a \oplus b)))
\]

I will leave a more complete demonstration until Chapter 8.

As mentioned above, Landman maintains a ‘group-forming’ operator ↑ to block distribution: his explicitly intensional individuals cannot control distributive behaviour. He keeps the notion of intensional individuals separate from that of the blocking of distribution, so that he can represent the detailed interpretation of intermediate group readings using the same method as suggested in [Link, 1983]. He is thus not able to explore the possibility of using this same account to permit stage level individuals to be equated with the sums of their parts, nor to control distribution into mass terms. Once we drop any attempt to represent the intermediate distributive readings explicitly (at least in the way adopted by Landman) we can explore the use of intensional “individuals” to control inferences exemplified by distribution, whilst allowing (though not requiring) a choice of ontology where different substances may have the same extension — dirt and water may be parts of dirty water, for example — and stage level individuals may be equated with the sum of their parts. Inferences involving distributive properties can cause problems with such ontologies, because we want to distribute to the correct parts. Landman’s insights enable us to embellish the application of a property to a term so that we can recover these parts. As an example, we might interpret the sentence “some water is liquid”, as predicating “liquid-water” of the mereological sum of some water. That predicate can then distribute to those parts that are water (as opposed to arbitrary parts of the sum).

How do we know under which aspect the application of a property to a term should be restricted? That is, in Landman’s terminology, where do roles come from? With sentences where the subject is not a proper name, one possibility suggests itself: if we take “the man is tall”, then the subject noun-phrase can provide the information that enables us to interpret this as “the man is a tall man”. However, a sentence may be ambiguous as to whether such a device is appropriate:

The cleaners are badly paid.
CHAPTER 7. INTENSIONAL SUBJECTS

may mean either that the cleaners are badly paid cleaners, or that the cleaners are badly paid people. If we take the latter reading, where does the aspect of being a person come from? Perhaps a theory of context and salience could provide the answer. I see several possibilities:

(i) take “badly paid” as ambiguous between being *badly paid as an X* where the subject fills in the X and just plain *badly paid (as a person)*;

(ii) take “The cleaners” as ambiguous between *the cleaners as cleaners, and the cleaners as individuals*;

(iii) have some sortal hierarchy, which says *a cleaner is a person*, and allow individuating terms, optionally, to be generalised;

(iv) some other, more general, theory of context and salience.

It should be noted that, under whichever role we treat the cleaners, we can obtain the correct individuation. It is not, however, possible to take the sentence as ambiguous between providing a role and not, since we must know how to obtain the appropriate nice parts.

Considering sentences involving proper names, when we say that “John is tall” and “Everest is tall”, it should be clear that we are not asserting that John and Everest are both in the class of tall objects. Presumably, we mean “John is a tall person” and “Everest is a tall mountain”, respectively. These could then be taken as forms of ellipsis. But what information do we use to complete the interpretation? The translation of proper nouns can include the role under which they are to be considered, so we could have, in some sense, *the person John* and *the mountain Everest* as our representations of these names. I realise this is something of a cheat: I will offer no theory as to how these interpretations are to be acquired, and how proper names are to be interpreted when first encountered. Indeed, my final theory (Chapter 8) is only concerned with sentences involving quantified noun phrases.

With quantified noun phrases, we can just alter the representation of the determiner so that the appropriate modified predication occurs. The sentence “a cleaner strikes” could be represented as:

$$\exists c (\text{cleaner}'c \land (\text{strikes}'(\text{cleaner}')c))$$

This is true should there be a cleaner who is a striking cleaner (or alternatively, who is striking, as a cleaner).

I should point out that in order to stop distribution at an appropriate point, it is not necessary to use the full intensionality of the property modifier regime: all that is required is for the property which is modified to provide the correct individuation. Any additional intensionality can, in principle, be thrown away. Thus we can always consider cleaners as cleaners, and John as John. However, I think that Landman provides sufficient indication as to why we might not want to do this.
CHAPTER 7. INTENSIONAL SUBJECTS

It may be noted that theories of individuals under different roles may allow descriptors with no extension in the denotable terms to denote the bottom element \( \bot \). Bottom could then be considered under different guises, preventing different non-denoting terms being equated. Note that my formalisation of Landman’s individuals under roles will not allow \( \bot; p \) to be a restricted term \( R \) unless \( T(p, \bot) \).

If guises are added to the theory, either with the aid of property modifiers, or new terms, then presumably these can be inherited by anaphoric reference. For example:

There is [a judge];
[He]; is strict.

should gives rise to:

There is a strict judge.

This would seem easier to formalise if Landman’s restricted individuals are used, as this would be a simple extension of existing discourse theories (for example [Kamp, 1981, Heim, 1982]), rather than property modifiers. However, the property modifiers need only arise in the truth conditions. If the anaphora resolution occurs before this stage, then there may be no call for additional complexity.
Chapter 8

The Final Theory

We are in a position to explore the details of how predicating properties of terms under some guise can be used to: control distribution into mass terms; permit stage level individuals to be equated with the sum of their parts; enable different nominals to have the “same” extension without semantic mishap. I will only examine, in detail, sentences with intransitive verbs and unjoined quantified noun phrases. Proper nouns could be treated along the lines suggested in Chapter 7. Coverage of transitive and ditransitive verbs would require some additional effort (not least in empirically determining the acceptable readings) as there is the additional problem of quantifier scoping ambiguity. This theory produces only non-generic interpretations of NL sentences.

First, the representations of NL sentences used in this chapter are introduced. This is followed by rules that type these representations so, if appropriate, they can be proven to be propositions, and hence carriers of truth values. Axioms are then presented which equate the truth conditions of sentences with the truth conditions of terms involving logical constants. The rest of the chapter is concerned with strengthening the theory to obtain controlled distributive inferences, using property modifiers, together with other appropriate entailment patterns.

I shall not examine conjoined nouns and noun phrases in this semantic theory. This is not because it is impossible to do so, but because it avoids further combinatorial complexity, which would make central aspects of the theory hard to grasp. It is possible to view the distributive/collective distinction over conjoined noun phrases as a case of ellipsis. Note that it is not only verb phrases that can give rise to such ellipsis. The sentence:

The man and woman died.

can be interpreted as:

The man died and the woman died.
where both “died” and “the” to distribute over the conjoined noun. This could be given a syntactic or a semantic treatment. Rather than offer a formal theory of ellipsis here, I shall take it that all representations of sentences have any such ellipsis resolved. The grammar outlined in Appendix B gives a syntactic treatment of ellipsis using a context free grammar with argument-value features. This grammar is not intended to constitute a linguistic theory, but merely to indicate that it is possible to obtain interpretations of sentences where the ellipsis is resolved, and so together with the theory presented here, produce distributive entailments with conjoined noun phrases, syntactic plurals and mass terms.

Representations of Sentences

This chapter assumes that sentences are represented by terms that disambiguate the predicate argument structure. In a sentence like:

The man dies.

the verb is taken to be the main functor, with the noun phrase as its argument. Similarly, the noun phrase itself consists of the determiner, with the noun as its argument:

dies'(the'man')

Adjectives take the noun they modify as an argument, so:

The tall man dies.

is represented by:

dies'(the'(tall'man'))

Adjectives and nouns appearing after the copula are treated in the same manner as verbs, and take the noun phrase as argument:

The bull is black.
The puddle is water.

being represented by:

black'(the'bull')
water'(the'puddle')

respectively.

As mentioned, this theory does not treat conjoined (and disjoined) nouns and noun phrases directly, although one of the later examples does involve a disjunction of adjectives. Conjunction (or disjunction) can be represented by $\oplus$ ($\ominus$, respectively), regardless of the conjoined (disjoined, respectively) categories. For example, assuming that there is no ellipsis, the sentences:
The drunk and alcoholic died.
The blue and red bike disappeared.
All phosphorus is red or white.

can be represented as:

\[
died'(\text{drunk}' \oplus \text{alcoholic}')
\]
\[
disappeared'((\text{blue}' \oplus \text{red}')\text{bike}')
\]
\[
(\text{red}' \odot \text{white}')\text{all'}\text{phosphorus'}
\]

For completeness, closure conditions for arbitrary conjoined and disjoined semantic categories are presented. To some extent, this demonstrates that there is no call for type-shifting rules in the semantics.

As we are interested in the truth conditions of the terms representing sentences, we should first be able to prove that they are terms capable of having truth values, that is, we should be able to prove that they are propositions. To do this, we need to type the semantics of the categories. The simplicity of the semantics proposed requires terms to belong to more than one type. This is possible in PT. Simplifying the typing would require these expressions to be more complex.

**Types**

In principle, we could give arbitrary types to terms representing the objects from the various syntactic categories, so long as the representations of well-formed sentences were propositions. However, when we come to consider the truth conditions of sentences — which will be given using expressions that more closely mirror Montague style semantics (as in [Turner, 1992]) — it is useful if the syntactic categories are represented by terms of the more conventional types. Thus, although the proposal above is to represent nouns with terms that do not have any arguments, when the truth conditions are given, they will behave as predicate-like objects, as is usual in a Montague-style treatment.

**Lemma 8.1** The denotations of nouns are of the type $\text{Pty}_{\Delta}$.

Using the same motivation, the type corresponding to the category of the determiners will be akin to that used in more conventional representations (restricted to the denotables). This will result in noun phrase being represented by quantifiers (of denotables). However, the definite descriptor “the” must be given a more complex type, compatible with the proposed treatment of non-denoting terms.

**Lemma 8.2** Except for the definite descriptor, the interpretations of determiners are of the type $\text{Det}_{\Delta}$.
The type of the definite descriptor “the” should be restricted so that it conforms to the goal of not being able to prove propositionhood of a sentence when it contains a definite descriptor which “fails to denote”. Thus the denotation of “the” should only form a quantifier with a property of denotables which has an extension in the denotable terms.

**Lemma 8.3** The interpretation of the definite descriptor is of the type

\[ \forall p((\text{Pty}_\Delta(p) \& \exists x (\text{T}(px))) \rightarrow \text{Quant}_\Delta(\text{the}p)) \]

Sentences are to be represented by propositions (type P), and are formed by a verb phrase and a noun phrase (type \( \text{Quant}_\Delta \)). As, in the suggested representation, the noun phrase is the argument of the verb phrase, then verb phrases should be represented by terms of type \( \text{Quant}_\Delta \rightarrow P \). The only verb phrases I shall consider (ignoring copula phrases) are intransitive verbs, so these will also be represented by terms of that type. This means that the syntactic rule which allows an intransitive verb to be considered as a verb phrase has no effect on the semantic type.

**Lemma 8.4** Intransitive verbs are of the type \( \text{Quant}_\Delta \rightarrow P \).

As, in the simple representation that I have proposed, nouns can be turned into verbs with the copula, then they can also be in this type. The weak typing of PT allows terms to belong to more than one type. An alternative would be to give the copula an explicit representation with a term that has a type changing function.

**Lemma 8.5** Nouns are of the type \( \text{Quant}_\Delta \rightarrow P \).

Adjectives can modify nouns. In the semantics, an adjective should take a property (of denotables) and produce a new property (of denotables).

**Lemma 8.6** Adjectives are in the type \( \text{Pty}_\Delta \rightarrow \text{Pty}_\Delta \).

In the proposed representation, adjectives can also behave like a verb phrase, when they appear after the copula, so they can also be of the same type as intransitive verbs. As mentioned before, a more linguistically principled theory might give the copula the explicit function of modifying the type of adjectives, and nouns, to that of verb phrases.

**Lemma 8.7** Adjectives are in the type \( \text{Quant}_\Delta \rightarrow P \).

This is sufficient to prove propositionhood of the semantics of sentences without conjoined nouns. Other axioms are required to type the sums and products of terms. These will strengthen the closure conditions of types. The motivation is to mirror the syntactic closure conditions for conjoined
and disjoined categories in a grammar — like that given in Appendix B — with closure conditions on the semantic types that will correspond to the representations of these categories. When we conjoin or disjoin two nouns, the result is a noun. As the representation of a noun, at some level, corresponds to a property of denotables, then if, in the semantics, we form the sum or product of two properties of denotables, then the result will be a new property of denotables. This is expressed in the following two axioms.

**Axiom 8.1** The sum of two properties of denotables, is a property of denotables.

\[(\text{Pty}_\Delta(r) \& \text{Pty}_\Delta(w)) \rightarrow \text{Pty}_\Delta(r \oplus w)\]

**Axiom 8.2** The product of two properties of denotables, is a property of denotables.

\[(\text{Pty}_\Delta(r) \& \text{Pty}_\Delta(w)) \rightarrow \text{Pty}_\Delta(r \otimes w)\]

It will be useful to indicate how to derive the truth conditions of complex properties of denotables, applied to denotables. The extension of a conjunction (or disjunction) of properties of denotables is the intersection (union, respectively) of the extensions of the conjoined (disjoined, respectively) properties.

**Axiom 8.3** The sum of properties of denotables, applied to a denotable, is true if the conjunction of those properties, applied to that denotable, is true.

\[(\text{Pty}_\Delta(r) \& \text{Pty}_\Delta(w) \& \Delta(x)) \rightarrow (T(r \oplus w)x \leftrightarrow T(rx) \& T(wx))\]

**Axiom 8.4** The product of properties of denotables, applied to a denotable, is true if the disjunction of those properties, applied to that denotable, is true.

\[(\text{Pty}_\Delta(r) \& \text{Pty}_\Delta(w) \& \Delta(x)) \rightarrow (T(r \otimes w)x \leftrightarrow T(rx) \lor T(wx))\]

The closure conditions for the conjunction and disjunction of terms of the type \(\text{Quant}_\Delta\) are given next.

