Some Notes on Mass Terms and Plurals

Ann Copestake
Computer Laboratory, University of Cambridge
New Museums Site, Pembroke Street, Cambridge CB2 3QG, UK

November 1989

Abstract

This report describes a short investigation into some possible treatments of mass nouns and plurals.

Acknowledgement

This work was supported by SERC grant GR/D/57713 held by Steve Pulman and Mike Gordon and draws on earlier work on the same project by Per Hasle and Thomas Forster. I would also like to thank Dick Crouch for his extensive help and advice.

Contents

Chapter 1

Introduction

This project involves an attempt to find a linguistically plausible and computational implementable theory which dealt with a range of problems associated with the treatment of mass nouns and plurals. The aim was to provide a grammar and axiomatisation with a reasonable coverage of these phenomena, so that a range of sentences could be parsed and certain types of inference could be made automatically.

The previous work on the project, reported in Hasle(1988), was based on two theories described by Link (1983, 1984, 1986) and Pelletier and Schubert (1986). In Chapter 2 and Chapter 3 of this report I review Hasle's work and the papers on which it was based. This review is intended to make apparent some limitations of both the original theories and Hasle's implementation which meant that I could not simply extend the previous work. I do not attempt to give any sort of comprehensive coverage of other work on mass terms and plurals, but I discuss some more recent work, particularly that relevant to Link's theory, in some detail, and informally introduce my own treatment of some of the issues discussed.

In Chapter 4 I describe the more significant aspects of my own grammar and axiomatisation. These cannot be regarded as an implementation of any particular theory, but mostly draw on Link's papers and those of Krifka(1987) and Roberts(1987). In Chapter 5 I discuss the problems with my approach (the ones I know about!); in many cases it seems that the possible solutions must be considered in a much wider context than I have time or ability to investigate. I do not claim that my approach itself is of any great significance, rather the theme of these chapters is to show what sort of phenomena a relatively straightforward approach can treat, and to demonstrate the implications of some of the gadgetry introduced in more complex theories.

The implemented grammar covers everything that was treated by Hasle's implementation, extends that coverage in a variety of ways, and provides a better integration of the treatment of mass nouns and plurals than was achieved in the earlier work. It was written in the same CFG+ formalism as Hasle used; this was developed by Steve Pulman (Pulman, forthcoming). Some parts of the axiomatisation have been tested using the HOL system

(Gordon 1987). The main limitations on the grammar coverage are that the grammar is basically extensional and that it contains no treatment of any temporal phenomena; some relevant aspects are briefly discussed in the report.

Chapter 2

Plurals and mass terms — Link, Krifka and Pelletier and Schubert's theories

This chapter contains a very sketchy and partial review of some recent work on plurals and mass terms, concentrating on Link's theories and related work. In this chapter I assume familiarity with Link's papers on the logic of plurals, in particular Link(1983). However I start off by reviewing the semilattice concept used by Link, in order to introduce the notation and to make clear the relationship to purely set-theoretic concepts (to a large extent following Landman (1987) — Link (to appear) also draws attention to the parallels).

Although Hasle's work took Pelletier and Schubert's draft paper (1986) as the basis for the treatment of mass terms, I will leave detailed discussion of that theory to the end of this chapter. The reason for this is that the paper is more concerned with generic and kind readings of mass terms rather than ordinary quantity readings. For example the sentence:

John threw snow at Mary.

would be translated initially in terms of the kind snow; the ordinary reading involving a quantity of snow would have to be recovered by meaning postulates. These meaning postulates had not been implemented by Hasle, and appear not to be suitable for implementation without considerable modification. Furthermore the limitations on the coverage of Pelletier and Schubert's theory make it very difficult to see how it could be integrated with Link's theory without considerable modification and extension. Many central aspects of non-generic mass term behaviour, for example cumulativity, are not dealt with, and the theory presented in that paper does not deal with bare plurals, which have a similar pattern of behaviour to bare mass nouns. (Pelletier and Schubert (1988) does consider both.)

I have therefore attempted to deal with non-generic mass terms and non-generic plurals together, basing the treatment on Link(1983) and the modifications and extensions to this proposed by Krifka(1987), and tried to extend this to deal with sentences involving bare plurals and bare mass terms as time permitted.

So the first section of this chapter is a brief review of the semilattice concepts used in Link's logic of plurals, LP. The second section outlines Link's theory of mass terms, LPM. Link makes a distinction between individual entities and the quantities of matter which constitute them; in this section I consider whether this distinction does really provide a solution to the class of problem it was introduced to deal with. In the third section I introduce Krifka's treatment of nominal predicates which forms part of the basis of my own grammar. This covers mass terms, but is a simplification of Link in that Krifka does not make the individual/matter distinction. Krifka, however, extends Link's theory to cope with specified quantities in general, such as 'five ounces of gold', 'five head of cattle' and 'five cows'. I will then mention some parts of Pelletier and Schubert's theories that are relevant to the ordinary quantity readings, the issue of 'grinding' count nouns. I conclude this chapter with a brief overview of Pelletier and Schubert's theory and Hasle's implementation of it.

2.1 A brief recapitulation of Link's logic of plurals, LP

Link's logic of plurals, LP, is a first order logic which introduces a sum operation for its individual terms. This *isum* operation, $\sqcup_i$, has $\oplus$ as its syntactic counterpart. For example, $a \oplus b$ denotes an entity in the domain of individuals, made up of the individuals a and b, but itself an individual of the same type as a and b. The domain of discourse, E is closed under the isum operation and forms a complete join semilattice. [1]

For example if the atomic individuals in the domain are

$$A = \{j, m, b\}$$

then we have

$$E = \{j, m, b, j \sqcup_i m, j \sqcup_i b, m \sqcup_i b, j \sqcup_i m \sqcup_i b\}$$

The intrinsic ordering relationship on E induced by $\sqcup_i$ is $\sqsubseteq_i$, *ipart* (syntactic counterpart Π).

$$a \Pi b \iff a \oplus b = b$$

There is also a variant $^\bullet\Pi$ which is read *is an atomic ipart of.* Something is defined to be atomic if it has no iparts other than itself.

$$\forall x[\mathrm{At}(x) \iff \forall y[y \Pi x \Rightarrow y = x]]$$

[1] In Link(1983) E has a complete Boolean structure and is a lattice rather than an upper semilattice — that is it has a 0-element. However in Link(1986) E is taken to be a semilattice. I have assumed the semilattice structure here since the 0-element appears to be an unnecessary complication.

Because E is complete for any predicate P we can form an individual term $\sigma x P x$ which denotes the isum of all individuals that are P. If the extension of P is a singleton set then $\sigma x P x$ and $\iota x P x$ have the same denotation. There is also a *closure* operator, *, such that $^\star P$ forms all possible isums from the members of the extension of P, and a *proper plural* operator $^\bullet P$ which is equivalent to $^\star P$ with atomic individuals excluded. So if

$$[\![P]\!] = \{j, m, m \sqcup_i b\}$$

then

$$[\![{}^\star P]\!] = \{j, m, m \sqcup_i b, j \sqcup_i m, j \sqcup_i m \sqcup_i b\}$$
$$[\![{}^\bullet P]\!] = \{m \sqcup_i b, j \sqcup_i m, j \sqcup_i m \sqcup_i b\}$$
$$[\![\sigma x P x]\!] = \{j \sqcup_i m \sqcup_i b\}$$

There is an isomorphic set-theoretic model. If we start with a basic set A of individuals then the denotations of individual constants are not the elements of A, but sets of elements of A, the power set of A, $\mathrm{pow}(A)$. We want to exclude the empty set from the domain of individuals (although the empty set is relevant to the denotation of predicates). For example,

$$A = \{j, m, b\}$$
$$\mathrm{pow}(A) - \{\} = \{\{j\}, \{m\}, \{b\}, \{j, m\}, \{j, b\}, \{b, m\}, \{j, b, m\}\}$$

So $\mathrm{pow}(A) - \{\}$ has the structure of a complete atomic join semilattice, without the zero element, with union corresponding to Link's isum operator and subset to ipart.