**Axiom 8.5** The sum of two quantifiers of denotables is a quantifier of denotables.

\[(\text{Quant}_\Delta(f) \& \text{Quant}_\Delta(g)) \rightarrow \text{Quant}_\Delta(f \oplus g)\]

**Axiom 8.6** The product of two quantifiers of denotables is a quantifier of denotables.

\[(\text{Quant}_\Delta(f) \& \text{Quant}_\Delta(g)) \rightarrow \text{Quant}_\Delta(f \otimes g)\]

Finally, the closure conditions on propositions are:
Axiom 8.7 The sum of two propositions is a proposition.

\[(P(p) \& P(q)) \rightarrow P(p \oplus q)\]

Axiom 8.8 The product of two propositions is a proposition.

\[(P(p) \& P(q)) \rightarrow P(p \otimes q)\]

The sums and products of propositions will correspond to the conjunctions and disjunctions of sentences. Sentential conjunction and disjunction will have the standard interpretations. This can be enforced with:

Axiom 8.9 The sum of two propositions is a true proposition if both propositions are true.

\[(P(p) \& P(q)) \rightarrow (T(p) \& T(q) \leftrightarrow T(p \oplus q))\]

Axiom 8.10 The product of two propositions is a true proposition if one of the factor propositions is true.

\[(P(p) \& P(q)) \rightarrow ((T(p) \lor T(q)) \leftrightarrow T(p \otimes q))\]

Truth Conditions

Were we not using property modifiers, plurals and mass terms, then the truth conditions of the representations of simple sentences — for example, of the form: “every q p” and “some q p” — could be given as reflecting the more conventional interpretations of the quantifiers as terms involving logical constants as suggested by Turner [Turner, 1992]. Thus we would have, for example:

\[
P(p(\text{every}'q)) \rightarrow (T(p(\text{every}'q)) \leftrightarrow T(\Theta k x(qx \Rightarrow px)))
\]

\[
P(p(\text{some}'q)) \rightarrow (T(p(\text{some}'q)) \leftrightarrow T(\Xi k x(qx \land px)))
\]

As we are using property modifiers — in particular to keep track of the relevant property for determining the nice parts to which properties can distribute — the truth conditions must be complicated somewhat. Further, quantifiers like “all” and “some” must not result in properties distributing to the parts of the representation of the noun phrase. The quantifier “some” in a sentence should give rise to quantification over collections in the sentence’s truth conditions. The quantifier “all” is more problematic: in the truth conditions of a sentence it should not give rise to universal quantification over collections, as this gives rise to distributivity. It must thus be treated more akin to the definite descriptor. I shall allow sentences like “all q p” to be true if there are no qs. The truth conditions can be stated with the following axioms:
Axiom 8.11

\[ P(p(\textit{every'}q)) \rightarrow (T(p(\textit{every'}q)) \leftrightarrow T(\Theta x(qx \Rightarrow pqx))) \]

Axiom 8.12

\[ P(p(\textit{some'}q)) \rightarrow (T(p(\textit{some'}q)) \leftrightarrow T(\Xi x(qx \land pqx))) \]

Axiom 8.13

\[ P(p(\textit{all'}q)) \rightarrow (T(p(\textit{all'}q)) \leftrightarrow (\Delta(\sigma xqx) \rightarrow T(pq(\sigma xqx)))) \]

Axiom 8.14

\[ P(p(\textit{the'}q)) \rightarrow (T(p(\textit{the'}q)) \leftrightarrow T(pq(\sigma xqx))) \]

The truth conditions for \textit{every'}pq implicitly require the notion of atomicity (or singularity), that is, \( q \) only ranges over singular individuals, not mass terms or plurals. This could be guaranteed by a semantic restriction in the truth conditions. Indeed, I define a semantic notion of singularity later. Here, however, I shall assume that any grammar used to obtain the representations has syntactically restricted the use of “every” to semantically singular terms.

In order to make the right-hand sides of the biconditionals in these axioms provably propositions, independently, the verbs must be able to act as property modifiers:

Lemma 8.8 The denotations of verbs are in the type \( Pty_{\Delta} \implies Pty_{\Delta} \)

As nouns can also act as verb-like terms when appearing after the copula, they can also be in this type.

Lemma 8.9 The denotations of nouns are in the type \( Pty_{\Delta} \implies Pty_{\Delta} \)

If verb phrases had been typed independently of intransitive verbs, and the type changing function of the copula was made explicit, then this last lemma would become redundant.

Property Modifiers

Considering terms of the form \textit{died’men’x}, as mentioned before, I shall take \textit{died’} here to be the semantic equivalent of a subselective adjective. Although I use the word \textit{adjective}, I am referring to a semantic category which need not be limited to the representation of natural language adjectives. This is not new. For example, Hoepelman has suggested that nouns may, in the semantics, be treated as a special case of adjective [Hoepelman, 1983].

The apparent similarity between the representation of roles and that of adjectives is important. In particular, it is fruitful to look at the kinds of behaviours defined in work on the semantics of adjectives, as it transpires that
these behaviours are also useful in explaining how to obtain intuitively useful inferences in a semantic theory which uses property modifiers to represent Landman’s notion of roles. In the remainder of this chapter, in order to make some of the examples work, property modifiers are given behaviours which, in isolation, might seem rather ad hoc, however the suggested inferences have independent motivation from considerations of the semantics of NL modifiers, and would already be required in a comprehensive semantics for NL.

As PT is weakly typed, we can, to some extent, avoid deciding whether adjectives should be exclusively property modifiers, or properties.\footnote{For a comparison of these two approaches (considering ‘predicates’ rather than ‘properties’) see [Kamp, 1975].} There is no need for formal devices, such as operators or dummy predicates, to change the type of a subsective adjective (and general property modifiers).

I shall assume that properties introduced in the representation of natural language are typically properties of denotables, and that property modifiers are properties of denotables, which can also modify other properties to produce a new properties.\footnote{However, I do feel that it would be fruitful to formalise a version of this theory where the property modifiers are constructed from properties using some operator. This has been suggested to me by Gennaro Chierchia. In such a theory, rather than represent ‘John is a strict judge’ with something like (strict’judge’\( j \)), where strict’ is both a property of denotables, and a property modifier, we would use (as’judge’strict’\( j \)), where judge’ is only a property of denotable, and as’ is a function that turns such a property into a property modifier. I suspect that this would simplify some aspects of the theory, though perhaps at the cost of a unification with a treatment of NL modifiers.}

To motivate the formal definition of a property modifier, if we take “red” to be a subsective adjective (that is, a property modifier), then it may produce a proposition, given some natural language denotable term, which is true if that denotable is red. But given some other property, like “book”, a new property is created, that of being a red book. Unless extra conditions are added, a red book need not be red, but it is a book.

**Definition 8.1** A property modifier \( r \) is a property of denotables which, given a property of denotables \( p \), forms a new property of denotables \( rp \), and if that new property holds of some denotables, then the original property of denotables \( p \) also holds of it.

\[
\mathcal{PM}(r) =_{df} \text{pty}_{\Delta}(r) \& \forall p(\text{pty}_{\Delta}(p) \rightarrow (\text{pty}_{\Delta}(rp) \& \forall x(T(rpx) \rightarrow T(px))))
\]

As modifiers (both semantic and syntactic) can be conjoined and disjoined, it is necessary to add typing rules to make \( \mathcal{PM} \) closed under \( \oplus \) and \( \odot \).

**Axiom 8.15** The sum of two property modifiers is a property modifier.

\[
(\mathcal{PM}(r) \& \mathcal{PM}(w)) \rightarrow \mathcal{PM}(r \oplus w)
\]
Axiom 8.16 The product of two property modifiers is a property modifier.

\[(\mathcal{PM}(r) \& \mathcal{PM}(w)) \rightarrow \mathcal{PM}(r \otimes w)\]

I will not show how these complex modifiers affect the truth conditions of propositions in which they appear until Example 8.4, towards the end of this chapter.

Taking the suggestion that some of the behaviours of adjectives are desirable for property modifiers, we can define Bunt’s notions of distributive, collective, and homogeneous adjectives [Bunt, 1979].

Definition 8.2 A term \( r \) is cumulative (a cumulative modifier) iff it is a property modifier, and when restricting a property (of denotables) \( p \), if it holds of denotables \( x, y \), it also holds of the sum \( x \oplus y \).

\[
\mathcal{C}_{\mathcal{PM}}(r) =_{df} \mathcal{PM}(r) \& \forall p \forall x \forall y ((\text{Pty}_{\Delta}(p) \& T(rp) \& T(rp)) \rightarrow Trp(x \oplus y))
\]

This can be generalised to arbitrary (definable) suprema.

Definition 8.2’

\[
\mathcal{C}_{\mathcal{PM}}(r) =_{df} \mathcal{PM}(r) \& \forall p \forall q (\text{Pty}_{\Delta}(p) \& \text{Pty}_{\Delta}(q) \& \\
\forall \Delta x (T(qx) \rightarrow T(rp)) \rightarrow T(rp(\sigma_{\Delta}qx)))
\]

The following notion of distributive property modifiers is central to this theory’s treatment of distributive inferences into denotable terms. Essentially, distribution must be restricted to the appropriate nice parts of a term. When a denotable \( x \) appears as the argument of the property consisting of the modifier \( r \) and the property \( p \) — that is \( rpx \) — then if \( r \) is distributive, the appropriate nice parts of \( x \) to distribute to are those that have the property \( p \). In general, we must not assume that \( r \) itself distributes to these parts, but allow \( rp \) to distribute (although for some pairs \( r, p \), if \( rp \) holds, then we may infer \( r \) holds).

Definition 8.3 A term \( r \) is distributive (a distributive modifier) iff it is a property modifier, and when restricting any property (of denotables) \( p \), if it holds of denotables \( x \), it also holds of any part \( y \) of the denotable \( x \), which is also in the extension of \( p \).

\[
\mathcal{D}_{\mathcal{PM}}(r) =_{df} \mathcal{PM}(r) \& \forall p \forall x \forall y ((\text{Pty}_{\Delta}(p) \& T(rp) \& T(py) \& y \leq x) \rightarrow Trp(y))
\]

\(^3\)Obviously, I assume the adjectives in question are subsective, and that they act as property modifiers.
CHAPTER 8. THE FINAL THEORY

If we say “the water is liquid”, and take the truth conditions of this to be interpreted as:

\[ T(\text{liquid}'\text{water}'\sigma \text{water}'x) \]

then — assuming that liquid’ is distributive — any part of \( \sigma \text{water}'x \) that is water, will be liquid water. There is a potential problem with this, depending upon our view of the meaning of “liquid”. For this treatment of distribution to work, it must be assumed that the smallest denotables that have the property of being “water” must still be large enough to possess any distributive properties — like ‘liquid water’ — that a significant amount of that denotable can have. This treatment of distribution would fail if we consider a molecule of water to be water, yet only allow a more substantial body of water to be liquid. Further, if we consider “being liquid” to be a property of a collection of molecules that are a physically continuous body, then this account would fail if denotable water could consist of the fusion of discontinuous fragments of water, each fragment being too small to be attributed the property of being liquid.\(^4\) Such criticisms seem to assume a reductive account of the extension of mass terms to notions in modern physics. We can maintain that the denotables that we can refer to by a term like “water” are precisely those terms to which we can attribute properties such as “being liquid”. We could view an attempt to ascertain the smallest such denotable in terms of physical molecules as rather like trying to give a reductive account of the meaning of “a table” in terms of a physical arrangement of components.

We might look more closely at the notion of smallest part in modern physics. It tells us that the smallest element of water is a water molecule, precisely because a water molecule is the smallest physical component of water that possesses the properties that distinguish it as water, for the purposes of theories in physical science. Should we accept that the property of “being liquid” cannot be attributed to molecules in isolation, then for the purposes of this theory of semantics, the smallest component of water that counts as water will have to be something more than a molecule. This does not prevent reference to molecules of water, as this sentence verifies. According to this theory, we might then say “the water is liquid” without entailing that “the molecules of water are liquid”.

**Definition 8.4** A term \( r \) is homogeneous (an homogeneous modifier) iff it is cumulative and distributive.

\[ \mathcal{H}_{\mathcal{PM}}(r) =_{df} \mathcal{C}_{\mathcal{PM}}(r) \& \mathcal{D}_{\mathcal{PM}}(r) \]

Note that these distributive and cumulative notions are different from those given by \( * \) and \( D \), which produce collective and distributive properties respectively [Link, 1991a, Lønning, 1989]. However, it seems to be the case

\(^4\)These criticisms were suggested to me by Hans Kamp at a talk I gave on plurals and mass terms, in Stuttgart, February 1993.
that whenever the representation of a word is a cumulative property modifier, as defined above, it is typically a cumulative property of denotables, in the following sense:

**Definition 8.5** A *strongly cumulative term* is a cumulative property of denotables.

\[ \mathcal{C}_S(s) =_{df} \mathsf{Pty}_\Delta(s) \& \forall x t (T(st) \leftrightarrow t = \sigma_x(x \land x \ll t)) \]

Typically, those words that are considered to be homogeneous will be homogeneous modifiers, and cumulative properties:

**Definition 8.6** A *strongly homogeneous term* is an homogeneous modifier, and a strongly cumulative property.

\[ \mathcal{H}_S(S) =_{df} \mathcal{H}_M(s) \& \mathcal{C}_S(s) \]

**Singles and Plurals**

Some plurals are also strongly homogeneous. We may wish to relate the interpretation of a plural to that of its syntactic singular. A singular term is not homogeneous. This result can be achieved as follows:

**Definition 8.7** A *singular* is a property of denotables which, if it holds of a denotable, does not hold of any of its parts.

\[ \mathsf{Sing}(s) =_{df} \mathsf{Pty}_\Delta(s) \& \forall x (T(px) \rightarrow \exists y (y \leq x \& T(py) \& y \neq x)) \]

If \( s \) is a singular property, we can represent its (improper) plural form with \( \tau s \).\(^5\) The intention of the next axiom is to state that any supremum of terms that have the singular property \( s \) will have the plural property \( \tau s \).

**Axiom 8.17** If some denotables each have the singular property \( p \), then the supremum of those denotables has the plural property \( \tau p \).

\[ \forall p (\mathsf{Sing}(p) \rightarrow \forall q (\mathsf{Pty}_\Delta(q) \& \forall x (T(qx) \rightarrow T(px)) \rightarrow T(\tau p(\sigma x q x)))) \]

This is a generalisation of the intuition that the sup of two terms with the singular property \( s \) have the plural property \( \tau s \):

\[ \forall p (\mathsf{Sing}(p) \& T(pa) \& T(pb) \rightarrow T(\tau p(a + b))) \]

If a plural property holds of a term, but does not hold of any proper part of that term, then it must be the case that the underlying singular property also holds of the term. This is effectively the converse of the previous axiom. It means that part of a term in the extension of a plural property must have the underlying singular property.

\(^5\) The improper plural form subsumes the singular form: \( \tau p \) will hold of one or more \( ps \), as opposed to the two or more with English plurals.
Axiom 8.18 If a denotable has the property \( \neg p \), (where \( p \) is a singular property) and no proper part of it does, then it also has the property \( p \).

\[
\forall p \forall x (\text{Sing}(p) \land T(\neg px) \land \exists y (y \leq x \land y \neq x \land T(py))) \rightarrow T(px)
\]

Taking these last two axioms together, a plural property holds of a term if, and only if, that term is the fusion of terms having the singular property.

The equivalence:

\[
\sigma px = \sigma x' px
\]

can be proved.

These axioms would be satisfied by the cumulative operator \(^*\). However, in this formal theory, we also wish to be able to have plural forms of property modifiers.

Axiom 8.19 Property modifiers have plural forms.

\[
\forall p ((\exists M(p) \land \text{Sing}(p)) \rightarrow \exists M(\neg p))
\]

Typically, when a plural property modifier is modifying a singular property, and the resultant property holds of a term, then the singular property modified by the singular property modifier also holds of that term. This is formalised in the following axiom:

Axiom 8.20

\[
\forall pq \forall x ((\exists M(p) \land \text{Sing}(p) \land \text{Pty}(q) \land \text{Sing}(q) \land T(\neg px)) \rightarrow T(pqx))
\]

This results in a useful inference, allowing the representation of “the men die” to entail the representation of “every man dies” (assuming there are some men). This will be elaborated in Example 8.1, below.

If the semantics of certain words is ambiguous with respect to whether they are distributive or not, operators could be added that relate the distributive term to the non-distributive representation. I have not done this: it is trivial but combinatorially complex.

Examples and Further Axioms

Example 8.1 I shall demonstrate the use of these notions with the sentence “the men die”, by showing that its truth conditions (assuming there are some men, so that it forms a proposition):

\[
T(\text{die'men'}(\sigma e \text{men'x}))
\]

allow the derivation of the truth conditions for “every man dies”:

\[
\forall x (T(\text{man'x}) \rightarrow T(\text{die'man'x}))
\]
CHAPTER 8. THE FINAL THEORY

I will take the two terms, men' and die' to be strongly homogeneous property modifiers. Further, I will take it that men' = "man", and die' = "dies". From the assumption that die' is a property modifier, we can infer:

\[ T(\text{die'} \text{men}'(\sigma_\text{e}x \text{men}'x)) \rightarrow \]
\[ \forall y(T(\text{men}'y) & y \leq (\sigma_\text{e}x \text{men}'x) \rightarrow T(\text{die'} \text{men}'y)) \]

Given the homogeneous behaviour of men, and its relation to its singular, we can show:

\[ T(\text{die'} \text{men}'(\sigma_\text{e}x \text{man}'x)) \rightarrow \]
\[ \forall y(T(\text{man}'y) & y \leq (\sigma_\text{e}x \text{man}'x) \rightarrow T(\text{die'} \text{man}'y)) \]

From the axioms concerning \( \sigma \), this gives:

\[ T(\text{die'} \text{men}'(\sigma_\text{e}x \text{man}'x)) \rightarrow \forall y(T(\text{man}'y) \rightarrow T(\text{die'} \text{man}'y)) \]

From the fact that die' is the plural of dies', Axiom 8.20 allows us to infer that the internal consequent implies:

\[ T(\text{dies'} \text{man}'y) \]

Thus:

\[ T(\text{die'} \text{men}'(\sigma_\text{e}x \text{men}'x)) \rightarrow \forall y(T(\text{man}'y) \rightarrow T(\text{dies'} \text{man}'y)) \]

This example shows that even with a semantic theory intended to cover some awkward examples, our intuitions concerning simpler cases are supported, if rather indirectly.

It is possible to define notions corresponding to transparent, or predicative adjectives, like those presented by Bunt and Roeper [Bunt, 1985, Roeper, 1983]. The behaviour defined is required to cope with Bunt’s “wet puddle” argument, which is formalised in the next worked example.

**Definition 8.8** A term \( r \) is transparent (a transparent modifier) with respect to a property (of denotables) \( p \), iff it is a property modifier, and when restricting \( p \), if it holds of denotables \( x \), it also holds of \( x \) by itself. Further, if \( r \) and \( p \) hold of \( x \), by themselves, then \( rp \) holds of \( x \).

\[ T_p(r) =_{def} \mathcal{P}\mathcal{M}(r) & P\text{ty}_\Delta(p) & \forall x(T_p x \leftrightarrow (Tr x & T p x)) \]

This is rather like Hoepelman’s notion of strongly predicative adjectives, except that here transparency is indexed to particular properties of denotables, whereas strongly predicative adjectives are transparent with respect to all properties (of denotables). His notion of weakly predicative adjectives cannot be expressed in this theory as it stands, as there is no formal notion of
polar opposites (tall v. short, for example).\(^6\) We may wish to maintain that all property modifiers are transparent with respect to themselves, echoing one of Landman’s axioms (given as axiom (ii) in Chapter 7), which can be paraphrased as “John as a judge is a judge” [Landman, 1989]:

\[ \forall p (\mathcal{M}(p) \rightarrow T_p(p)) \]

This does not cause the collapse of compound property modifiers, which might be useful in a theory of adjectives: Hoepelman, for example, would use tall’tall’man’\(x\) to indicate a man who is tall for a tall man. If, however, we take chessplayer’ to be a transparent property modifier, it would cause the collapse of chessplayer’chessplayer’\(x\) to just chessplayer’\(x\). Hoepelman would prefer to use this to indicate chessplayers who are relatively good at chess, in a body of other chessplayers [Hoepelman, 1983].

We are now in a position to consider some more examples. The next two show how the theory can address Bunt’s desiderata for a theory of mass terms §2.3 [Bunt, 1985].

**Example 8.2** Considering the argument:

All water is wet  
The puddle is water
\[ \therefore \text{The puddle is wet} \]

Assuming that there is a puddle, the truth conditions of the three sentences are given by:

\[ T(\Theta_{\phi}x (\text{water'}x \Rightarrow \text{wet'water'}x)) \]
\[ T(\text{water'}\text{puddle'}(\sigma_{\phi}\text{puddle'}x)) \]
\[ T(\text{wet'}\text{puddle'}(\sigma_{\phi}\text{puddle'}x)) \]

After applying the axioms of truth, the argument becomes:

\[ \forall_\Delta x (T(\text{water'}x) \rightarrow T(\text{wet'}\text{water'}x)) \]
\[ T(\text{water'}\text{puddle'}(\sigma_{\phi}\text{puddle'}x)) \]
\[ \therefore T(\text{wet'}\text{puddle'}(\sigma_{\phi}\text{puddle'}x)) \]

The terms wet’ and water’ can be taken to be transparent, with respect to each other and puddle’.\(^7\) From the transparency of water’ with puddle’, and the second premise, we have:

\[ T(\text{water'}(\sigma_{\phi}\text{puddle'}x)) \]

\(^6\)Hoepelman suggests we might take ‘red’ to be weakly predicative: a red tomato may be red compared to the existent tomatoes, but we might not wish to class it as unadorned “red”. It is weakly predicative as a red tomato is definitely not “unred”, the polar opposite of “red”, (green, blue, black, etc.), and a tomato that is truly red is a red tomato.

\(^7\)Note that if we additionally take wet’ to be distributive, it does not mean that “wet” distributes to all parts. The distribution is only motivated when the distributive ‘property’ appears as a property modifier, so it can still only distribute to relevant parts.
and thus, from the first premise, we obtain:

\[ T(\text{wet'} \text{water'}(\sigma_5 \text{xpuddle'}x)) \]

From this and the transparency of wet' with water', we have:

\[ T(\text{water'}(\sigma_5 \text{xpuddle'}x)) \]

and from the transparency of wet' with puddle':

\[ T(\text{wet'} \text{puddle'}(\sigma_5 \text{xpuddle'}x)) \]

which is what was wanted.

\[ \bullet \]

**Example 8.3** Taking the sentence “all water is water”, the truth conditions of its representation are given by:

\[ \Delta(\sigma x \text{water'}x) \rightarrow T(\text{water'} \text{water'}(\sigma x \text{water'}x)) \]

Assuming that there is some water, the truth conditions of the sentence are dependent upon:

\[ T(\text{water'} \text{water'}(\sigma x \text{water'}x)) \]

From the transparency of water', with respect to itself\(^8\), and our assumption that it can also be treated as a property modifier, we have:

\[ \forall \Delta x (T(\text{water'}x) \leftrightarrow T(\text{water'} \text{water'}x)) \]

From the homogeneity of water', and axioms for the supremum operator (assuming that there is some water), we can infer that:

\[ T(\text{water'}(\sigma x \text{water'}x)) \]

Taking these two results together, we can infer:

\[ T(\text{water'} \text{water'}(\sigma x \text{water'}x)) \]

on the assumption that there is some water, thus we obtain the desired result (cancelling the assumption):

\[ \Delta(\sigma x \text{water'}x) \rightarrow T(\text{water'} \text{water'}(\sigma x \text{water'}x)) \]

Thus “All water is water” is true.

\[ \bullet \]

\(^8\)We cannot claim that it is transparent, period, as there is a counter example: a water meadow is not water.
CHAPTER 8. THE FINAL THEORY

So the results that Bunt considers to be essential for a theory of mass terms are obtained.

Next, it is shown that the theory can cope with Roeper’s ‘phosphorus’ example, containing disjunction of properties, provided the theory is strengthened to indicate how the disjunction (product) of property modifiers affects the truth conditions. Following this, I elaborate on the behaviour of conjoined (summed) property modifiers.

Example 8.4 In the theory as it stands, the truth conditions of the sentence “all phosphorus is red or white” [Roeper, 1983]:

\[ T(\Theta_b x (\text{phosphorus}' x \Rightarrow (\text{red} \land \text{white}') \text{phosphorus}' x)) \]

cannot be unpacked as far as we might like. To give a full elaboration of the truth conditions requires an additional axiom for the product of subsective adjectives.

**Axiom 8.21** A property of denotables \( p \) modified by the product of two property modifiers \( r, w \) holds of a denotable \( x \), iff either (1) the property of denotables, modified by either property modifier, holds of that denotable; or (2) that denotable can be divided in two \( u, v \), such that the property of denotables holds of one part \( u \), when modified by the first property modifier, \( r \), and the other part \( v \), when modified by the second, \( w \).

\[
\forall x \forall w \forall \Delta x (\varphi_M(r) \land \varphi_M(w) \land \text{pty}_\Delta p \Rightarrow
\quad (T((r \land w) x r) \leftrightarrow \begin{cases} T(r w x) v \\
T(w r x) v \\
\exists \Delta \lambda w (x = u \land v \& T(r u x) \land T(w v x)) \end{cases})
\]

Thus, “all phosphorus is red or white” is true, on this narrow scope reading of the disjunction, if it is the case that any fusion of phosphorus is red; or it is white; or if it has two parts that are phosphorus, and one part is white, and the other is red. Note that this narrow scope reading subsumes the wide scope reading of the disjunction (“all phosphorus is red, or all phosphorus is white”).

I think that we can say little about the case of summed property modifiers. If something is black and white, it may be acceptable to say that there are parts of it that are black, and parts that are white. However, in dissecting the object, these attributions of colour may be invalid.\(^9\) If we take the sentence:

John is an angry and hateful person.

\(^9\)Confusion is increased with uses of “black and white” such as “black and white television”, where the object may be neither be truly said to be black, or white.
it is not easy to contemplate the idea that there are necessarily parts of John which are angry, and other parts which are hateful. We could account for this sentence by assuming movement has occurred (in the syntax) from “John is an angry person and John is a hateful person”. However, we may then posit movement in cases of exclusive properties, black and white, for example. These readings could be ruled-out by semantic considerations. In which case, we might also want the constraint that \((r \oplus w) p\) can only be a proposition if \(r p, w p\) are exclusive properties.

With or without this constraint, I am fairly certain that we can have the following:

**Axiom 8.22** If a property of denotables \(p\), modified by a property modifier \(r\), holds of a denotable \(u\), and modified by another property modifier \(w\), holds of another denotable \(v\), then the property of denotables, modified by the sum of the property modifiers \(r \oplus w\), holds of the (denotable) sum of denotables \(u \oplus v\).

\[
\forall wp \forall \Delta x ((\mathcal{P}M(r) \& \mathcal{P}M(w) \& \mathcal{P}t(Y, p)) \rightarrow (\mathcal{T}(r w u) \& \mathcal{T}(w v)) \rightarrow \mathcal{T}((r \oplus w)p(u \oplus v)))
\]

In the case when \(r p, w p\) are exclusive properties, we may also have the converse of this.

**Axiom 8.23** If two property modifiers \(w, r\) are exclusive, when modifying a property of denotables \(p\), then if the sum \(w \oplus r\) modifying \(p\) holds of a denotable \(x\), then \(x\) is the sum of denotables \(u, v\), where \(w p\) holds of \(u\) and \(r p\) holds of \(v\).

\[
\forall wrp \forall \mathcal{P}w \& \mathcal{P}r \& \mathcal{P}t(Y, p) \& \forall \Delta x (\mathcal{T}(w x ) \leftrightarrow \sim \mathcal{T}(r x ) ) \rightarrow \\
\forall \Delta x (\mathcal{T}(w \oplus r \Theta x ) \rightarrow \exists \mathcal{P}_{u \oplus v}(x = u \oplus v) \& \mathcal{T}(w p u) \& \mathcal{T}(r p u)))
\]

The next example shows how the theory allows contradictory properties to be attributed to objects, even when their extensions are equated. Although with the example chosen, we might prefer not to equate the extensions for philosophical reasons, it will serve as an illustration.

**Example 8.5** The well-worn sentences:

The gold ring is new.
The gold is old.

become:

\[
\text{new}'(\text{the}'(\text{gold}'\text{'ring'}))
\]
\[
\text{old}'(\text{the}'\text{'gold'})
\]

Assuming there is some gold and a gold ring, the truth conditions of these terms are given by:

\[
\mathcal{T}(\text{new}'(\text{gold}'\text{'ring'})(\sigma_x(\text{gold}'\text{'ring'}), x))
\]
\[
\mathcal{T}(\text{old}'\text{'gold'}(\sigma_x\text{'gold'} x))
\]
Even if the gold is realised by the gold ring\textsuperscript{10}:

\[ \sigma_5 x (\text{gold'} \text{ring'}) x = \sigma_5 x \text{gold'} x \]

there is no intrinsic contradiction in the truth conditions of these sentences. ●

The final example shows how the theory copes with the comparative sentences which Landman uses to argue against a property modifier treatment of roles [Landman, 1989].

**Example 8.6** Take the comparative sentence:

The judge and the cleaner earn different incomes.

Essentially, its treatment in the theory rides on the meaning of “earn different incomes”. In a strongly typed logic, it may be hard to see how “earn different incomes” can be expressed. Weakly typed logics make it easier, the expression is just earn’(different’incomes’). We can represent the sentence as:

\[ \text{earn'}(\text{different' incomes'})((\text{the'} \text{judge'}) \oplus (\text{the' cleaner'})) \]

where:

\[ T(\text{earn'}(\text{different' incomes'})((\text{the'} \text{judge'}) \oplus (\text{the' cleaner'}))) \leftrightarrow \\
T(\exists a b \text{ (income'} a \land \text{income'} b \land \\
\text{earn' } a(\text{the'} \text{judge'}) \land \text{earn' } b(\text{the' cleaner'}) \land \\
\text{different' } a b)) \]

The meaning postulate concerning “earning different incomes” may be generalised to arbitrary summation, and sums of plurals. However, along with other collective properties, it may be preferable not to attempt a general reduction of its meaning. ●

The treatment of some of the examples is perhaps rather indirect and involved. However, the main reason for giving the details is to show that the ideas developed can be used in compositional semantics. Clearly some additional effort must be expended to extend the coverage of this theory to transitive and ditransitive verbs, and to incorporate a dynamic component to produce a more reductive analysis of intermediate distribution as suggested by Schwarzschild [Schwarzschild, 1990, Schwarzschild, 1992].