I mention this here to emphasise that we can treat the semilattice structure as though it were set-theoretic, with the exception that a semilattice structure need not be atomic (see Landman 1987). I regard the importance of the semilattice to be that we need not assume the existence of atomic individuals. This makes it much easier to provide an integrated treatment of mass nouns and plurals, since it is convenient to be able to remain neutral on the issue of the atomicity of quantities (from a linguistic viewpoint, of course). I do not think any of Link's other arguments against the use of sets are of any practical significance although, because of the isomorphism, there can be no empirical evidence either way.

Bunt(1985) gives an alternative approach to the treatment of mass terms which also avoids the assumption of atomic individuals. He introduces a theory of *ensembles* which takes the concept of 'part-of' as fundamental. Ensembles need not have atomic parts — sets are, in effect, the special case of ensembles which do have atomic parts. This approach is elegant in that many of the set operations are special cases of the more general ensemble operations (subset is a special case of part-of for example). But, for current purposes at least, this is not necessarily an advantage, since the structure of the domain of individuals (given in LP by the lattice operations) is quite distinct from the denotation of predicates (which are sets in LP in the standard way).

Link's representation of the NP "some cows" would be:

$$\lambda P \exists x[{}^{\bullet}\mathrm{cow}(x) \wedge P(x)]$$

However in later papers Link does not consistently use the proper plural operator and insisting that plurality is truth-conditional does cause some problems, which are discussed in Section 2.3. So to avoid complicating the issue here I will assume that the representation of "some cows" is in fact:

$$\lambda P \exists x[{}^{\star}\mathrm{cow}(x) \wedge P(x)]$$

and the representation of both "the cow" and "the cows" is equivalent to:

$$\lambda P \exists y[y = \sigma x \mathrm{cow}(x) \wedge P(y)]$$

$\oplus$ is used to represent 'and', so "John and Mary" will be translated as

$$\lambda P[P(\mathrm{John} \oplus \mathrm{Mary})]$$

Link does assume that the distributive or Boolean 'and' within an NP will also be needed, at least for constructions such as:

John and every other student left. (Link 1986)

Partitive constructions are represented using Π. Predicates such as common nouns and verbs like 'die' are assumed to only take atoms in their extension, others take both atoms and isums; these latter are referred to as having 'mixed extension'. This concept formed part of Link's original treatment of distributivity which will be discussed in the next chapter; in this chapter I will concentrate on nominal representations.

Group terms such as "the committee" and "the team" are not represented as isums in Link's theory. Instead they are simply atoms and there no way is given of representing their relationship with the individual entities which constitute them (other than that they are made of the same 'stuff', see below). The semilattice constructions in LP represent syntactically marked plurality and conjunction alone. In the next section I illustrate how Link does use a semilattice to represent material constituency, but only when considering mass terms and 'stuff', not individual objects.

2.2 Link on mass terms

In the theory LPM (the logic of plurals and mass terms) Link proposes a distinction between an entity and the 'stuff' which constitutes it (constituency is represented by $\triangleright$). Material constituency of 'stuff' is represented using a semilattice but this is distinct from the semilattice of ordinary individuals. So there is an operator, $+$, *material fusion*, such that $a + b$ constitutes but is not identical to, $a \oplus b$. The *material part* relation between the stuff constituting entities is denoted by $\top$:

$$a \Pi b \Rightarrow a \top b$$

So $+$ and $\top$ are defined on the whole domain of individuals E, but their semantics depend on D, the set of portions of matter, which is a subset of A, the atoms of E. All entities which constitute another entity are members of D; members of D have themselves as constituents. D itself forms a semilattice under the $+$ operation, but it is not necessarily atomic with respect to $\top$.

Link claims that the distinction between objects and their constituent stuff provides a solution to the puzzle of how a sentence like:

> This ring is new but the gold which it is made out of is old.

can be true. Link would give this the following sort of representation:

$$\text{old}(\iota x[\text{gold}(x) \wedge x \triangleright a]) \wedge \neg\text{old}(a)(\text{where } a \text{ is the ring})$$

New can be true of the object and old can be true of the stuff out of which the object is made; the object and the stuff out of which it is made are not identical but are related by the constitution relation.

However there are problems with this. Bach(1986) comes to the conclusion that if this approach is adopted an indefinite number of levels of constituency may be needed. His example is of a snowman, which is made out of snow, which is made out of water molecules. The H_2O molecules may be old, yet the snow is new. I think that the problem is deeper than this, however. My interpretation of the problem sentence is something like:

> This ring was made recently out of gold that was mined a long time ago.

The interpretation cannot be that the gold considered purely as stuff is old, because considered in terms of its atoms all gold is of roughly the same age. So we are not really applying 'old' to the stuff at all; really 'old' is being applied to some previous object (the nugget that was mined, some previous artifact, or whatever). This actually suggests that although we could currently equate the ring with the gold from which it is made, we will have to be careful when introducing predicates such as 'old' and 'new' which refer back to a time when the object did not exist, but the stuff which constitutes it did.

I think that a formal treatment of such puzzles will involve allowing predicates to takes context-dependent aspects of an entity as arguments. Obviously we sometimes utter sentences that make distinctions between an artifact and its constituents, but there seem to be an indefinite number of aspects of an object that we can distinguish other than constituency. For example:

We've got a new car.

can be said (in certain contexts) even if the car is second-hand. We cannot enumerate all the various aspects of an object that we may need to consider; the particular aspect will normally be made clear by the predicate and the context, and I believe that interpretation may involve an indefinite amount of inference to find the appropriate aspect of an entity. There are parallels with the relationship between groups and the individuals constituting them (see the discussion in Section 5.3) which should be brought out in a full treatment (Link's constitution operator does not apply to these relationships).

A further problem with Link's paper is that it does not give any indication of how to treat mass terms such as "furniture", which are not so obviously true of stuff as mass nouns like "gold". It seems as though we would have to treat them both in the same way in this theory, but it does not seem appropriate to claim that "this furniture" constitutes "these chairs" and if this approach were adopted we could not specify that "this wood" constituted "this furniture".

So Link's distinction between individual objects and the stuff that constitutes them does not seem to be adequately motivated. My own treatment is more similar to that of Krifka, discussed in the next section, which does not make the individual/stuff distinction. An interim solution to the particular problem discussed by Link is simply not to treat 'new' and 'old' as purely intersective adjectives, but to interpret them relative to the noun. (This would of course be desirable for independent reasons, compare 'an old gymnast' and 'an old woman'.)

Link does not treat the more generic uses of mass terms and specifically excludes from consideration the problem of 'nominal mass terms', by which he means, for example, the use of gold in

Gold has atomic number 79

2.3 Krifka's theory of nominal reference

Krifka assumes that objects, including quantities of matter, can be characterised by a predicate O, the extension of which has the structure of a complete, complementary, join semilattice without the 0-element. I will adopt his notation, $\sqcup_o$ for join, $\sqsubseteq_o$ for part and $\sqsubset_o$ for proper part. The fusion operation generalises the join operation, mapping a set to its lowest upper bound.

$$\forall x, P[FU_o(P) = x \iff \\ \forall x'[P(x') \Rightarrow x' \sqsubseteq_o x] \wedge \forall x''[\forall x'[P(x') \Rightarrow x' \sqsubseteq_o x''] \Rightarrow x \sqsubseteq_o x'']]$$

Various higher order predicates and relations are introduced to characterise different reference types. The most important of these for my purposes are repeated here.

$$\forall P[\mathrm{CUM_o}(P) \iff \forall x, y[P(x) \wedge P(y) \Rightarrow P(x \sqcup_o y)]]$$

P has cumulative reference; if P is true of two entities it will be true of the join of the two entities.

$$\forall P[\mathrm{QUA_o}(P) \iff \forall x, y[P(x) \wedge P(y) \Rightarrow \neg(y \sqsubset_o x)]]$$

P has quantised reference; if P is true of an entity it will not be true of any proper part of that entity. Except in the case where a predicate P has singular reference

$$\forall P[\mathrm{SNG_o}(P) \iff \exists x[P(x) \wedge \forall y[P(y) \Rightarrow x = y]]]$$

it cannot be both cumulative and quantised. Mass nouns like gold′ and plurals like horses′ have cumulative reference, but singulars like horse′ and measured objects like (five ounces of gold)′ have quantised reference.

$$\forall P, x[\mathrm{ATOM_o}(x, P) \iff P(x) \wedge \neg\exists y[y \sqsubset_o x \wedge P(y)]]$$

x is a P-atom iff x is P and there is no proper part of x which is P.

$$\forall P[\mathrm{ATM_o}(P) \Rightarrow \forall x[P(x) \Rightarrow \exists y[y \sqsubseteq_o x \wedge \mathrm{ATOM_o}(y, P)]]]$$

If P has atomic reference then every x which is P is either a P-atom itself or has a P-atom as a proper part.