---

\textsuperscript{10}The grammar presented in Appendix B offers no compositional derivation of this.
Chapter 9

Conclusions

The thesis explores the semantics of mass terms and plurals in an axiomatic property-theoretic framework.

Summary

The initial theory of plurals in PT (Chapter 5) demonstrates how an axiomatic approach to NL semantics allows the formalisation of a deliberately incomplete theory which is undecided as the nature of sentences involving ‘non-denoting’ definite descriptors. This weak theory is compatible with a treatment of helpful answers to questions, and supports a theory of presuppositions.

The theory can be further weakened (Chapter 6), so that it is incomplete with respect to the existence of atomic individuals. This appears to be a suitable representation for the extensions of both mass and plural nominal terms. It is used for this purpose in the final theory.

This mereological theory does not seem suitable for representing mass terms directly as it would require the extensions of mass terms to be homogeneous in order to allow for distributive inferences without distributing properties to inappropriate parts. It is not clear that homogeneity satisfies our intuitions, especially when considering composite mass terms. Rather than accept homogeneity of extensions merely to obtain the correct behaviour in one formal framework, it seems desirable to find a framework that can directly control distributive inferences, thus separating the issues of homogeneity and distribution. To this end, it seems desirable for the representation of NL nominals to allow us to infer the relevant nice parts of its extension to which a property can distribute. Landman’s intensional individuals — where plain individuals (the underlying extensions of nominal terms) are combined with the role under which they are being considered — appear to give sufficient information for this. The weakly typed PT allows us to simplify Landman’s formal theory, as there is no call for type-shifting. However, rather than adopt a universe of new individuals modified by their various roles, it is pos-
sible to represent roles using property modifiers (Chapter 7). Such modifiers are already required for the semantics of adjectives and adverbs. Again, PT’s weak typing allows a simple treatment of such modifiers.

In the final theory (Chapter 8), the semantics of the verb phrase (acting as a property modifier) modifies the representation of the subject noun (a property). Distributive inferences are performed by allowing the resultant modified property to hold of any part of the denotable term to which it is applied, provided that the parts are in the extension of the noun. So, the representation of the sentence:

The water is wet.

would predicate wet’water’ of the extension of “the water” σxwater’x. This could be paraphrased as “the water is wet-as-water”. The modified property can distribute to any part of the extension of “the water”, provided that it is water. To infer, for example, that each of these parts are also wet, not just wet-as-water, the behaviour of the property modifier must be strengthened. Such possible strengthenings are already motivated by considerations of the semantics of NL adjectives and adverbs. In effect, the final theory extends ideas from the semantics of NL modifiers to address the problem of controlling distribution with mass terms.

As distributive inferences can now be made regardless of whether there is an homogeneous, or atomic ontology, the same theory can be used for plural terms. This has the advantage over the original theory of plurals (Chapter 5) in that it gives a treatment of the examples that Landman uses to motivate his intensional individuals.

The proposed representation for sentences, used in the examples of Chapter 8, uses terms that do not involve the logical constants. This is a further weakening of the theory which means that particular truth conditions for sentences are not forced upon us merely by virtue of the truth conditions of the logical constants.

**Axiomatics**

The weak typing of PT simplifies the task of producing a semantics for NL, as there is no strong hierarchy of types, which can require various type-lifting strategies over semantic terms, to allow, for example, one term to appear as both a property, and a property modifier.\(^1\) This is just one result that flows from the axiomatic approach to formal semantics adopted in the thesis.

Strong typing is often used in NL semantic theories to avoid the self-predication paradoxes, whereas PT avoids these paradoxes axiomatically, making the theory too weak to prove that paradoxes of self-predication are propositions.

---

\(^1\) Perhaps to be distinguished from type-lifting as used in categorial grammars to account for ellipsis [Dowty, 1988].
CHAPTER 9. CONCLUSIONS

The thesis uses weak axioms to address the category mistake apparent in predication of non-denoting terms. Incompleteness is extended to proofs of the existence of atomic mass terms, satisfying our intuitions that mass terms can be used regardless of one's theory of substances and matter. Further, this allows the effects of different ontological choices on the semantics of nominals to be explored within the one theory: the final theory neither forces nor prevents the formal equality of certain terms, like “the mud and the water” with “the muddy water”.

Implementation

PT is a first order theory, thus it has a semi-decidable proof-theory. This means that it lends itself to implementation. As the behaviour of NL representations are described directly as axioms in PT, rather than as restrictions on a model of the representation, a system that uses this theory for semantic representation can perform useful inferences. This is surely a desirable objective for formal semantics: not just to provide a symbolic representation of sentences, but to indicate how intuitively acceptable inferences can be performed, preferably in a computationally tractable framework.

The weak treatment of non-denoting terms is compatible with a proof-theoretic implementation of a NL system. As an example, the failure to ascribe a truth value to the sentence “the present king of France is bald” can be mechanically demonstrated to be due to the non-existence of “the present king of France”. Thus, in an implementation of a question answering system, the helpful answer:

There is no present king of France.

could be generated automatically, in response to the question:

Is the present king of France bald?

Clearly, such an implementation would be dependent upon a suitable formal theory of questions and answers.

Further Work

In the chapter on plurals in PT (Chapter 5), a sketch of a theory of presupposition with definite descriptors was presented. This could be given a full formalisation, and might prove to be extendible to cover other examples of presupposition.

The thesis does not address the contextual, dynamic effects that seem to affect intermediate distributive readings [Schwarzschild, 1990, Schwarzschild, 1992]. It was hinted that the strengthened PT that would be needed to account for presuppositions can be used to embody Martin-Löf's type theory
CHAPTER 9. CONCLUSIONS

[Martin-Löf, 1982, Martin-Löf, 1984] (Chapter 5). This can, in turn, be used to model some of the dynamic aspects of NL semantics [Ranta, 1991]. It then seems that there is scope for further work directed at treating dynamic effects in PT, using ideas from Martin–Löf’s type theory, which might be used to account for the contextual effects required to obtain a reductive analysis of intermediate distributive readings.

The final theory offers no treatment of committee-like objects, or proper names. Both of these additions should be fairly straightforward. In particular, it should be a simple matter to add Barker’s account of collective nouns [Barker, 1992].

In order to improve the maintainability of a program which makes use of this theory, work is required in developing a presentation of the semantic theory that lends itself to a more transparent application in grammars for natural language. As many terms belong to several types, their different behaviours when acting as a particular type must be made explicit. It would be interesting to see whether there are some generalisations that might be found. Insights from Martin-Löf’s theory of types might prove useful to this aim. Alternatively, just expressing some of the axioms and definitions directly in terms of the expressions suggested for the semantics of NL may simplify the theory.
Chapter 10

Afterthoughts

In this chapter I discuss some points and issues concerning measures and amounts; mention some complications for existence presuppositions; and speculate about a possible connection between property inheritance and Landman’s roles.

10.1 Measure and Amount

In the body of the thesis I have neglected to mention the notion of measure, as it occurs in such sentences as:

The sand weighs five tons.
The weight of the sand is five tons.

where the amount terms are *predicative* — as one of the arguments of the verb, or as the value of a function noun — and phrases like:

Five tons of sand.

where the amount term is *attributive* [Bunt, 1985]. Parsons calls these *isolated* amount terms, and *applied* amount terms, respectively [Parsons, 1970]. Cartwright has also given some consideration of measures and amounts [Cartwright, 1975]. Here I shall present Bunt’s account of measures and amount terms [Bunt, 1985], together with some criticisms.

Bunt seeks to address the problem of referring to the same measurement using different units. He notes that a function that takes two units of measure and gives the multiplication factor — a positive real number (\(\mathbb{R}^+\)) — to translate a measurement between them will be *transitive*: that is, if we translate a measurement from one unit to another, we should obtain the same result if we translate through some third unit of measure:

**Definition 10.1** A function \(f \in A \times A \rightarrow \mathbb{R}^+\) is transitive when, for any \(x, y, z \in A\):

\[f(x, y) \cdot f(y, z) = f(x, z)\]
Definition 10.2 \( T \) is the set of transitive functions from \( A \times A \rightarrow \mathbb{R}^* \).

He defines a dimension as follows:

**Definition 10.3** A dimension \( D = \langle SU_d, F_d \rangle \) where:

\[
SU_d = \{ x \mid x \text{ is a measure unit} \}
\]

\[
F_d \subseteq T r(SU_d)
\]

The elements of \( SU_d \) are D-units. \( F_d \) is the conversion function of \( D \).

An amount in dimension \( D \) is given by a pair \( \langle n, u \rangle \) where \( n \in \mathbb{R}^* \) and \( u \in SU_d \). As the same amount may be given by different measures, an equivalence between measures in dimension \( D \) is defined by: \( =^D \), must be given:

\[
\langle n_1, u_1 \rangle =^D \langle n_2, u_2 \rangle \quad \text{iff} \quad n_1 \cdot F_d(u_1, u_2) = n_2
\]

**Definition 10.4** A D-amount is an element of the partition of \( \mathbb{R}^* \times SU_d \) into equivalence classes defined by \( =^D \).

Elements of an amount represent that amount. Bunt also defines an ordering on amounts:

\[
\langle n_1, u_1 \rangle >^D \langle n_2, u_2 \rangle \quad \text{iff} \quad n_1 \cdot F_d(u_1, u_2) > n_2
\]

\[
a_d > a_d' \quad \text{iff} \quad \exists x \in a_d, x' \in a_d' \text{ s.t. } x >^D x'
\]

and the null amount for a dimension \( D \), \( \theta_d = \langle 0, u_i \rangle \) for an arbitrary D-unit \( u_i \).

Here is one of Bunt’s examples of a dimension:

\[
V = \langle SU_v, F_v \rangle \text{ where}
\]

\[
SU_v = \{ \text{litre, pint, gallon} \}
\]

\[
F_v = \{ \langle \text{litre, pint}, 1.76 \rangle, \langle \text{litre, gallon}, 0.22 \rangle, \langle \text{litre, litre}, 1 \rangle, \langle \text{pint, litre}, 0.57 \rangle, \langle \text{pint, gallon}, 0.13 \rangle, \langle \text{pint, pint}, 1 \rangle, \langle \text{gallon, litre}, 4.55 \rangle, \langle \text{gallon, pint}, 8 \rangle, \langle \text{gallon, gallon}, 1 \rangle \}
\]

He gives as an example of a V-amount as:

\[
\{ \langle 1.14, \text{litre} \rangle, \langle 2, \text{pint} \rangle, \langle 0.25, \text{gallon} \rangle \}
\]

Translating the intuitions of Bunt’s ideas from expressions in terms of ensemble theory to a Boolean algebra, a measure function \( \mu \) for dimension \( D \), for an algebra \( B \) will have as its domain the power set of parts of \( B \), \( \mathcal{P}(B) \), and will range over amounts of \( D \), \( A_d \). According to Bunt, \( \mu \) should have the following properties:
(i) \( \mu(\perp) = \emptyset_d \) and it is \( \emptyset_d \), and \( \mu(x) > \emptyset_d \) for \( x \neq \perp \).

(ii) When applied to the sum of non-overlapping terms, its value is the sum of the values when applied to the parts.

Bunt would then have it that:

\[ x \text{ weighs } n \text{ pounds.} \]

means:

\[ \mu_w(x) = (n, \text{pounds})_w \]

and:

\[ n \text{ pounds of } S. \]

denotes \( y \in S \) such that:

\[ \mu_w(y) = (n, \text{pounds})_w \]

I think there are several points that can be criticised in this account. First, it is not clear that “\( n \) pounds of \( S \)” denotes some \( S \). We might compare this with “the group of boys”. In the latter example, it would seem that the head of the noun phrase — at least semantically — may be either the group, or the boys that constitute the group. If we adopt a Platonic view of measures and amounts, as Bunt assents to\(^1\), why can these amounts not be the denotations of applied amount terms? After all, we may say:

Fifty pounds is too much to carry.

Bunt’s notion of a dimension is rather reductive. It takes a dimension to be defined by the existing units (and the conversion ratios between them). Should a new unit of measure be created for an existing dimension, then the dimension is changed. A dimension surely exists independently of the units used to measure it. For a naive account of measurements it might not be necessary to consider a real number representation: we might take denotables to just have the dimensions they have, and axiomatise the relationship between the dimensions of denotables and those of their parts and fusions.

We might also question what is being attributed a particular measure. In the relation:

\[ \mu(x) = a \]

\(^1\)Quoting from his book:

“From a semantic point of view, the treatment of amount terms which is presented here has a Platonic basis in that amount terms are considered as having denotations which are abstract mathematical objects, similar in many ways to numbers.” [Bunt, 1985, page 76].
what is $x$? For measures of volume and weight it could be anything with a physical extension (or the physical extension — of a substance or object — itself). If we measure such things as length or distance, then it is not so clear that $x$ is an object in the conventional sense. We might measure the distance between two objects. Is it then the case that the distance between two objects is itself an object? It is certainly a denotable, but only in terms of the objects that it is between. We can also measure to arbitrary (denotable) points:

The edge is 1 meter from the centre.

Perhaps in measuring length, $x$ might be the denotation of some vector, given in terms of two denotable points or planes.

Areas can also have a vector-like nature (given by the normal to the plane of measurement). This only seems to be of use when measuring the area of some abstract plane. Surface areas are not defined by a normal vector. Areas can also be summed without taking account of direction, unlike length.