It is important to note that the definition of atomicity is with respect to a particular predicate. This an essential difference in the formal apparatus between this theory and Link's LP. Link defines atomicity with respect to Π but Krifka remains neutral on whether the semilattice is atomic with respect to $\sqsubseteq_o$. Krifka's join and part operators are used in such a way as to make them operators on material rather than individuals; this means that Krifka's theory has closer parallels with mereological theories than Link's does.[2] However there is still no question of making the denotation of predicates anything other than a set.

[2]This is something which I would have liked to investigate if I had the time and the competence. There are certainly strong superficial similarities with the formal apparatus developed by Lesniewski and reformulated by Leonard and Goodman, described in Bunt(1985).

Krifka gives the following representations for phrases with specified quantities at the Nbar level:

five ounces of gold
$\lambda x[\mathrm{gold}'(x) \wedge \mathrm{ounce}'(x) = 5]$
five head of cattle
$\lambda x[\mathrm{cattle}'(x) \wedge \mathrm{NU}(\mathrm{cattle}')(x) = 5]$
five cows
$\lambda x[\mathrm{COW}(x) \wedge \mathrm{NU}(\mathrm{COW})(x) = 5]$

These are similar in that they are all treated as though a measure phrase is applied to the head noun. The measure function may be specified directly (ounce′ for example) but in the case of classifier constructions and ordinary count noun constructions it will depend on the head noun. In this case the function symbol NU is applied to the nominal predicate to yield a measure function. (Krifka discusses the properties that are necessary for measure functions and the conditions under which a measure phrase is well formed, but these are not important to this discussion.) The representation of 'five cows' uses COW which is assumed to be a nominal predicate similar to cattle′, which underlies the count noun cow, but with no direct representation in English. It seems clear that Link's *cow could fill this role in his system. However in a grammar based on Krifka's treatment it appears that we have no need for a predicate which is only true of individual cows since 'a cow' and 'one cow' can both be represented at the NP level [3] as:

$$\lambda P \exists x[\mathrm{COW}(x) \wedge \mathrm{NU}(\mathrm{COW})(x) = 1 \wedge P(x)]$$

From now on I will assume that all nominal predicates introduced in the translations in Krifka's theory can be true of singular and plural entities. It seems redundant to mark all nominal predicates in the same way, wherever they are introduced. If we did have a predicate P true only of singular count individuals then I assume that the denotation of the predicate introduced in the translations would be equivalent to $^{\star}P$, since in this case closure over $\sqcup_o$ is equivalent to closure under $\oplus$. Since Link sometimes has to introduce atomicity separately into the translations anyway, even a grammar fragment based on Link that did not contain any means of numerical specification could dispense with the unstarred predicates.

Plural phrases which are not numerically specified, for example, 'some cows', are translated by Krifka in such a way that they could be true of singular entities.

$$\lambda P \exists x, n[\mathrm{cow}(x) \wedge \mathrm{NU}(\mathrm{cow})(x) = n \wedge n > 0 \wedge P(x)]$$

[3] Krifka's paper excludes any discussion of quantifiers and his thesis (Krifka 1986), where they are discussed, is inaccessible to me. I have therefore assumed the standard treatment for non-generic sentences.

The reason for not excluding singular entities is basically that phenomena such as dependent plurals make syntactic plurality a poor guide to semantic plurality. Grice is invoked to explain why such a phrase is not normally used by a speaker who has sufficient evidence that there is only one cow. It does seem that 'some cows' can be used in contexts where only a single individual is referred to, but where the existence of any such individual is important.

Some cows can undo bolts on feed bins.

seems acceptable even if only one individual cow has ever performed this feat of bovine sagacity. In Link(1983) plurality is truth conditional and the proper plural operator would be used in the representation of "some cows", so the sentence would actually be regarded as false in this situation.

Krifka's translation of the definite article makes use of the fusion operator so "the gold" is:

$$\iota x[x = FU_o(\text{gold}') \wedge \text{gold}'(x)]$$

Krifka therefore makes use of similar lattice operations to Link. However the underlying ontology of his theory seems to be very different. In Krifka's theory the lattice operations do represent material constituency. The domain of matter is structured by the join and part-of relationships. No concept of atomic individual is built into the model. Certain predicates, normally corresponding to singular count nouns, will only be true of entities that correspond to what we normally think of as singular individuals, but there is nothing to prevent an entity from being singular with respect to one predicate and plural with respect to another. Other predicates, corresponding to mass nouns, will not carry any connotations of individuality at all. So a particular entity in the semilattice might be equally well be referred to as "this technical report", "65 sheets of paper" and "garbage".

In Krifka's theory we need extra information, besides that given in the logical forms, to make inferences such as:

Some men are students.
$\exists x[\text{student}(x) \wedge \text{NU}(\text{student})(x) > 0 \wedge \text{man}(x) \wedge \text{NU}(\text{man})(x) > 0]$
$\Rightarrow$ A man is a student.
$\exists y[\text{student}(y) \wedge \text{NU}(\text{student})(y) = 1 \wedge \text{man}(y) \wedge \text{NU}(\text{man})(y) = 1]$

I will assume that the value of a measure function such as NU(man) is a positive integer[4], but I also need the information that NU(man) = NU(student)

[4]Krifka apparently intended phrases like "half a cow" to be translated by using non-integral values of natural units, but this does seem to lead to two halves of different cows being indistinguishable from one whole cow. I have therefore assumed that the natural unit measure function does reflect the integrity of individuals. This modification does require a change to Krifka's axioms for measure functions and it is possible that I am overloading the concept of natural units here.

and also that both predicates distribute down to individual entities. Krifka states the measure functions for all living individuals will probably be identical.

Contrast this with Link's theory. In effect Link makes the assumption that all count nouns have identical ways of individuating, because atomicity is defined by the lattice operations and is not relative to a particular predicate. The inference above is simplified because we only need the information that "student" and "man" are both distributive predicates (see Section 3.1).

Some men are students.
$\exists x[{}^{\star}\text{student}(x) \wedge {}^{\star}\text{man}(x)]$
$\Rightarrow$ A man is a student.
$\exists x[\text{student}(x) \wedge \text{man}(x)]$

But there do seem to be some cases where equality is asserted between entities which are individuated differently by different predicates.

These students are the college's two football teams.

If (as Link suggests) group terms such as "football team" are atomic then we can derive

A student is a football team.

in exactly the same way as we derived

A man is a student.