Considering measures of length as vectors suggests that a notion of negative amounts might be useful. A more obvious case of negative amounts occurs with measures of temperature. Also, the conversion between different units of temperature is not given by a simple multiplication function. Converting from Kelvins to degrees Celsius requires the subtraction of 273.15. Temperature is also a problematic example as it is not clear how the measure of temperature of two objects can be reduced to simple addition. The only examples of measure that Bunt discusses are weight and volume. These are unproblematic as they are both additive, and require little consideration of the sorts of objects that they apply to. If we consider density, defined as mass divided by volume, we obtain another non-additive dimension: the density of the fusion of two things is not the sum of their densities, but is the sum of their masses divided by the sum of their volumes:

$$
\mu_{\text{density}}(x) = \frac{\mu_{\text{mass}}(x)}{\mu_{\text{vol}}(x)}
$$

$$
\mu_{\text{density}}(x_1 \oplus x_2) = \frac{(\mu_{\text{mass}}(x_1) + \mu_{\text{mass}}(x_2))}{(\mu_{\text{vol}}(x_1) + \mu_{\text{vol}}(x_2))}
\neq \mu_{\text{density}}(x_1) + \mu_{\text{density}}(x_2)
$$

where $x_1, x_2$ are not overlapping. What is missing is the notion of prime dimensions, which are additive, and derived dimensions which need not be. However, even if we find some set of prime dimensions — such as length, mass, and time — it is still not clear how to find the appropriate amounts for the dimensions of fused objects in some cases. Even if it is possible to reduce the dimension of temperature to the complex dimension\textsuperscript{2} $(\text{distance} \cdot \text{time}^{-1})^2$, it does not help in determining a reduction of the temperature of some fusion to the temperature of the parts. Elaborating on this, we may have a class

\textsuperscript{2}According to modern physics, temperature is a function of the mean speed of the molecular components of an object.
of prime additive dimensions \( p(1), \ldots, p(n) \) such that for non-overlapping
denotables \( x, y \):

\[
\mu_{p(i)}(x \oplus y) = \mu_{p(i)}(x) + \mu_{p(i)}(y)
\]

Derived dimensions \( d(1), \ldots, d(i), \ldots \) will be given by some function \( f_{d(i)} \) of
the prime dimensions:

\[
\mu_{d(i)}(x) =_{def} f_{d(i)}(\mu_{p(1)}(x), \ldots, \mu_{p(n)}(x))
\]

Then for a derived dimension, when applied to a fusion:

\[
\mu_{d(i)}(x \oplus y) = f_{d(i)}(\mu_{p(1)}(x \oplus y), \ldots, \mu_{p(n)}(x \oplus y))
\]

This clearly works for density, defined in terms of mass and volume. However
for temperature, and other dimensions such as velocity, it is not clear that this
is correct: what is the velocity of the fusion of two things? We must remember
that this is an abstract, not a physical fusion under consideration. Velocity
is a measure of something not possessed by an object in itself, but relative
to other objects. It is a measure of change (with respect to other objects)
in time. We should note that temperature, as a derived dimension, also
requires the prime dimension time. Considering time as a prime dimension:
although it can appear as a component in derived dimensions of denotable
objects, by itself it applies to events. Although additive when applied to
events, it is not additive considering just denotable objects, as it does not
even apply to such terms. Thus a caveat must be added to the above that
the reduction — of measures of fusions to measures of their parts — only
holds when considering dimensions derived from prime dimensions that can
apply directly to the type of object under consideration.\(^3\)

### 10.2 More on Existence Presuppositions

As mentioned in Chapter 5, the treatment of existence presuppositions with
definite descriptors may be too weak in general. Although it does seem
correct to say that the sentence:

The present King of France is bald.

has no truth value, there are examples of the use of such definite descriptors
where some authors feel a truth value can be assigned, namely:

The exhibition was visited by the present king of France.

\(^3\)It might be argued that the notion of weight involves the dimension of time — weight
is definable in terms of force which in turn depends upon the dimension of time — yet
weight is clearly additive. I think in this case it can be argued that the dimension of
time can be factored-out of weights: weights are measured in the same gravitational field.
What is referred to as weight is in reality dependent upon mass alone, which is additive.
which Strawson considers to be false [Strawson, 1964], and:

The king of France is sitting in that chair.

which Lasersohn takes to be false, when the chair is known to be empty [Lasersohn, 1993].

Strawson suggests splitting a sentence into a topic — what the sentence is about, akin to the focus of the sentence — and a comment — what is being attributed to the topic. According to Strawson, truth value gaps arise when a non-denoting term appears in the topic, but that no such truth value gap need arise if that term appears in the comment. This accounts for the passive example, where the topic is “the exhibition”, and the comment is “was visited by the present king of France” [Strawson, 1964].

With the property-theoretic treatment, this example causes no great problems, as there is nothing in the theory to prevent the representation of “visited by the present king of France” as a property of denotables. This can then form a proposition with a denoting noun phrase, and hence have a truth value.

Lasersohn believes that this treatment is less plausible for sentences of the second, non-passive form. He accounts for them using data set semantics, which act as a model of partial states of information [Landman, 1986a; Veltman, 1981]. If the current state of information is revised, non-monotonically, to verify the counter-factual “there is a king of France”, then it is consistent to extend the information state to include either that the king of France is bald, or that he is not bald. Akin to some models of free logic, when either extension is possible, a truth value gap occurs [van Fraassen, 1966]. With the sentence “the king of France is sitting in that chair”, when the chair is empty, there is no consistent way of extending the information state to include the proposition, so according to Lasersohn, the sentence must be false [Lasersohn, 1993].

This idea seems to be capturing the notion that if an agent is in a complete state of information with regard to the extension of a property of denotables, then when that property is attributed to anything not in its extension, it must lead to a false proposition, even when predicated of non-denoting terms. With the empty chair example, we are in a complete state of information with regard to ‘sitting in that chair’-hood.

As a first attempt, we could strengthen the theory with the axiom:

\((\text{Pty}_\Delta(p) \& \text{Pty}_\Delta(q) \& \sim \exists x T(qx)) \rightarrow \sim T(q(\text{the}'p))\)

The existing theory does not allow a proof of the propositionhood of properties of denotables predicated of non-denotables, but neither does it allow

\(^4\)However, when I presented these sentences to ‘non-academic’ native speakers, they took them to be as nonsensical as “the present king of France is bald”.}
a proof that they are not propositions, thus no contradiction results from adding this axiom.

Such an axiom accounts for the empty chair example. However, Lasersohn makes passing reference to another example:

The king of France is on the University of Rochester faculty.

which he takes to be false. He offers this as an example that indicates that the distinction between sentence with truth value gaps and those without does not lie in an individual/stage level distinction [Carlson, 1977]. Lasersohn offers no formal account of this example. It could be treated as a generalisation of the empty chair example: we know that even if the denotable individuals change, so that there is a king of France, it is unlikely that he is on the University of Rochester faculty; or, perhaps counter to Lasersohn’s views, as a generalisation of the topic/comment distinction, where the topic is “the University of Rochester faculty”.

An initially plausible generalisation of the above axiom that might seem to cope with this example is:

\[(\text{Pty}_\Delta(p) \& \text{Pty}_\Delta(q) \& \forall x (T(qx) \rightarrow \sim T(px))) \rightarrow \sim T(q(\text{the}'p))\]

More careful consideration reveals that this is too strong: any property that does not have a non-denoting nominal in its extension will form a false proposition with such terms. This appears to miss two points:

(i) There is a counter-factual element in these examples: even if there were a present king of France, he would not be in the faculty;
(ii) The example seems to rely on propositional attitudes: we know (or believe) that the existence of a present king of France is unlikely to have a bearing on the extension of the appropriate property.

We could take Lasersohn’s revision of data sets as a model of reasoning with knowledge (or belief), and counter-factuals. Without considering theories of knowledge, belief and counter-factuals, it is hard to make more progress with this example.\(^5\)

\(^5\)The following axiom fails to capture the appropriate intuitions:

\[(\text{Pty}_\Delta(p) \& \text{Pty}_\Delta(q) \& \exists x (T(px) \rightarrow \sim T(qx))) \rightarrow \sim T(q(\text{the}'p))\]

as the material implication is true when the inner antecedent is false, forcing “the king of France is bald” to be a false proposition. Relativising to an agent’s beliefs, a more intensional conditional might be captured by something of the form:

\[\text{bel}(\exists x (px \Rightarrow \sim qx)) \rightarrow \text{bel}(\sim q(\text{the}'p))\]

depending upon a suitable theory of belief. See [Davies, 1990] for a property-theoretic treatment of knowledge and belief.
10.3 Inheritance Hierarchies

To provide further justification for Landman’s individuals under roles [Landman, 1989] (or at least another application of them), I shall consider inheritance hierarchies, where properties are inherited via a variety of sortal hierarchy [Etherington and Reiter, 1983, Touretzky, 1984]. In such a logic, we would have a directed acyclic graph, consisting of a set of nodes, and a set of arcs between those nodes. Properties are inherited down the arcs, so that in the hierarchy of Figure 10.1 the individual nick will have the property of being a raven, and will in turn inherit all the properties that ravens have.

![Diagram](image)

Figure 10.1: A simple hierarchy.

Thus, nick will have the properties: raven, black, bird; and fly. We can formalise this notion of property inheritance in a proof-theory for directed acyclic graphs: if there is list of nodes:

\[ L = \langle x_1, \ldots, x_n \rangle \]

such that for each adjacent pair \( x_m, x_{m+1} \) in \( L \), there is a directed arc from \( x_m \) to \( x_{m+1} \), then we can write:

\[ L \vdash x_n(x_1) \]

This describes a variety of default reasoning.

There may be exceptions to the general rules captured by the graph. To indicate this, there are two kinds of directed arcs. There are those as above, that indicate that a property is inherited, and additionally, crossed arcs, to indicate exceptional cases. This is illustrated in Figure 10.2, where two exceptions are indicated: one, that penguins do not fly, although they are a kind of bird, and birds, in general, do fly; the second, that albino’s are not black.
CHAPTER 10. AFTERTHOUGHTS

Figure 10.2: Example of exceptions.

So, if a bird is a raven, it will be black, unless it is an albino. A crossed arc will block property inheritance, so if a crossed arc appears in a proof, it must be the last arc. We may have express this as a proof rule. If there is a list of nodes, connected by uncrossed arcs, except for the last arc:

$$L = \langle x_1, \ldots, \neg x_n \rangle$$

where $$\neg x_n$$ indicates the crossed arc from $$x_{n-1}$$, then we may write:

$$L \vdash \neg x_n(x_1)$$

Considering just the part of the Figure 10.2 given in Figure 10.3, it is apparent that with the two rules as stated, we may prove that tweety flies, and that tweety does not fly.

Figure 10.3: Extract from Figure 10.2.

We may avoid this problem by saying that the crossed arc between penguin and fly (and hence $$\neg$$fly) is more relevant than the uncrossed arc from bird.
to fly, as it applies at a level more specific to tweety. A constraint could be added to the proof rules so that the more specific proof succeeds. With the ‘diamond’ part of Figure 10.2, given in Figure 10.4, the inference:

\[
\text{nick, albino, } \neg \text{black} \vdash \neg \text{black}(\text{nick})
\]

should over-ride the inference:

\[
\text{nick, raven, black} \vdash \text{black}(\text{nick})
\]

to avoid the apparent inconsistency of supporting a proposition and its negation.

\[
\text{black} \\
\text{raven} \\
\text{albino} \\
\text{nick}
\]

Figure 10.4: Extract from Figure 10.2.

We might proceed by asserting that albino-hood is more specific than raven-hood, so inferences involving albino-hood should take precedence over those involving raven-hood. Alternatively, the first proof rule could be modified so that it does not apply if there is a proof to the contrary using the new proof rule. Caution should be taken in making this a general principle. Consider the so-called 

\text{Nixon diamond}, of the same form in Figure 10.5.

\[
\text{pacificist} \\
\text{quaker} \\
\text{republican} \\
\text{Nixon}
\]

Figure 10.5: The Nixon diamond.

It is not at all clear that there is a principled reason for the inference of \(\neg \text{pacificist}(\text{Nixon})\) to over-ride that of \(\text{pacificist}(\text{Nixon})\). I suggest that Landman’s roles may give an escape route out of this problem. When contradictory outcomes are possible and there appears no principled way of avoiding
them, then rather than allow inferences which attribute a property to the bare individual, the entailments could be modified so that property in question holds of an individual under a particular guise, or role. Thus the conclusions from the above ‘diamond’ might be:

\[
pacifist(Nixon \text{ as quaker})
\]
\[
\neg pacifist(Nixon \text{ as republican})
\]

This can be done by recording the point of divergence in the proof as the appropriate role. Sketching this idea, if we have the two sequences:

\[
L_1 = \langle i, x_1, \ldots, x_n, p \rangle
\]
\[
L_2 = \langle i, y_1, \ldots, y_m, \neg p \rangle
\]

then we take the first corresponding pair of elements \( x_r, y_r \) that differ — \( r \) is the lowest \( j \) such that \( x_j \neq y_j \) — as the appropriate roles, and allow the following proofs:

\[
L_1 \vdash p(i \text{ as } x_r)
\]
\[
L_2 \vdash \neg p(i \text{ as } y_r)
\]

This need not conflict with the notion of specificity mentioned above: for example, if the first difference between \( L_1, L_2 \) occurs only at the end of one of the lists, then the corresponding inference may be taken to be more specific, and hence over-rule the other putative inference (this does not cover the albino raven example).

Clearly, if a node inherits a property \( p_1 \) via a property \( p_2 \) that it has only under a particular role, then we may wish to say that it has the first property \( p_1 \) under the same role that it obtains the second \( p_2 \). So from Figure 10.6, we might like to prove:

\[
\text{anti-nuclear}(Nixon \text{ as quaker})
\]

\[
\text{anti-nuclear}
\]
\[
\text{pacifist}
\]
\[
\text{quaker}
\]
\[
\text{republican}
\]
\[
\text{Nixon}
\]

Figure 10.6: An extended Nixon diamond.
CHAPTER 10. AFTERTHOUGHTS

If this approach were adopted, then it might be possible for terms to be considered under iterated roles as in Figure 10.7, where there are two choice points.

\[ q \]
\[ r_3 \]
\[ r_4 \]
\[ p \]
\[ r_1 \]
\[ r_2 \]
\[ i \]

Figure 10.7: Double diamond.

The following might then be valid:

\[ \neg p(i \text{ as } r_2) \]
\[ q((i \text{ as } r_1) \text{ as } r_3) \]
\[ \neg q((i \text{ as } r_1) \text{ as } r_4) \]

This is in no way intended to be a thorough account of reasoning with property-inheritance graphs, but it serves to illustrate that the notion of individuals under roles may do service in other areas. It also indicates how a proof-theory concerning the construction of such terms might be developed. However, the notion of individuals under roles is more general than the above example suggests. We can capture the essence of:

- Judges are strict.
- John is liberal.
- John is a judge.

with the hierarchy in Figure 10.8.
but it is not clear how a directed acyclic graph could — or even whether it should — capture the meaning of the non-generic sentences:

The judge is strict.
John is liberal.
John is the judge.

which are of central concern for Landman’s theory of intensional individuals.
Appendix A

A Model of PT with Boolean Terms

We can show that PT+$\Delta$, PT+II are consistent if we can provide a (non-trivial) model which verifies all their axioms.

First of all we need a model for the Lambda Calculus. This can be used to build a model of PT. We then require the model of PT to be strengthened to satisfy the axioms for plural/denotable terms. Link and Lønning have both effectively shown that their axioms, with atomicity, are satisfiable if the denotable domain is a (definably) complete atomic Boolean algebra (where a definably complete atomic Boolean algebra is an atomic Boolean algebra) [Link, 1991a, Lønning, 1989]. As my axioms for plural terms and denotables (Chapter 5 and Chapter 6 respectively) are both essentially a weaker version of Link’s axioms in §2.2 then a Boolean algebra should verify them also.

We thus need a model of PT, where the natural language denotable/plural terms belong to a (definably) complete (atomic) Boolean algebra. The models we shall present for PT shall ‘naturally’ satisfy full completeness, as opposed to definable completeness, but, as we are only interested in showing that PT+$\Delta$, PT+II are consistent, then the model can be stronger than these theories: if we only need definable completeness for both theories, then it does not matter if we actually have full completeness in the model, nor, for PT+$\Delta$, does it matter that the Boolean algebra is atomic.

A Model of the $\lambda$-Calculus with Summed Terms

Following an existing approach [Scott, 1973], we shall build a model of the lambda calculus from domains consisting of complete lattices. In the limit we have a domain $D_\infty$ isomorphic to its own continuous function space, so $D_\infty \cong [D_\infty \rightarrow D_\infty]$. We can define mappings $\Phi : D \rightarrow [D \rightarrow D]$ and $\Psi : [D \rightarrow D] \rightarrow D$. 

149
**Definition A.1** A Scott Model is a triple $\mathcal{D} = \langle D, \Phi, \Psi \rangle$ with $D$ a domain and $\Phi, \Psi$ as above.

The terms of Lambda Calculus can be interpreted in such a structure relative to an assignment function $g$ which assigns elements of $D$ to variables, and interpretation function $i$ which assigns elements of $D$ to constants. The function $g[d/x]$ is the function $g$ except that $d$ is bound to $x$. Reference to $\mathcal{D}$ is dropped in the following, $i$ is assumed to be fixed.

$$\begin{align*}
\mathcal{I}[x]_g &= g(x) \\
\mathcal{I}[c]_g &= i(c) \\
\mathcal{I}[\lambda x t]_g &= \Psi(\lambda d, \mathcal{I}[t]_g[i/d]) \\
\mathcal{I}[t(t')]_g &= \Phi(\mathcal{I}[t]_g, \mathcal{I}[t']_g)
\end{align*}$$

We want to be able to give a model of the lambda calculus extended with sums and products of terms. We can do this with:

$$\begin{align*}
\mathcal{I}[t \oplus t']_g &= \bigsqcup \{\mathcal{I}[t]_g, \mathcal{I}[t']_g\} \\
\mathcal{I}[t \otimes t']_g &= \bigsqcap \{\mathcal{I}[t]_g, \mathcal{I}[t']_g\}
\end{align*}$$

**A Model of PT with Boolean Terms**

Following [Aczel, 1980]:

**Definition A.2** A model for PT shall be taken to be a Frege structure $\mathcal{M} = \langle \mathcal{D}, T, P \rangle$ where $\mathcal{D}$ is a model of the Lambda Calculus and

$$\begin{align*}
T : D &\longrightarrow \{0, 1\} \\
P : D &\longrightarrow \{0, 1\}
\end{align*}$$

Where $T$ and $P$ satisfy the structural requirements in [Aczel, 1980].

The characteristic functions $T$ and $P$ provide the extensions of the truth predicate, and the proposition predicate, respectively. The structural requirements they conform to verify the appropriate axioms of PT. For example, the function $T$ characterises a subset of $P$. Thus the terms have a subclass consisting of terms that correspond to propositions. This subclass, in turn, has a subclass of terms corresponding to the true propositions.

The language of wff can now be given truth conditions.

$$\begin{align*}
\mathcal{M} \models _3 s = t &\text{ iff } \mathcal{I}[t]_g = \mathcal{I}[s]_g \\
\mathcal{M} \models _3 T(t) &\text{ iff } T(\mathcal{I}[t]_g) = 1 \\
\mathcal{M} \models _3 P(t) &\text{ iff } P(\mathcal{I}[t]_g) = 1 \\
\mathcal{M} \models _3 \varphi \& \psi &\text{ iff } \mathcal{M} \models _3 \varphi \text{ and } \mathcal{M} \models _3 \psi \\
\mathcal{M} \models _3 \varphi \lor \psi &\text{ iff } \mathcal{M} \models _3 \varphi \text{ or } \mathcal{M} \models _3 \psi \\
\mathcal{M} \models _3 \varphi \rightarrow \psi &\text{ iff } \mathcal{M} \models _3 \varphi \text{ implies } \mathcal{M} \models _3 \psi \\
\mathcal{M} \models _3 \sim \varphi &\text{ iff } \mathcal{M} \models _3 \text{ not } \varphi \\
\mathcal{M} \models _3 \forall x \varphi &\text{ iff for all } d \in D \mathcal{M} \models _{g[d/x]} \varphi \\
\mathcal{M} \models _3 \forall^*_x \varphi &\text{ iff for some } d \in D \mathcal{M} \models _{g[d/x]} \varphi
\end{align*}$$
A wff $\varphi$ of PT is valid in a model $\mathcal{M}$ iff $\mathcal{M} \models_g \varphi$ for all assignment functions $g$.

For a model of PT+$\Delta$ (and PT+$\Pi$) we need a stronger base model than $\mathcal{M}$. We require those terms representing denotables to form a complete (atomic) Boolean algebra. Models of the Lambda Calculus, in general, do not possess these properties.

This problem can be addressed by giving a substructure of $\mathcal{M}$ the desired properties, and letting the natural language denotable terms (or plural terms) denote appropriate objects in this structure. Thus, denotable objects will form a sub-domain.

**A Complete Atomic Boolean Algebra**

A Complete Atomic Boolean Algebra $B$ is given by the following axioms:

1. $B$ contains at least 2 elements.
2. If $a, b \in B$ then $a' \in B$ and $a \sqcup b \in B$.
3. If $a, b \in B$ then $a \sqcup b = b \sqcup a$.
4. If $a, b, c \in B$ then $a \sqcup (b \sqcup c) = (a \sqcup b) \sqcup c$.
5. For all $a, b \in B$, if there is some $c \in B$ such that $a \sqcup \cdot = c \sqcup c'$
   then $a \sqcup b = a$.
6. For all $a, b, c \in B$ if $a \sqcup b = a$ then $a \sqcup b' = c \sqcup c'$.
7. For all $a \neq (c \sqcup c') \in B$ there exists $u \in B$, such that $u \sqcup a = q$
   and for all $i \in B$ such that $i \sqcup u = u$ either $i = (c \sqcup c')$ or $i = u$.
8. Any (non-empty) set $E \subseteq A$ has a least upper bound, $\sqcup E \in A$.

The axioms (i)-(vi) — adapted from [Hughes and Cresswell, 1973] — give a Boolean algebra. Axiom (vii) makes the algebra atomic. Axiom (viii) makes it complete. The notions *bottom* $0$: *top* $1$: $a \sqcap b$ and $a \setminus b$ can be defined:

\[
0 =_{df} a \sqcup a'
\]
\[
1 =_{df} 0'
\]
\[
a \sqcap b =_{df} (a' \sqcup b')'
\]
\[
a \setminus b =_{df} a \sqcap b'
\]

There is one well-known theorem that will be of use later:

**Lemma A.1** In an atomic Boolean algebra, every element is the supremum of the atoms it dominates (each element of $B$ is defined by the atoms it dominates).

**Proof:** From [Halmos, 1963]: each element $p \in B$ is an upper-bound of the set of atoms $E \subseteq B$ that it dominates. We must demonstrate that if $r$ is an arbitrary bound of $E$, then $p \subseteq r$. Assume otherwise, that $p \setminus r \sqsubseteq 0$. From
atomicity it follows there is a \( q \in B \) where \( q \subseteq p - r \). As \( p - r \subseteq p \), the atom \( q \in E \). But since \( (q \cap r) \subseteq ((p - r) \cap r) \) this contradicts that \( r \) is an upper-bound of \( E \).

We can draw two corollaries from this result:

**Corollary A.1** Different collections of atoms have different suprema.

**Corollary A.2** Different suprema dominate different atoms.

**Definition A.3** Let \( \mathcal{M}_\Delta = \langle D, T, P, B \rangle \) (\( \mathcal{M}_\Pi = \langle D, T, P, B \rangle \)) be a model of \( PT + \Delta \) (\( PT + \Pi \)). Where \( B : D \to \{0, 1\} \) characterizes those elements of \( D \) that are in a complete atomic Boolean sub-domain of \( D \), with the same ordering and join operator.

We can now express the conditions for \( \Delta t, t \leq t' \):

\[
\mathcal{M}_\Delta \models \Delta(t) \iff B[I[t]] = 1
\]
\[
\mathcal{M}_\Delta \models t \leq t' \iff I[t] \subseteq I[t']
\]

Thus, natural language denotable terms denote items in the complete atomic Boolean algebra. To give a model for the plural denotables, as opposed to the potentially atomless, mereological denotables, \( \Delta \) can be replaced by \( \Pi \) in the above.

As it does not matter what ‘non-denoting’ definite descriptors denote, the definite descriptor \( \sigma x \varphi \) can be given an interpretation:

\[
I[\sigma x \varphi] = \bigcup \{ \alpha | T[I[\varphi]_{\alpha \alpha/z}] = 1 \}
\]

It can be demonstrated that the model satisfies the axioms given for plural/denotable terms. I show that the axioms I have given for plural denotables in PT are satisfied by the model. As the axioms for a potentially atomless mereology in PT are weaker, they too will be satisfied by the model.

**Theorem A.1** The summation operator \( \oplus \) is symmetric, idempotent and associative (Axioms 5.1; 5.2; 5.3):

**Proof:** Trivial: the summation operator is modelled by the lattice theoretic join \( \cup \) which is also symmetric, idempotent and associative. \( \square \)

**Theorem A.2** The domain of plurals is closed (Axiom 5.5).

\[
\Pi a \& \Pi b \to \Pi(a \oplus b)
\]

**Proof:** From the axioms of the Boolean algebra: if \( a, b \in B \), then \( a \cup b \in B \). \( \square \)

**Theorem A.3** The domain of plurals is (definably) complete (Axiom 5.6).

\[
\forall p((\text{Pty}(p) \& \forall x(T(px) \to \Pi(x)) \& \exists x(T(px))) \to \Pi(\sigma x px))
\]
**Proof:** From the completeness of the Boolean algebra: $X \subseteq B$, then $\bigsqcup X \in B$. The axiom is weaker, as it requires that there is a plural denotable in the extension of the property. □

**Theorem A.4** Different plural denotables have different atoms (Axiom 5.7).
\[ \forall \Pi xy(x \leq y \rightarrow \exists \Pi u (Ju \& u \leq x \& u \leq y)) \]

**Proof:** From Corollary A.2. □

**Theorem A.5** Plurals within the extension of a property (of plurals) must have an upper-bound (Axiom 5.8).
\[ \forall \Pi \forall y ((\text{Pty}_{\Pi}(p) \& T(py)) \rightarrow y \leq \sigma_{\tau} x px) \]

**Proof:** The axioms of the Boolean algebra define $\bigsqcup X$ as an upper-bound on the members of the set $X \subseteq B$. The supremum operator $\sigma$ is modelled by $\bigsqcup$. Further, all PT definable properties of plurals will have an extension within $B$. □

**Theorem A.6** The supremum of the extension of a property of plurals is the smallest plural, which dominates all the terms in the extension (Axiom 5.9).
\[ \forall \Pi \forall y ((\text{Pty}_{\Pi}(p) \& \forall \Pi x (T(px) \rightarrow x \leq y)) \rightarrow \sigma_{\tau} x px \leq y) \]

**Proof:** From the axioms of the Boolean algebra: $\bigsqcup X$ is the least upper-bound of $X$, where $X \subseteq B$. □

**Theorem A.7** Different atoms in the plural domain have different suprema (Axiom 5.10).
\[ \forall \Pi \forall u ((\text{Pty}_{\Pi}(p) \& Ju \& u \leq \sigma_{\tau} x px) \rightarrow \exists \Pi z (T(pz) \& u \leq z)) \]

**Proof:** From Corollary A.1. □
Appendix B

A Grammar

The following grammar generates some sentences with intransitive verbs. The semantic expressions that are created are intended to be given a stage level interpretation. This may, depending upon one’s intuitions, result in some liberties being taken in forming the semantic representation. It should be noted that this is very much a toy grammar, that leaves much to be desired. The aim is not to present a linguistically cogent analysis of syntax, rather, it is to indicate that a weakly typed theory, embellished with a part-whole structure and property modifiers, can in principle be used for the semantics of natural language.

The grammar will give a syntactic treatment of distribution across conjoined noun phrases, by means of a syntactic analysis of ellipsis, whereby “the old man and woman died”, for example, is to be read as “the old man died, and the old woman died”. This might seem counter to the arguments given in §2.1, that the distributive/collective distinction is not a scoping effect. However, the grammar only treats distribution (of intransitive verbs) across syntactic conjunction in this manner: it is not a general treatment of distribution into lexical plurals, and mass terms.

A further objection might be that this is syntactic analysis of distribution, whereas the body of the thesis is concerned with a semantic treatment of distributive inferences. It is possible to obtain these inferences exclusively in the semantics, thereby what is a potential syntactic ambiguity here, becomes a semantic ambiguity, or indeterminacy. Indeed, the arguments presented by Schwarzschild would seem to suggest that a semantic treatment (as opposed to a syntactic treatment) of the distributive/collective ambiguity with conjoined noun phrases is desirable, as NL syntax may be misleading as to the appropriate inferences [Schwarzschild, 1990, Schwarzschild, 1992]. However, it should be noted that the main target of Schwarzschild’s criticisms is compositionally derived structured representations for reciprocal readings. It is my intention that the truth-conditions of the non-distributive collective readings here should subsume those of the intermediate distributive readings and reciprocals.

154
The aim of capturing ellipsis is not to make explicit particular intermediate distributive interpretations of collective predication, but to provide essential coverage of the ellipsis of determiners, adjectives and noun phrases as will be exemplified below. Although it would be possible to prevent the verb phrase abstraction from conjoined noun phrases (in order to avoid a syntactic treatment of distributive inferences) such ellipsis will be retained. If the grammar ensures that properties with a distributive interpretation are forced to distribute across conjunction, then the axioms concerned with distributive inferences in the semantic theory do not have to contend with a combination of lexical plurals and conjunction. In particular, distribution for noun phrases exclusively involving proper names is dealt with entirely within the grammar, although as mentioned in Chapter 7, proper names do present problems for determining the appropriate property modifiers in the truth conditions.