Krifka's theory does allow us to specify NU(footballteam) $= 11 * \text{NU}(\text{human})$ so we can deduce that each football team has eleven students as members, if we have the extra information that although two individual football teams may have overlapping parts, two individual students do not. Individual football teams have quantised reference:

$$\forall x, y[\text{ft}(x) \wedge \text{NU}(\text{ft})(x) = 1 \wedge \text{ft}(y) \wedge \text{NU}(\text{ft})(y) = 1 \Rightarrow \neg y \sqsubset_o x]$$

but I have to introduce a stricter condition for individual students.

$$\begin{aligned}&\forall x, y[\text{student}(x) \wedge \text{NU}(\text{student})(x) = 1 \wedge \\ &\quad \text{student}(y) \wedge \text{NU}(\text{student})(y) = 1 \wedge x \neq y \\ &\quad \Rightarrow \neg\exists z[z \sqsubseteq_o x \wedge z \sqsubseteq_o y]]\end{aligned}$$

So although Krifka's theory does increase the complexity of the inferences this does actually seem to be potentially useful. It provides some indication of why it is difficult to count the 'things' on my desk, for example. When the entities involved are similar then the measure functions will be the same, and there are other cases where the context provides sufficient information about the individuation to make counting possible. It seems to me that it is desirable to avoid the assumption that there is some consistent notion of individuality. Even considering living organisms there are severe problems in deciding whether something like a stand of bamboo or a Portuguese man-of-war is an individual or a colony of individuals[5]. In languages with a count/mass distinction, mass terms are normally used when intuitions about individuation and countability are unclear, but there is no agreement about this between languages. A theory such as Krifka's which makes individuation dependent on particular predicates seems to account for this much more naturally than a theory in which atomicity is built into the model.

Krifka remains neutral as to whether the semilattice of matter is really atomic or not. Because of this, the effect of my adopting a particular position on the minimal parts problem will be limited; an alternative viewpoint could be taken without fundamentally changing the theory. In fact I accept Bunt's *homogeneous reference hypothesis*:

> Mass nouns refer to entities having a part-whole structure without singling out any particular parts and without making any commitments concerning the existence of minimal parts. (Bunt, 1985, p46)

I will assume, for example, that all parts of a quantity of gold are gold.

$$\forall x[\text{gold}(x) \Rightarrow \forall y[y \sqsubseteq_o x \Rightarrow \text{gold}(y)]]$$

(It may be worth pointing out that although we may not want to assume that mass terms are atomic linguistically there is also no reason to assume that $\sqsubseteq_o$ is used to refer to sub-molecular parts. Therefore there is no reason to conclude that "electrons are gold", for example.)

I have made a further extrapolation from Krifka(1987) in my treatment of "ground" count nouns. It is well known that most count nouns can be used as mass nouns; even without a grinding machine there are some contexts where we may only be interested in the substance and not in the existence of any individual; food contexts are the most obvious. There are a few nouns, those used in partitives, which only seem to restrict physical shape or quantity, for example, "puddle", "cube" and "ounce" are difficult (if not impossible) to "grind". But since the vast majority of count nouns can be used as mass nouns it seems unreasonable to give dual lexical entries. Pelletier and Schubert (1986) suggest two treatments; one, the s-theory (presumably due to Schubert), makes use of "lexical extension rules" to form a mass sense from count nouns. The mass sense is marked as being in some way abnormal. But if the idea that these senses are "extended" is to have any use, there would

[5]Gould(1987) discusses this from a biologist's viewpoint and comes to the conclusion that there is a continuum rather than a dichotomy between the concepts.

have to be degrees of abnormality, corresponding to the likelihood of use in a mass sense, and these would have to be assigned arbitrarily and permanently in a static lexicon. The other treatment described, the p-theory (presumably Pelletier's) avoids making any distinction between the mass and count sense at all, but this seems to go too far.

Krifka's theory suggested a slightly different possibility to me, that the predicate could be true of mass and count senses, but that a count sense would have a specified value of natural units. So, for example, "some lamb" would be translated as:

$$\lambda P \exists x[\mathrm{lamb}(x) \wedge P(x)]$$

which could be true either of an individual lamb or of a quantity of lamb, but "some lambs" would be:

$$\lambda P \exists x, n[\mathrm{lamb}(x) \wedge \mathrm{NU}(\mathrm{lamb})(x) = n \wedge n > 0 \wedge P(x)]$$

I introduced this treatment (without any great expectations of it working) because it seemed true to Pelletier's motivation in removing the mass/count distinction but avoided some of the obvious problems. It worked well enough to be retained in the grammar that I will discuss in Chapter 4, but there are difficulties, some of which will be discussed in Section 5.2. One effect of this decision is that for all nominal predicates P (not just conventional mass predicates), all parts of an individual which is P are also P, that is they are fully distributive.

$$\forall P[\mathrm{NOM}(P) \Rightarrow \forall y[P(y) \Rightarrow \forall x[x \sqsubseteq_o y \Rightarrow P(x)]]]$$

So all parts of a man are man and so on.

It is reasonable to assume that for most count nouns in most contexts there is a default assumption that:

$$P(x) \Rightarrow \exists n[\mathrm{NU}(P)(x) = n]$$

and therefore that "the man" for example will normally be assumed to refer to an individual. A similar argument would hold for mass nouns like furniture which do have obvious individuating parts; "this chair leg is furniture" would be odd because it contradicted that default assumption. So, rather than Schubert's lexical extension rules, this places the decisions about the oddness of the mass sense in the inference system. This is preferable, because the degree of abnormality would be deduced from reasoning about the entire context on the basis of the meaning of the words, and would not depend on particular lexical items.

2.4 Pelletier and Schubert on kinds

As I mentioned at the beginning of this chapter Pelletier and Schubert's theory, as described in Pelletier and Schubert (1986), only really makes sense as an attempt at a general treatment of bare noun phrases, both mass terms and plurals (and their general ideas are much better motivated in Pelletier and Schubert(1988)). I should reiterate that the earlier paper is only a draft and it is reasonable to assume that the authors may have refined their ideas. In this section I will briefly discuss the theory and some of the problems that arise when attempting to implement it, which are mostly caused by the difficulties of deriving non-kind readings from the initial kind reading. This part of the theory is very badly specified, so while it is possible to implement the grammar, it seems almost impossible to formalise the meaning postulates correctly. I will also attempt to motivate my very crude treatment of kind readings which is described more fully in the next chapter; possibilities for a more sophisticated treatment are described in Section 5.4. I am following Hasle's interpretation of Pelletier and Schubert here; the paper is sufficiently vague in places that the authors may have intended something different, but I do not think that this affects my main criticisms.

Pelletier and Schubert's theory introduces several predicate modifiers; an unmodified predicate P is assumed to have both kinds and quantities in its extension. μ is the name forming operator, so μbeer names the *kind* 'beer'. βP denotes the *conventionally recognised subkinds* of P (bitter, Guiness for example), γP denotes the *conventionally recognised servings or portions* of P (as in "Do you want a beer?"), and δP denotes a quantity of P (eg the pint of beer in my glass). Underlying all this is supposed to be a semilattice of kinds. However Pelletier and Schubert never make use of the union operator and its suggested interpretation as 'mixed with' appears completely wrong to me, because it would mean that

1966 Dow is vintage port.
vintageport(μDow1966)
1977 Fonseca is vintage port.
vintageport(μFonseca1977)
$\Rightarrow$ 1966 Dow mixed with 1977 Fonseca is vintage port.
vintageport((μFonseca1977) $\cup$ (μDow1966))

which is obviously invalid. As I will suggest later a join operator may be needed to interpret 'and' in the kinds domain but Pelletier and Schubert seem to intend something different. I will therefore assume that there is a *hierarchy* of kinds, since the structure of the domain that they appear to assume can be more simply described in this way. So I use $<_k$ to indicate the partial ordering on the domain of kinds:

$$\forall x, y, z[x <_k y \wedge y <_k z \Rightarrow x <_k z]$$

Hasle introduces the notation ξP, which denotes the subkinds of P, whether conventionally recognised or not. ξP can be defined as:

$$\forall P, x[\xi P(x) \iff x <_k \mu P]$$

It is difficult to know what to make of the conventional serving reading because the discussion in the paper is very muddled. In Hasle's implementation there are three types of things; kinds (conventional and otherwise), quantities, and conventional servings. Pelletier and Schubert seem to suggest that this is correct in one place, in another to suggest that conventional servings are kinds, and in another that they are quantities. (I would have thought that they were just a quantity with a conventionally understood measure function, but Pelletier and Schubert hardly discuss ordinary quantity readings and don't say anything about measured quantities.) There seems to be a suggestion that the conventional portion reading can be used for the ordinary individuals as well as conventional servings of their mass ("a lamb"). There seems to be no way in the p-theory (where no distinction is made between mass and count nouns) to distinguish between "lamb" (the animal) and "lamb" (the meat). Because of these deficiencies the theory would need a great deal more work before it could be taken seriously as a treatment of mass nouns in general.

The theory attempts to translate bare mass nouns, in any sentence position, as kinds. This immediately leads to the problem of accounting for the truth of

> Cheap wine is wine.

while

> Cheap wine is a wine.

seems false. Pelletier and Schubert account for this by introducing the concept of *conventional* kinds; cheap wine is a *kind of* wine but not a conventional kind of wine. (The semilattice only applies to kinds, not conventional kinds.) It seems to me that their notion of kind has by now become completely abstract; it obviously bears little relationship to the use of 'kind' in ordinary language since:

> Cheap wine is a kind of wine.

doesn't seem to me to be true.

Pelletier and Schubert's use of this concept of kind might be acceptable if they specified how ordinary quantity readings are to be extracted from sentences where the initial translation of the bare mass noun is as a kind. For example "snow" is translated as a kind in the sentence:

John threw snow at Mary.

and somehow a quantity reading has to be recovered. The meaning postulates that are supposed to do this have not been formally specified, and attempts at giving them a formal specification seem to lead to contradictions with the earlier part of the paper. It also seems to me that incorrect predictions might be made about the "kind"; it just doesn't seem to me to be true that "the kind snow" has been thrown and if this doesn't seem too serious consider:

Police discovered heroin during the raid.

as opposed to:

Fleming discovered penicillin during the 1920s.

As I stated earlier, I have been mainly concerned to give reasonable logical forms for ordinary quantity readings. In order to treat the same inferences as Hasle did I adopted a very crude strategy. I assumed that all bare mass nouns and bare plurals were ambiguous between three readings at the NP level; one existentially quantified over quantities, one universally quantified over quantities, and one nominal kind reading. Disambiguation between these three possibilities relied on the verb phrase; some predicates are assumed to only be true of kinds, for example,

The poodle is a breed of dog.
Claret is a (kind of) wine.
Clarets are wines.
Bell invented the telephone.
Water is referred to as H_2O.

Other predicates are assumed only to be true of objects (there are one or two verbs like "discovered" which do seem to be true of both, although possibly with somewhat different interpretations; these could be given dual lexical entries). Stative VPs, such as "is wine" force the universal sense to be selected, whereas non-stative VPs, such as "is dripping from the tap" will force the existential sense.

I am aware of problems with this treatment; a better treatment, involving mereological sums rather than universal quantification, is discussed in Section 5.4. However I should point out that there is only one of Pelletier and Schubert's criteria which their treatment could satisfy which mine, as stated here, does not, and that is the treatment of sentences such as:

Claret is wine and is in short supply.

In the treatment outlined above

Claret is wine.

would be translated as being universally quantified but

Claret is in short supply.

would have to involve the kind claret. (This could be dealt with by type-raising, see Section 5.4). For my current purposes such objections did not seem to outweigh the advantages of having an implementable treatment which could be integrated with a treatment of mass term quantities and plurals.

So instead of assuming that an unmodified predicate P applies to kinds and quantities I will assume that most unmodified predicates apply only to quantities and objects, but that a predicate that is true only of kinds can be formed by the predicate modifier "Type". (This is intended to have an interpretation similar to the ξ operator mentioned earlier but it is rather more machine readable.) I will use "Kind" instead of μ as the name forming operator. The domain of kinds, K, is an atomic join semilattice, which is disjoint from the domain of objects, O, but essentially the same join and part operators apply, since it is useful to be able to conjoin kinds, to form plural kinds, and so on, in much the same way as in the domain of objects.

Claret and sauternes are wines.
$\text{Type}(\text{wine})(\text{Kind}(\text{claret}) \sqcup_k \text{Kind}(\text{sauternes}))$

The hierarchy of kinds is implemented by a further structuring relationship, $<_k$ which is not the same as the semilattice part operator since:

$$\text{Kind}(claret) <_k \text{Kind}(\text{wine}) \quad \text{but} \quad \neg[\text{Kind}(\text{claret}) \sqsubseteq_k \text{Kind}(\text{wine})]$$

Similar axioms to those introduced earlier apply:

$$\forall x, y, z[x <_k y \wedge y <_k z \Rightarrow x <_k z]$$
$$\forall P, x[(\text{Type}(P))(x) \iff x <_k \text{Kind}(P)]$$

but the theorem which derives from them, which will be used in the natural language inferences, is:

$$\forall P, Q, R[\text{Type}(P)(\text{Kind}(Q)) \wedge \text{Type}(Q)(\text{Kind}(R)) \\ \Rightarrow \text{Type}(P)(\text{Kind}(R))]$$

The only connection between the kinds and the objects domain which I need to assume is[6]:

$$\forall P, Q, x[\text{Type}(P)(\text{Kind}(Q)) \Rightarrow \forall x[Q(x) \Rightarrow P(x)]]$$

[6] If the theory was modified along the lines suggested in Section 5.4 this might be better reformulated in terms of the mereological sum.

$$\forall P, Q, x[\text{Type}(P)(\text{Kind}(Q)) \Rightarrow P(FU_o(Q))]$$

Chapter 3

Distributivity

This chapter consists of a discussion of some of the recent suggestions for treatments of distributivity, concentrating on Link's theories and related work. I will first outline my interpretation of the treatment of distributivity given in Link(1983). I will then briefly describe some problems with Hasle's interpretation of Link with respect to distributivity, which indicates why I have not followed the approach taken in the earlier report (Hasle 1988). The third section discusses some inadequacies in Link's approach to distributivity, and in Landman's work (Landman 1987) which is based on Link. The fourth and last section gives an informal description of what I believe is a better approach to distributivity, on which the treatment described in more detail in Chapter 4 is based. This section especially is influenced by Roberts(1987).

3.1 Link(1983) on distributivity

Link's discussion of distributivity is based on the idea that predicates can be of two types. *Atomic* predicates are those which basically only include atoms in their extension; these include all common nouns (including group nouns) and verbs like "die". Predicates with *mixed extension* take both atoms and isums; these include verbs like "gather" and "convene" (which will only take atoms in their extension if these atoms correspond to groups) and also verb phrases like (lift a piano)$'$, which can be true of non-group individuals, as well as isums. In order to form predicates which can take isums out of the atomic predicates the closure operator has to be applied. The representation of

John and Mary die.

is therefore

$$^{\star}\text{die}(j \oplus m)$$

The closure operator is not applied to mixed extension predicates, so

John and Mary convene.

is represented as

convene($j \oplus m$)

Distributivity will then only arise for predicates with atomic extension. This is accomplished by defining 'Distr' as

(D.19) $\quad \text{Distr}(P) \iff \forall x[P(x) \Rightarrow \text{Atom}(x)]$

Distr is true of all atomic predicates and the following theorem can be derived:

(T.10) $\quad \text{Distr}(P) \Rightarrow \forall x, y[{}^\star P(y) \wedge x \Pi y \Rightarrow {}^\star P(x))]$

So from ${}^\star\text{die}(j \oplus m)$ and Distr(die) it follows that ${}^\star\text{die}(j) \wedge {}^\star\text{die}(m)$. Since

(T.8) $\quad \forall x[\text{Atom}(x) \Rightarrow [P(x) \iff {}^\star P(x)]]$

we can derive $\text{die}(j) \wedge \text{die}(m)$, that is:

John dies and Mary dies.

3.2 Hasle's implementation of Link

There are several aspects of Hasle's implementation of Link's 1983 theory, as described in the previous report, which I find unsatisfactory. The most important problems, which are the only ones I will consider in detail here, concern the successive redefinitions of Link's closure operator, in an attempt to use it to represent distributivity directly. (In the discussion that follows I will use the term 'star operator' to refer to the redefined closure operator.)