There are two steps in determining the truth conditions of a given sentence. The role of the grammar, as presented here, is to generate an explicit representation of the predicate-argument structure of an English sentence as a PT term $s$. Assuming that the semantic representation forms a proposition, the truth conditions of a sentence are obtained in terms of the truth conditions of $T(s)$. It is possible in many cases, to give the semantics of a sentence as a term involving logical constants (as in [Turner, 1992]). The truth conditions of $T(s)$ are then simply the truth conditions of the logical constants given in Chapter 4. This has the potential to be too reductive: it need not be the case that we would want the truth conditions obtained in this manner. As an example, expressing the representation of NL quantifiers as the intension of logical quantifiers may force a particular reading of a sentence, whereas quantifier scoping ambiguity [Cooper, 1983] suggests that there is often more than one way of satisfying the truth conditions of sentences involving more than one quantifier. This might, of course, be given a syntactic treatment. My suggestion is to represent NL sentences in terms of non-logical constants. Rather than using the axioms of truth of basic PT to obtain the truth conditions of a sentence, extra axioms — concerning the truth conditions of various NL-derived semantic structures — can be added as in Chapter 8, giving the possibility of ensuring that the system is never too strong. This approach has the additional advantage of avoiding an overdose of lambda expressions that would otherwise be required to accommodate the property modifiers used in the truth conditions. It also allows the grammar to be integrated more readily with different semantic theories by just altering the axioms that give the truth conditions of the relevant structures, without a major reworking of the semantic expressions in the grammar, or the truth conditions of basic PT.

The grammar is presented in the form of context-free rules with attribute-value features. Only those attributes concerning the semantics and ellipsis are made explicit. I assume that in an effective implementation the grammar
rules would have additional attribute-value feature structures that impose constraints on number, and count agreement, for example. These constraints could involve the properties of the semantics of the categories, so that it is not necessary to state separately in both the syntax and semantics that, for example, a word is plural.

Basic Grammar Rules

The basic grammar rules are in Figure B.1. They are of the form:

\[ C \rightarrow C_1 \ C_2 \ldots \ C_n \]

where \( C, C_1, \ldots, C_n \) are syntactic categories with attribute-value features, or lexical items. The relevant attribute-values pairs of categories are made explicit, in the form:

\[
\langle \text{category}\rangle \left[ \begin{array}{c}
\{\text{attribute}_1\}:\{\text{value}_1\} \\
\{\text{attribute}_2\}:\{\text{value}_2\} \\
\vdots \\
\vdots \\
\end{array} \right]
\]

As mentioned, the attributes that will be made explicit are those for the semantics \( sem \), and later, those for ellipsis \( elip \).

Variables with the same name are to be equated, or unified, when occurring as attribute values within a rule. This can have the effect of constraining the application of a rule. Rather than give the features of each category, these constraints could be expressed as equations associated with the rule. However, later there are rules which have categories of the same name appearing on the left and right sides: giving the attribute-values with each category avoids a source of ambiguity.

This grammar has many failings. As an example, the analysis of “is” constructions is perhaps naive, and it results in some complications of the semantic typing of lexical items, but it is adequate for the current purpose.

Grammar Rules for Ellipsis

I will present the grammar rules that involve elliptic categories as metagrammar rule schemata. This avoids the need to present all the possible rules involving ellipsis explicitly. As the semantics of the grammar are very close to the structure of the sentence, the reader may notice how semantic manipulations might be used to achieve the same effect as this syntactic treatment of ellipsis.

For a syntactic account of ellipsis, an extra attribute-value feature:

\[ elip : L \]

can be added to the categories. The value \( L \) will be a list of elliptic categories \([C_1, \ldots, C_n]\) that have been abstracted, or an empty list \( \emptyset \) if there
\[
\begin{align*}
\text{DET} \[ \text{sem:the}\] & \rightarrow \text{the} \\
\text{DET} \[ \text{sem:all}\] & \rightarrow \text{all} \\
\text{DET} \[ \text{sem:some}\] & \rightarrow \text{some} \\
\text{DET} \[ \text{sem:a}\] & \rightarrow \text{a} \\
\text{CONJ} \[ \text{sem:}\lambda q(p \land q)\] & \rightarrow \text{and} \\
\text{CONJ} \[ \text{sem:}\lambda q(p \lor q)\] & \rightarrow \text{or} \\
N \[ \text{sem:}\{\text{nouns}\}\] & \rightarrow \langle \text{noun} \rangle \\
N \[ \text{sem:}\{\alpha\}\beta\] & \rightarrow \text{ADJ} \[ \text{sem:}\alpha\] N \[ \text{sem:}\beta\] \\
\text{ADJ} \[ \text{sem:}\{\text{adjectives}\}\] & \rightarrow \langle \text{adjective} \rangle \\
NP \[ \text{sem:}\{\alpha\}\beta\] & \rightarrow \text{DET} \[ \text{sem:}\alpha\] N \[ \text{sem:}\beta\] \\
NP \[ \text{sem:}\{\text{proper nouns}\}\] & \rightarrow \langle \text{proper noun} \rangle \\
VP \[ \text{sem:}\{\text{verbs}\}\alpha\beta\] & \rightarrow \langle \text{verb}\rangle \alpha\beta\alpha\beta \\
VP \[ \text{sem:}\alpha\] & \rightarrow \text{is} A \text{DJ} \[ \text{sem:}\alpha\] \\
VP \[ \text{sem:}\alpha\] & \rightarrow \text{is} N \[ \text{sem:}\alpha\] \\
S \[ \text{sem:}\{\text{sentences}\}\alpha\beta\] & \rightarrow \text{NP} \[ \text{sem:}\alpha\] VP \[ \text{sem:}\beta\] \\
C \[ \text{sem:}\gamma\alpha\beta\] & \rightarrow C \[ \text{sem:}\alpha\] CONJ \[ \text{sem:}\gamma\] C \[ \text{sem:}\beta\] \\
\end{align*}
\]
Where \( C \) is any category.

Figure B.1: The Basic Grammar
are no elliptic categories. As all the previous rules do not involve ellipsis, the attribute-value \( \text{clip} : \emptyset \) can be added to their features. At times it is convenient to refer to the head \( H \) and tail \( \varphi \) of a list \([H | \varphi]\), where \( \varphi \) is a list.

If in the grammar we have a rule:

\[
X \left[ \text{clip} : \emptyset \right]_{\text{sem} : S} \rightarrow A \left[ \text{clip} : \emptyset \right]_{\text{sem} : \alpha} B \left[ \text{clip} : \emptyset \right]_{\text{sem} : \beta}
\]

then the following rules can be added to the grammar:

\[
X \left[ \text{clip} : [A] \right]_{\text{sem} : \lambda x(S|\alpha := x)} \rightarrow B \left[ \text{clip} : \emptyset \right]_{\text{sem} : \beta}
\]

\[
X \left[ \text{clip} : [B] \right]_{\text{sem} : \lambda x(S|\beta := x)} \rightarrow A \left[ \text{clip} : \emptyset \right]_{\text{sem} : \alpha\cdot\alpha'}
\]

To allow multiple iteration of ellipsis, if we have the rule:

\[
X \left[ \text{clip} : \varphi \right]_{\text{sem} : S} \rightarrow A \left[ \text{clip} : \emptyset \right]_{\text{sem} : \alpha'\cdot\alpha''}
\]

and there is a rule that allows the formation of \( A[\text{clip} : Y]\):

\[
A \left[ \text{clip} : [Y] \right]_{\text{sem} : \alpha'\cdot\alpha''} \rightarrow P
\]

we can add the rule:

\[
X \left[ \text{clip} : [Y | \varphi] \right]_{\text{sem} : \lambda x(S|\alpha' := \alpha'\cdot\alpha')} \rightarrow P
\]

Thus the list of ellipsis categories that is the value of the feature \( \text{clip} \), can grow, like a stack, with the most recently abstracted category at the head of the list.

To resolve ellipsis, the following rules can be added:

\[
X \left[ \text{clip} : \varphi \right]_{\text{sem} : \beta\cdot\alpha} \rightarrow A \left[ \text{clip} : \emptyset \right]_{\text{sem} : \alpha} X \left[ \text{clip} : [A | \varphi] \right]_{\text{sem} : \beta}
\]

\[
X \left[ \text{clip} : \varphi \right]_{\text{sem} : \alpha\cdot\beta} \rightarrow X \left[ \text{clip} : [B | \varphi] \right]_{\text{sem} : \beta} B \left[ \text{clip} : \emptyset \right]_{\text{sem} : \alpha\cdot\beta}
\]

This is assuming that the grammar contains the rule \( X \rightarrow A \cdot B \) as before.

The rules in the grammar so far support the creation of the following basic rules involving ellipsis:

\[
S \left[ \text{clip} : [VP] \right]_{\text{sem} : \lambda x(x(x(\alpha)))} \rightarrow NP \left[ \text{clip} : \emptyset \right]_{\text{sem} : \alpha}
\]

\[
NP \left[ \text{clip} : [DET] \right]_{\text{sem} : \lambda x(x(x(\alpha)))} \rightarrow N \left[ \text{clip} : \emptyset \right]_{\text{sem} : \alpha\cdot\alpha'}
\]

\[
N \left[ \text{clip} : [ADJ] \right]_{\text{sem} : \lambda x(x(x(\alpha)))} \rightarrow N \left[ \text{clip} : \emptyset \right]_{\text{sem} : \alpha\cdot\alpha''}
\]
APPENDIX B. A GRAMMAR

Iteration of ellipsis, according to the schema above, gives:
\[
NP \left[ \textit{clip}: [\textit{ADJ}, \textit{DET}] \quad \text{sem m: } \lambda y \lambda x (\varepsilon (y(x))) \right] \quad \rightarrow \quad N \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \gamma \right]
\]
\[
S \left[ \textit{clip}: [\textit{DET}, \textit{VP}] \quad \text{sem m: } \lambda y \lambda x (\varepsilon (y(x))) \right] \quad \rightarrow \quad N \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \gamma \right]
\]
\[
S \left[ \textit{clip}: [\textit{ADJ}, \textit{DET}, \textit{VP}] \quad \text{sem m: } \lambda y \lambda x (\varepsilon (y(x))) \right] \quad \rightarrow \quad N \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \gamma \right]
\]

The meta-rule for reducing ellipsis produces the following new rules:
\[
N \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \gamma \right] \quad \rightarrow \quad ADJ \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \beta \right] \quad N \left[ \textit{clip} :: \textit{ADJ} \right]
\]
\[
NP \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \gamma \right] \quad \rightarrow \quad NP \left[ \textit{clip} :: N \quad \text{sem m: } \beta \right] \quad N \left[ \textit{clip} :: \emptyset \right]
\]
\[
NP \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \gamma \right] \quad \rightarrow \quad DET \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \beta \right] \quad NP \left[ \textit{clip} :: \textit{DET} \right]
\]
\[
S \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \gamma \right] \quad \rightarrow \quad NP \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \beta \right] \quad S \left[ \textit{clip} :: \textit{NP} \right]
\]
\[
S \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \gamma \right] \quad \rightarrow \quad S \left[ \textit{clip} :: \textit{VP} \quad \text{sem m: } \beta \right] \quad VP \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \beta \right]
\]
\[
S \left[ \textit{clip} :: \textit{VP} \quad \text{sem m: } \gamma \right] \quad \rightarrow \quad S \left[ \textit{clip} :: \textit{N}, \textit{VP} \quad \text{sem m: } \beta \right] \quad N \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \beta \right]
\]
\[
S \left[ \textit{clip} :: \textit{VP} \quad \text{sem m: } \gamma \right] \quad \rightarrow \quad DET \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \beta \right] \quad S \left[ \textit{clip} :: \textit{DET}, \textit{VP} \right]
\]

Some of these rules may appear somewhat obscure outside their context of use. It is possible to derive other elliptic rules, but not all of them are relevant for the fragment of English.

A major reason for which I require a treatment of ellipsis is to allow distribution of determiners, adjectives and verbs across conjoined categories. The syntactic analysis of ellipsis across conjunction, presented here, has the unfortunate consequence of requiring different rules for conjunction, depending upon how many syntactic objects have been abstracted from conjoined categories:
\[
A \left[ \textit{clip} :: \emptyset \quad \text{sem m: } \gamma \right] \quad \rightarrow \quad A \left[ \textit{clip} :: B \quad \text{sem m: } \alpha \right] \quad CONJ \left[ \text{sem m: } \gamma \right] \quad A \left[ \textit{clip} :: B \quad \text{sem m: } \beta \right]
\]
\[
A \left[ \textit{clip} :: C, B \quad \text{sem m: } \gamma \right] \quad \rightarrow \quad A \left[ \textit{clip} :: [C, B] \quad \text{sem m: } \alpha \right] \quad CONJ \left[ \text{sem m: } \gamma \right] \quad A \left[ \textit{clip} :: [C, B] \quad \text{sem m: } \beta \right]
\]
\[
A \left[ \textit{clip} :: [D, C, B] \quad \text{sem m: } \gamma \right] \quad \rightarrow \quad A \left[ \textit{clip} :: [D, C, B] \quad \text{sem m: } \alpha \right] \quad CONJ \left[ \text{sem m: } \gamma \right] \quad A \left[ \textit{clip} :: [D, C, B] \quad \text{sem m: } \beta \right]
\]

;
This can be generalised: if we have the following notation for arbitrary iterations of \( \lambda \)-abstraction and application:
\[
(\lambda x)^0 y =_{def} y \\
(\lambda x)^n y =_{def} \lambda x_n.(\lambda x)^{n-1} y \\
t x^0 =_{def} t \\
t x^n =_{def} (t x_n) x^{n-1}
\]
and arbitrary lists of syntactic abstractions:
\[
E^0 =_{def} \emptyset \\
E^n =_{def} [E_n | E^{n-1}]
\]
where \( E_1, \ldots, E_n \), are syntactic categories, then we can give the following rule schema for conjunction:
\[
C[\begin{array}{c}
\epsilon \bar{y} p \\
{sem:} \gamma \end{array} E^n] \longrightarrow C[\begin{array}{c}
\epsilon \bar{y} p : E^n \\
{sem:} \alpha \\
\end{array} CONJ[\begin{array}{c}
\gamma \\
{sem:} \beta \\
\end{array}]]
\]
For the sentences this grammar is intended to cover, we can add the restriction that ellipsis can only arise within the context of a conjunction: that is, in the fragment of English under investigation, categories are only abstracted from conjoined phrases.