3.2.1 The one-place closure operator

In Hasle(1988) Chapter 9 an initial implementation of Link(1983) is presented which does not cover transitive verbs. In this implementation the star operator is always assumed to license distributivity, that is Hasle introduces as an axiom

Axiom 9 $\quad \forall x, y[{}^\star P(x) \wedge y \Pi x \Rightarrow {}^\star P(y)]$

instead of Link's theorem

(T.10) $\quad \text{Distr}(P) \Rightarrow \forall x, y[{}^\star P(x) \wedge y \Pi x \Rightarrow {}^\star P(y)]$

which is derived from the definition of 'Distr' as

(D.19) $\quad \text{Distr}(P) \iff \forall x[P(x) \Rightarrow \text{Atom}(x)]$

In Hasle's translations, as in Link's, all distributive verbs are starred, but no non-distributive verbs are. [1]

John and Mary die.	$^{\star}\mathrm{die}(j \oplus m)$
John and Mary convene.	$\mathrm{convene}(j \oplus m)$

Now getting rid of the predicate 'Distr' might seem to improve the elegance of Link's theory. We could then make the inference

John dies.	$\mathrm{die}(j)$

from

John and Mary die.	$^{\star}\mathrm{die}(j \oplus m)$

without needing any other source of information such as Distr(die). To do this however we have to avoid making predicates such as 'convene' distribute down to individuals. If we simply make dissectiveness an inherent property of the closure operator, as Hasle does, we have:

$\mathrm{convene}(j \oplus m)$	(premise)	"John and Mary convene"
$\Rightarrow {}^{\star}\mathrm{convene}(j \oplus m)$	(Link's T.7)	
$\Rightarrow {}^{\star}\mathrm{convene}(j)$	(Hasle's Axiom 9)	
$\Rightarrow \mathrm{convene}(j)$	(Link's T.8)	"John convenes"

Although convene is not starred in the translation, Link's definition of the closure operator makes:

$$\forall P, x[P(x) \Rightarrow {}^{\star}P(x)]$$

a theorem.

Hasle does not in fact get these undesirable inferences but he accomplishes this by never applying his star operator to collective predicates and excluding from his axiomatisation any connection between predicates modified by the star operator and the unmodified predicate. He does not allow the inference

$$P(x) \Rightarrow {}^{\star}P(x)$$

Even singular noun phrases are translated in terms of starred predicates; the only predicates which are not starred are those corresponding to 'collective' predicates such as convene.

It is self-evident that by taking this step Hasle has completely redefined the star operator so that it is no longer the closure operator, but he does not

[1] I have simplified this slightly; Hasle's translations in this chapter (but not later) actually involve the proper plural operator, for example, $^{\bullet}\mathrm{die}(j \oplus m)$. Hasle discusses the problems with the inferences that this causes, but actually I can see little evidence that Link ever intended the use of the proper plural operator here.

mention the implications of this in the report, or give any semantics for the new operator. The effect is that there are two disjoint classes of predicates; predicates like 'student' and 'die' never occur unmodified by star in the logical form of a sentence. In effect the star operator is just being used as a tag, to carry information about distributivity up to the inference engine. There may not be anything wrong with this technically, the example inferences are certainly valid, but it doesn't seem to me to constitute an implementation of Link, and certainly loses much of the elegance of the original theory; for example, concepts like 'cumulativity' which followed from the definition of the closure operator, now have to be stated as axioms.

Link does prevent non-distributive predicates being starred during translation, but he cannot mean that predicates like 'convene' cannot have the closure operator applied to them by application of the theorem mentioned above. Link presumably avoids starring convene because he would have no way of inferring, for example convene$(j \oplus m \oplus b)$ from a translation of the form *convene$(j \oplus m \oplus b)$ In contrast 'die' has to be starred because unstarred 'die' only applies to atomic predicates. (As I mention later I think it would actually be preferable to apply the closure operator uniformly.)

3.2.2 Hasle's second definition of closure

Hasle does introduce a second definition of closure (in Chapter 10 of Hasle(1988)) which has fewer disadvantages then the treatment discussed above. The new operator is equivalent to closure over the atoms in the extension of P — the original closure operator includes any isums which may be in the extension of P. So, if for Link's closure operator we have:

$$[\![^{\star}P]\!] \equiv \mathrm{cl}_{\oplus}([\![P]\!])$$

where $cl_{\oplus}$ stands for the operation of closure under isum, then for Hasle's operator:

$$[\![^{\star}P]\!] \equiv \mathrm{cl}_{\oplus}([\![P]\!] \cap [\![A]\!])$$

where A denotes the set of atomic individuals.

For example suppose we have a predicate P with mixed extension (such as (lifted a rock)$'$) and

$$[\![P]\!] = \{j, m, j \sqcup_i b\}$$

then under Link's definition of closure:

$$[\![^{\star}P]\!] = \{j, m, j \sqcup_i b, j \sqcup_i m, j \sqcup_i m \sqcup_i b\}$$

but under Hasle's definition:

$$[\![^{\star}P]\!] = \{j, m, j \sqcup_i m, \}$$

The dissectiveness axiom given earlier does now follow from the semantic definition of Hasle's star operator. A limited connection can be made between the 'distributive' and 'collective' senses of 'mixed extension predicates'[2] which the earlier treatment would not allow for.

$$\forall x[\text{Atom}(x) \Rightarrow (P(x) \iff {}^{*}P(x))]$$

Hasle actually suggests restricting the application of this proposition to a particular class of verbs, to avoid its application to collective predicates when these take atomic, group, arguments. (This makes me wonder how seriously he is taking the semantic definition). I don't think this restriction is in fact desirable; there can be no harm in allowing the modified operator to be applied to a predicate taking any atomic argument since the application of the dissectiveness axiom can't give inappropriate results.

So the modification to the definition of closure does allow dissectiveness to apply without condition. Obviously this change produces other changes in Link's system, for example with respect to the definition of the supremum, but Hasle does not consider these effects in the report. The treatment of distributivity which results is very limited, as is discussed below, so I did not attempt to take this approach any further.

Hasle's redefinition of closure is quite similar to that of Landman(1987). Landman makes dissectiveness a property of his closure operator by making every basic predicate take only atomic arguments (rather than isums), which in the case of predicates such as convene are groups rather than ordinary individuals. (Some potential problems with this will be mentioned later.)

3.2.3 Hasle's treatment of transitive verbs

Link(1983) does not deal with transitive verbs. In order to represent distributivity in transitive verbs Hasle attempts to introduce a 'two-place closure operator' which can be used to translate verbs which are distributive in both argument places. I do not find his motivation for attempting this very convincing; it seems simpler to allow the one-place closure operator to act over complex predicates. Rather than discussing the motivation here I will just mention some problems that arise from Hasle's attempt to use a closure operator that can be applied to predicates which take multiple arguments.

The first problem is that overloading a single operator, so that it can be applied to predicates with any number of argument places, causes problems in a strictly typed inference system such as HOL, which meant that the inferences in this part of Hasle's report could not be made using the HOL system. Secondly, the two-place operator cannot be used to represent the

[2]Hasle treats mixed extension predicates as being lexically ambiguous between a distributive and a collective sense.

possibility of distributivity arising over either place of a two-place predicate. Hasle has to apply the single-place operator to complex predicates in his discussion of this.

But perhaps the most serious problem that arises for any attempt at extending Link(1983) to transitive verbs is accounting for the various scoping possibilities. For example

Some boys lifted a rock.

has a reading where there is one rock for each boy. These scope differences occur despite the fact that both noun phrases are existentially quantified. Hasle introduces quantifiers directly in the translation of distributive predicates, as well as the star operator, in order to get the correct scopings. In fact the two-place star operator is in effect not used in the translations given by Hasle's grammar, because quantifiers are always introduced in such a way that the operator is only applied to predicates with an atomic first argument! For example,

Some students see some colleges.

is represented as:

$$\exists x[{}^{\star}\mathrm{student}(x) \wedge \forall v[\mathrm{atomic}(v) \wedge v \Pi x \Rightarrow \exists y[{}^{\star}\mathrm{college}(y) \wedge {}^{\star}\mathrm{see}(v, y)]]]$$

Instead we might attempt to represent the scoping ambiguity by applying the ordinary one-place closure operator to complex predicates. If we were to represent the sentence:

Some boys lifted a rock.

as:

$$\exists x[{}^{\star}\mathrm{boy}(x) \wedge {}^{\star}(\lambda y \exists z[\mathrm{lift}(y, z) \wedge \mathrm{rock}(z)])(x)]$$

we could claim that the application of the closure operator to a complex predicate sets up a scoping context for Cooper Storage. This suggestion was made by Dick Crouch but, as he also stated, the explanation for this would presumably be that universal quantification is occurring over the isum of the boys. I did not adopt this approach in the current implementation, partly because higher order unification seems to be required in the inference system to support this approach.

In the next section I will briefly consider other attempts at modification of Link(1983) to account for distributivity.

3.3 The limitations of Link's theory of distributivity

The most serious problem that I find with Link's treatment of distributivity, and the modifications by Hasle and Landman, is the emphasis on verbs as the source of distributivity. The concept of the closure operator is potentially attractive in that it allows a representation of possibilities other than a fully distributive or fully collective reading. For example consider:

> One hundred students gathered.
> One hundred students gathered in groups of four.
> One hundred students gathered in the hall.
> One hundred students gathered in two classrooms.

Representing the first sentence as

$$\exists x[{}^{\star}\text{student}(x) \wedge \text{num}(x) = 100 \wedge {}^{\star}\text{gather}(x)]$$

allows for the various possibilities, without specifically enumerating them. (We could restrict the argument of gather to avoid the models where single students gather.) Note that an arbitrary amount of inference may be needed to decide how many groups of students there were, or it may not be possible to decide this at all; it is not possible to disambiguate this sentence into individual gathering events during parsing.

Link doesn't actually mention this sort of possibility. His discussion of distributivity in Link(1983) is centred around accounting for the distributivity of predicates, such as 'die', which are regarded as only taking atoms in their extensions. Of course, if die does only take atoms, the only model in which *die can be true of an argument is one in which die is true of all the atomic iparts of that argument. Where I think Link went wrong is in trying to avoid the closure operator being applied to predicates like 'gather' and 'convene' by making distributivity a lexical property. As I mentioned above the only reason I can see for making this step is to force the collective reading, not to prevent distributivity. I think that a better approach would have been to prevent the fully distributive readings by specifying that the arguments of 'gather' were non-atomic. The case of 'convene' is slightly more complex; it is not unambiguously collective, but rather requires that its argument be a recognised group. Consider:

> The shop stewards and the managers convened in two rooms at ACAS.
> The committees convened separately.

If 'convene' is not starred then only an incorrect, fully collective, reading, will be given for these examples. I can find no example of an unambiguously collective verb.

It seems these effects arise from a consideration of the underlying meaning of 'die', 'gather' and 'convene'; in many cases this sort of distributivity cannot be a lexical property. A detailed discussion can be found in Roberts(1987). One convincing pair of examples is:

> These men won a 100m sprint.
> These men won a 100m relay.

The first sentence is fully distributive because 'win a 100m sprint' only takes individuals in its extension but the second is not. We cannot give lexical entries which account for such a distinction; in general it may have to be made on the basis of an arbitrary amount of reasoning on real world knowledge.

Link's account of distributivity, in terms of predicates which only take atoms in their extension, does not extend intuitively to other cases of distributivity. Consider predicates of physical location, such as in the sentence:

> The cards are on the table.

This implies that each card (and each part of each card) is on the table; unlike 'die' it does not seem plausible to account for this by saying that 'is on the table' basically only takes atoms in its extension. Again the distributivity follows from the meaning of the predicate but it seems to arise in a completely different way.

Hasle's proposal, to restrict the closure operator to operating over atomic elements only, means he can only use it to account for the full distributivity of predicates such as 'die' and not for distribution over intermediate isums, nor for distributivity of predicates like 'convene' down to groups, even if these are regarded as atomic.

As mentioned earlier, Landman's proposals with respect to the use of the closure operator centre around regarding base predicates as applying to atomic entities only. Such entities may be simple individuals, in the case of predicates such as die, or groups in the case of predicates like convene and gather. The properties of distributivity and cumulativity then arise from the definition of the closure operator. Landman's approach has the advantage over Hasle's that distributivity of collective predicates down to groups can be accounted for. However, unless some sort of intermediate group formation is introduced, the equivalent of distribution down to individual isums is not possible. There seem to me to be other problems. The type proliferation resulting from the use of groups is discussed extensively in Landman(1987). He does not, however, discuss the treatment of scope ambiguities, where the same problems arise as discussed above in the context of Hasle's work.

Link(1984) also suggests that groups are needed as a separate level from isums, although their use is not central to his theory as it is to Landman's. In some ways Landman's proposals are a simplification of Link's and are developed in more detail, so if groups are needed Landman's paper might provide the basis for an implementation. I am not convinced that groups are needed however. A separate level of group representation certainly isn't a requirement for the examples discussed in this chapter and covered in my grammar, and so this issue is discussed later in Section 5.3.

In later papers Link seems to abandon the application of the closure operator to verbs and instead introduces the D-operator (which is apparently based on Dowty and Brodie's (1984) treatment of floated quantifiers). Link's papers are unclear, since the use of the D-operator changes, but it is an adverbial operator such that

$$^{D}\mathrm{VP} \equiv \lambda x \forall y[y {}^{\bullet}\Pi x \Rightarrow \mathrm{VP}(y)]$$

(where ${}^{\bullet}\Pi$ means "is an atomic part of"). In Link(1986), it is applied to complex predicates, for example:

$$^{D}(\lambda y \exists z[\mathrm{piano}(z) \wedge \mathrm{lifted}(y, z)])$$

this allows for scoping effects, which as mentioned above are problematic if the closure operator is used. It also allows predicates to be distributive on any combination of argument places. So distribution is always down to *atomic* entities; there is no use of the intermediate isums. The proposal in fact does not make any real use of the novel aspects of the original theory. In Link(1983) only one representation was given for verbs with 'mixed extension'. Since the vast majority of verbs appear to have 'mixed extension', sometimes in all argument places, avoiding ambiguity in this way is very attractive. However in the later papers the intention apparently is to give two representations, which differ according to whether the D-operator is applied or not, in such cases.

One further point is that Link's treatment of distributivity does not extend to mass terms or to distribution to individuals in group terms. For example, although by making the predicate 'is in the box' distributive, we could infer from:

The cards are in the box.

that the individual cards were in the box, we cannot infer anything from

The sand is in the box.

because 'the sand' is a quantity of stuff and is thus atomic with respect to the ipart operator.

3.4 Some suggestions

In the papers discussed above there is a certain confusion as to what these treatments of distributivity are supposed to account for. In Hasle's report there are references to individual events, and there seems to be the idea that distributivity should give us readings corresponding to individual events. Obviously there is an interaction but it is not possible in general to provide separate readings on such a basis. Consider a sentence like

Four boys lifted four rocks.