Categories with apparently similar syntactic behaviour, like \( VP[\epsilon \bar{y} p : \emptyset] \) and \( S[\epsilon \bar{y} p : NP] \), have dissimilar semantic behaviour when part of a conjunction: elliptic categories will have the semantics of the appropriate constituent mapped to the conjuncts, unlike the syntactically ‘equivalent’ non-elliptic categories. The conjunction rule has the effect of allowing arbitrary abstraction from conjoined phrases. When considering the abstraction of verb phrases from conjoined noun phrases, rules of this form have the effect of modelling a distributive inference over NL conjunction. The sentence:

The men and the women met.

can be parsed as either:
\[
met'(the'men' + the'women')
\]
or as an intermediate distributive reading:
\[
met'(the'men') + met'(the'women')
\]

If we assume no further restrictions on the attribute-value feature structure, then the possible semantics of “the old man and woman died” would be:
\[
died'(the'old'(man' + woman')) \\
died'(the'old'man' + the'woman') \\
died'(the'old'man' + the'old'woman') \\
died'(the'old'man') + died'(the'woman') \\
died'(the'old'man') + died'(the'old'woman')
\]
Intuitively, the first interpretation refers to something that is both man and woman. This might be acceptable with other, non-exclusive terms, such as drunk and addict. The second and third interpretation allow “the” to distribute over the conjunction, the third interpretation has also distributed the adjective over the conjunction. As the verb “to die” is conventionally distributive, only the last two syntactic readings seem desirable. If appropriate features were added to the grammar, the first three readings could be blocked, using the knowledge that died’ is always distributive. We could make these attribute-value restrictions dependent upon the semantics, which is, in any case, an attribute value. This might be objected to, as it gives rise to a syntactic classification of words on the grounds of their semantic entailments. It must be emphasised that this grammar is not offered as a theoretically principled solution.

If the grammar allowed the second and third readings, where died’ does not distribute, then it might be necessary to strengthen the semantic theory to force an equivalence with the last two parses, which would complicate the final semantic theory, in Chapter 8. As it stands, the final theory only concerns itself with explaining distribution into lexical plurals, as in “the old men died”.

With the sentence “the old man and addict met”, then the desirable results of parsing are:

\[\text{met}'(\text{the}'\text{old}'\text{man}' \oplus \text{the}'\text{addict}')\]
\[\text{met}'(\text{the}'\text{old}'\text{man}' \oplus \text{the}'\text{old}'\text{addict}')\]

The other readings could be ruled out by rules on the features that insist that relational predicates, such as meet', can only be applied to collections.

With conjoined noun phrases, the grammar obtains either an interpretation where the verb phrase applied to each noun phrase, or where the verb phrase applies to the sum of all the noun phrases. This is compatible with the view taken in the thesis that sentences are ambiguous between the fully distributive and the collective interpretations.

When a modifier is abstracted from conjoined categories, as in the noun:

\[\text{green man and woman}\]

it seems that in the semantics, the adjective should modify only atomic nouns, not conjoined nouns. Ambiguity lies in which nouns are modified. The rules given above do not prevent the interpretation:

\[\text{green}'(\text{man}' \oplus \text{woman}')\]

This can be avoided by adding an extra, Boolean feature to the categories, so that when categories are conjoined, a feature \textit{conj} is set to \textit{yes}. When we have a rule of the form \(X \rightarrow AX\), or \(X \rightarrow XA\), then the creation of ellipsis involving category \(A\) can be restricted to the case where \(X\) is not
conjoined (although within the context of a conjunction). For example, if we have the rule:

$$X[\text{clip}:\emptyset] \rightarrow A[\text{clip}:\emptyset] X[\text{clip}:\emptyset]$$

then the appropriate rule with ellipsis is modified to:

$$X[\text{clip} : \text{conj} : |A|_{\text{so}}]_{\text{sem} : \lambda x (S[x := x])} \rightarrow X[\text{clip} : \text{conj} : \text{no}]_{\text{sem} : \beta}$$

Using adjectives and nouns as an example, this has the result that:

- green man and woman and child

is interpreted as one of:

- green’man’ $\oplus$ green’woman’ $\oplus$ green’child’
- green’man’ $\oplus$ green’woman’ $\oplus$ child’
- green’man’ $\oplus$ woman’ $\oplus$ child’

To show that the representations for coherent sentences generated by the grammar are propositions requires extra typing rules (and lemmata) to be added to the semantic theory. The truth conditions of these sentences can be expressed in terms of more conventional logical connectives and quantifiers by adding further axioms. Chapter 8 discusses these issues.

Although not essential for the examples in the thesis, I shall elaborate on some extra rules that generalise the notion of ellipsis, and allow mutual abstraction, apparent in sentences like:

- The men and the women laughed and died.
- Each and every man and woman laughed.

To perform these more complex syntactic distributions, we need to allow categories with ellipsis to combine in a more complex fashion. This can occur in at least two ways. Considering rules with two constituents, of the form:

$$X[\text{clip}:\emptyset] \rightarrow A[\text{clip}:\emptyset] B[\text{clip}:\emptyset]$$

we must allow the categories $X[\text{clip}:A]$ and $X[\text{clip}:B]$ to combine by adding a rule like:

$$X[\text{clip}:\emptyset_{\text{sem} : S[\beta := \beta']_{\text{so}} := \alpha'}] \rightarrow X[\text{clip} : B] X[\text{clip} : A]$$

Presumably the application of this rule must be effectively restricted so that the categories with ellipsis were derived from the appropriate components, $A, B$ respectively. The semantics has the effect of swapping the application of the representations.

This meta-rule allows us to derive the rule:

$$S[\text{clip} : \emptyset_{\text{sem} : (\alpha)}] \rightarrow S[\text{clip} : \text{VP}] S[\text{clip} : \text{NP}]$$

which extends the grammar to cope with the example:
The men and the women laughed and died.

when the desired representation is:

\[
\text{laughed'}(\text{the'men'}) \oplus \text{died'}(\text{the'men'}) \oplus \\
\text{laughed'}(\text{the'women'}) \oplus \text{died'}(\text{the'women'})
\]

A second kind of mutual abstraction can be defined where the mutually abstracting categories themselves constitute a category with ellipsis. The third, external, abstracted category must combine with one of the constituents. If we have the rules:

\[
X[\text{clip}] \rightarrow A[\text{clip}] \ B[\text{e}]
\]
\[
A[\text{clip}] \rightarrow P[\text{clip}] \ Q[\text{clip}]
\]

then by this process, the following additional rules would be sanctioned:

\[
X[\text{clip};B] \rightarrow X[\text{clip};Q, B] \ A[\text{clip};P]
\]
\[
X[\text{clip};B] \rightarrow A[\text{clip};Q] \ X[\text{clip};P, B]
\]

Similarly with:

\[
X[\text{clip}] \rightarrow B[\text{clip}] \ A[\text{e}]
\]

We can consider the following correspondence in syntactic behaviour:

\[
X[\text{e}p;B] \cong A[\text{e}p]
\]
\[
X[\text{e}p;Q, B] \cong P[\text{e}]
\]
\[
A[\text{e}p;P] \cong Q[\text{e}]
\]

\[
X[\text{e}p;Q] \cong P[\text{e}]
\]
\[
X[\text{e}p;P, B] \cong Q[\text{e}]
\]

So both the new rules have the same syntactic behaviour as \(A \rightarrow P \ Q\). The difference will lie in the semantics. With conjoined categories, the semantic expressions should ‘redirect’ the semantics of the abstracted categories to the conjuncts. As with simple mutual abstraction, the order of application in the semantics must be reversed. However, if the argument in the resultant semantics contains the external abstraction, then the corresponding \(\lambda\)-abstraction in the semantics must have its scope widened. Following one possibility, if the semantics of the original rule \(X \rightarrow A \ B\) are as follows:

\[
X[\text{clip}; \emptyset, \text{sem}; (\alpha)\beta] \rightarrow A[\text{clip}; \text{sem}; (\alpha)] \ B[\text{clip}; \text{sem}; \beta]
\]

then the semantics of the new rules will be given by:

\[
X[\text{e}p; B] \rightarrow X[\text{clip};Q, B] \ A[\text{clip};P] \\
X[\text{e}p; B] \rightarrow A[\text{clip};Q] \ X[\text{clip};P, B]
\]
This allows the following rules to be added to the grammar:

\[
S \left[ \text{\textit{clip}}: \begin{array}{c}
\text{\textit{VP}} \\
\text{\textit{sem}}: \lambda x (\beta (\lambda y (\alpha y)))
\end{array} \right] \rightarrow S \left[ \text{\textit{clip}}: \begin{array}{c}
\text{\textit{NP}} \\
\text{\textit{sem}}: \alpha
\end{array} \right] \hspace{1em} NP \left[ \text{\textit{clip}}: \begin{array}{c}
\text{\textit{DET}} \\
\text{\textit{sem}}: \beta
\end{array} \right]
\]

\[
S \left[ \text{\textit{clip}}: \begin{array}{c}
\text{\textit{VP}} \\
\text{\textit{sem}}: \beta \alpha
\end{array} \right] \rightarrow NP \left[ \text{\textit{clip}}: \begin{array}{c}
\text{\textit{NP}} \\
\text{\textit{sem}}: \alpha
\end{array} \right] \hspace{1em} S \left[ \text{\textit{clip}}: \begin{array}{c}
\text{\textit{NP}} \\
\text{\textit{sem}}: \beta
\end{array} \right]
\]

which then allow the sentence:

Each and every man and woman laughed.

to be parsed, obtaining the semantic expression:

\[
\begin{align*}
& \text{laughed'each'man'} \oplus \text{laughed'each'woman'} \\
& \text{laughed' every'man'} \oplus \text{laughed' every'woman'}
\end{align*}
\]

### Categorial Grammar

Categorial Grammars\(^1\) may provide an alternative means of formalising the elliptic abstractions in the examples. In these grammars, the categories are functions that indicate the syntactic arguments they can combine with. For linguistic applications, the notation can indicate on which side the argument appears. For example a determiner can be represented by a function that produces a noun phrase, given a noun appearing on the right:

\[
NP / N + N \Rightarrow NP
\]

and, using Dowty’s ‘combinatory’ notation, a verb phrase can be considered as a function that produces a sentence, given a noun phrase on its left:

\[
NP + S \setminus NP \Rightarrow NP
\]

These examples illustrate the two rules for function application.

This is not the only manner in which categories can combine, for example, we can also define function composition:

\[
A / B + B / C \Rightarrow A / C
\]

\[
B \setminus C + A \setminus B \Rightarrow A \setminus C
\]

If we take \(C\) in \(B / C, B \setminus C\) to be an elliptic category, then these rules might indicate how an abstracted category can percolate through other categories. Thus categorial grammars seem already to contain the ingredients for exploring elliptic abstraction. Further, categorial grammars have already been used to represent the polymorphic behaviour of conjunction, and the ambiguity as to whether arguments (or functions) distribute to the syntactic conjuncts or not [Dowty, 1988, Geach, 1972]. In the meta-rules I have given, the apparent

\(^1\)For an overview, see [Steedman, 1993].
elliptic abstraction is realised in the semantics by adding λ-abstracts to the conjuncts, and making the conjunction map external argument to each of these abstracts. In categorial grammar, this process of abstraction can be implemented by means of type-raising, the effect of which is to abstract the distributing category from the conjuncts. The relevant type-raising rules are:

\[
A/B + B \Rightarrow A \\
A/B + A\backslash(A/B) \Rightarrow A
\]

\[
B + A\backslash B \Rightarrow A \\
A\backslash(A\backslash B) + A\backslash B \Rightarrow A
\]

Following Dowty, by raising nouns from type \(N\) to \(NP\backslash (NP/N)\) the noun phrase:

\[
\text{the man and woman}
\]

can be analysed as:

\[
\text{the} & \quad \text{man} & \quad \text{and} & \quad \text{woman} \\
NP/N & \quad NP\backslash (NP/N) & \quad Conj & \quad NP\backslash (NP/N) \\
\hline
\quad NP\backslash (NP/N) & \quad NP
\]

If the semantics of the type-raised nouns are embellished with appropriate λ-abstracts, so:

\[
\text{man} \quad \Rightarrow \quad \lambda F. F(\text{man}') \\
\text{woman} \quad \Rightarrow \quad \lambda F. F(\text{woman}')
\]

and the semantics of the conjunction, with type-raised categories, maps its argument (the determiner) to the conjuncts:

\[
\text{man and woman} \quad \Rightarrow \quad \lambda F. (\lambda F. F(\text{man}') \oplus \lambda F. F(\text{woman}'))
\]

then the semantics of the noun phrase will be:

\[
\text{the' man'} \oplus \text{the' woman'}
\]

For the singular noun phrase:

\[
\text{My friend and colleague}
\]

type shifting need not occur:

\[
\text{my} & \quad \text{friend} & \quad \text{and} & \quad \text{colleague} \\
NP/N & \quad N & \quad Conj & \quad N \\
\hline
\quad N & \quad NP
\]
resulting in the semantics:

\[ \text{my'(friend' } \oplus \text{ colleague') } \]

This approach can be used optionally to distribute the semantics of a verb phrase over syntactic conjunction, by type-raising NP to \( S\backslash(S\backslash NP) \):

\[
\begin{array}{ccc}
\text{NP} & \text{N} & \text{NP} \\
\hline
\text{the man} & \text{Conj} & \text{the woman} \\
\hline
\text{NP} & \text{NP} & \text{NP} \\
\hline
S\backslash(S\backslash NP) & S\backslash(S\backslash NP) & S\backslash(S\backslash NP) \\
\hline
S
\end{array}
\]

The double abstraction in examples like:

The man and woman died.

when interpreted as:

\[ \text{died'(the'man') } \oplus \text{died'(the'woman')} \]

would require more work, and perhaps some additional rules for combining categories.

The use of type-raising in the syntax here is reminiscent of type-lifting accounts of the distributive/collective ambiguity in the semantics [Lakoff, 1970, Lønning, 1989, Schwarzschild, 1990, Schwarzschild, 1992]. As noted in the thesis, these theories do not account for all the data, such as noun phrases read distributively with respect to one verb phrase, and collectively with respect to another; and, with Lakoff’s theory, readings where a collectively read noun phrase outscopes a noun phrase read distributively. A syntactic treatment of ellipsis covers some of the distinctions between distributive and collective readings, but it cannot provide a complete analysis.
Bibliography


[Cartwright, 1975] Helen M. Cartwright. Amount terms and measures of 
amount. _Noûs_, 9, 1975.

J.M.E. Moravcsik, and P. Suppes, editors, _Approaches to Natural Lan-

[Chierchia and McConnell-Ginet, 1990] 
Genaro Chierchia and Sally McConnell-Ginet. _Meaning and Grammar, 


semantics without types for natural languages. _Linguistics and Philosophy_, 

[Chierchia, 1982b] G. Chierchia. On plurals and mass nominals. In _Proce-
dings of the West Coast Conference on Formal Linguistics_, pages 243–255, 
1982.

[Chierchia, 1984] G. Chierchia. _Topics in the Semantics of Infinitives and 


[Cooper, 1983] R. Cooper. _Quantification and Syntactic Theory_. D. Reidel, 
Dordrecht, 1983.

[Davies, 1990] Nick Davies. Towards a first-order theory of reasoning 
agents. In _Proceedings of the European Conference on Artificial Intelli-
BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