Even if each boy and each rock must participate in exactly one lifting event, which is an implausible restriction, and assuming that there is no scope difference between the noun phrases, there are nine possible ways of partitioning the boys and rocks, corresponding to various subgroups of boys lifting various subgroups of rocks. So it seems that we need a representation which allows for these different possibilities, and such a representation will also subsume the fully distributive and fully collective readings over each argument place. In many cases we do not want to disambiguate to individual events or relationships; see for example Scha's(1983) discussion of cumulative readings and sentences of the type:

100 Dutch companies own 250 American computers.

As will be explained in Chapter 4, I give the sentence above a single group-group reading when scope is not considered.

$$\exists x[\mathrm{boy}(x) \wedge \mathrm{NU}(\mathrm{boy})(x) = 4 \wedge \exists y[\mathrm{rock}(y) \wedge \mathrm{NU}(\mathrm{rock})(y) = 4 \wedge \mathrm{lift}(x,y)]]$$

In this treatment predicates such as "lift" do not implicitly or explicitly refer to single events, or single relationships between individuals. In terms of a predicate ${}^{C'}$lift which did refer to individuated events we would have:

$$\begin{aligned} \mathrm{lift}(x,y) \quad \Rightarrow \quad & x = FU_o(\lambda z[\exists u[{}^{C'}\mathrm{lift}(z,u) \wedge z \sqsubseteq_o x \wedge u \sqsubseteq_o y]]) \wedge \\ & y = FU_o(\lambda v[\exists w[{}^{C'}\mathrm{lift}(w,v) \wedge v \sqsubseteq_o y \wedge w \sqsubseteq_o x]]) \end{aligned}$$

The unmodified predicate restricts the possible interpretations in terms of individual lifting events by specifying that all the material parts of x were material parts of subjects in some lifting event whose object was some material part of y, and all the material parts of y were material parts of objects in some lifting event whose subject was some material part of x. [3]

[3] Although the formal relationship given does stipulate full involvement I am aware that this is not correct, but I don't think this affects my main argument. It might be possible to refine the definition by making use of a context dependent quantifier. Bach introduced ENOUGH to deal with a similar problem with reciprocals, for example.

In general:

$$\begin{array}{lll} P(x) & \Rightarrow & x = FU_o(\lambda z[{}^C P(z) \wedge z \sqsubseteq_o x]) \\ P(x,y) & \Rightarrow & x = FU_o(\lambda z[\exists u[{}^{C'} P(z,u) \wedge z \sqsubseteq_o x \wedge u \sqsubseteq_o y]]) \wedge \\ & & y = FU_o(\lambda v[\exists w[{}^{C'} P(w,v) \wedge v \sqsubseteq_o y \wedge w \sqsubseteq_o x]]) \end{array}$$

I will not consider event semantics any further in this report, except to mention the parallels between the structure of the domain of objects and that of the domain of events, which are considered in Link(1987), Krifka(1987) and Bach(1986).

Link's treatment of distributivity in Link(1983), was based on whether particular verbs could only take atoms in their extensions. As was argued above whether a predicate only takes individuals or not is rarely a property of a single lexical item, but instead depends on an arbitrary amount of reasoning on real world knowledge. We could categorise all this as distributivity due to meaning. The modification to Link that I suggested above was to star all predicates, so that distributivity need not be a lexical feature. In fact the treatment I have implemented is in some ways similar to this idea, with the exception that I do not introduce a closure operator, but interpret the unmodified predicate in an equivalent way. Some predicates, like convene, by virtue of their meaning must refer to groups, and so no interpretation in which individuals convene will be true.

The students each convene.

will be given the reading

$$\forall x[\text{student}(x) \wedge \text{NU}(\text{student})(x) = 1 \Rightarrow \text{convene}(x)]$$

but this will always be false. Dowty's explanation, in terms of *distributive subentailments*, seem to me a plausible way of handling various aspects of distributivity due to meaning, including the distributivity of 'die'. Rather than restricting the base predicate 'die' so that it only applies to individuals we can deduce, from the meaning of die, that its application to groups will normally imply its application to the individuals in that group.

This explanation also accounts for something which Link's theory does not; there seem to be some predicates which are always fully distributive, in that they are always true of all the individuals involved, even on a collective reading. Consider:

John and Mary climbed the North Face of the Eiger.

Now climbing can be a highly collective activity (that is, unlike 'die', when it is true of a group it does imply significantly more than it would if it were just true of individuals within that group) but even on the collective reading of this sentence it is still true that:

John climbed the North Face of the Eiger and Mary climbed the North Face of the Eiger.

It appears that 'climb the North Face of the Eiger' is fully distributive from the point of view of these inferences, yet we do not want to exclude the collective reading by saying that the basic predicate can only take atoms in its extension or by always applying the D-operator.

So far my approach to distributivity has been to deny its existence at the level of parsing. Unfortunately there are some problems that I have to address. Firstly some syntactically plural NPs are quantificational rather than group denoting. For example:

Few men lifted a piano.

does not seem to have an interpretation where the lifting was collective. I accept Robert's division of determiners into individual denoting (ie either single individuals or groups) and quantificational, with the exception that 'both' and 'neither' seem to me to be quantificational. I will also treat bare plurals and numerically specified bare plurals as group-denoting, when they appear to be existential rather than generic.

Individual denoting noun phrases can have a distributive reading forced, for example by 'each'.

Four boys each lifted four rocks.
Four boys lifted four rocks each.

In both these cases "four boys" seems to have wide scope over "four rocks". This could be represented using Link's D-operator, if the notion of atomicity were built into the model. Since I take atomicity to be relative to a predicate I make the noun phrase ambiguous between a quantificational and a group reading.

Four boys lifted four rocks.

also has a reading where "four boys" has wide scope, which is distinct from the group-group reading discussed above. I treat this in just the same way as the sentences with "each". Roberts refers to this as "implicit each". I am not convinced that this is the best treatment but it does seem that the subject noun phrase can only have wide scope when it is fully distributive; for example there is no reading of the above sentence where the boys lifted eight rocks working in pairs. Furthermore in

Four boys working in pairs lifted four rocks.

there doesn't seem to me to be a reading where "four boys" has wide scope and there are eight rocks. I also do not think that readings where the object noun phrase has wide scope over the subject necessarily exist; they certainly are not normally very plausible.

So the two readings that my grammar produces for

> Four boys lifted four rocks.

are firstly the group-group reading as mentioned above:

$$\exists x[\mathrm{boy}(x) \wedge \mathrm{NU}(\mathrm{boy})(x) = 4 \wedge \exists y[\mathrm{rock}(y) \wedge \mathrm{NU}(\mathrm{rock})(y) = 4 \wedge \mathrm{lift}(x, y)]]$$

and secondly a reading where "four boys" is fully distributive and has wide scope over "four rocks".

$$\begin{aligned}\exists x[\mathrm{boy}(x) \wedge \mathrm{NU}(\mathrm{boy})(x) = 4 \wedge \forall z[\mathrm{boy}(z) \wedge \mathrm{NU}(\mathrm{boy})(z) = 1 \wedge z \sqsubseteq_o x \\ \Rightarrow \exists y[\mathrm{rock}(y) \wedge \mathrm{NU}(\mathrm{rock})(y) = 4 \wedge \mathrm{lift}(z, y)]]]\end{aligned}$$

(Note that this isn't the same as Hasle's fully distributive reading because the distributivity of "four rocks" isn't specified. It is left open whether each boy lifted all the four rocks at once or whether they were lifted individually or in subgroups.) In Chapter 4 I will describe the operation of the grammar in more detail. In Chapter 5 certain extensions are considered, particularly the treatment of group terms and the treatment of adverbials other than "each" which influence distributivity.

Chapter 4

A simplified approach to plurals and mass terms

It would take too long to systematically explain the grammar which I developed; in any case, due to the way it has evolved, it is unnecessarily complex. A representative extract from the grammar and the lexicon is given in the appendix. What I would like to do here is to pick up some of the aspects which I have introduced briefly and informally in the earlier chapters and give more examples of the representations and the types of inference requirement. All the representations given here are the result of parsing with a single version of the grammar. In most cases spurious equalities have been removed using code written by Dick Crouch. Some of the inferences shown have been verified by Mike Gordon using the HOL system, but unfortunately the original project goal of automatic theorem proving has had to be abandoned. Besides the examples specifically chosen to illustrate the points raised earlier I also include some examples for comparison with the earlier report. The axioms that I assume are collected in Section 4.4 for reference.

One limitation of the grammar is that no tense information is included. Despite this I have given many examples in the past tense, simply because they sound more natural. Obviously in these cases the logical form shown is at least incomplete, but I do not think there is any reason why incorporating tense information should affect my basic approach. Another caveat is that in some cases the grammar produces two equivalent logical forms for essentially spurious reasons. I have not had time to fully debug the grammar to prevent such occurrences and I have just ignored them in the examples that follow.

4.1 Quantities

I have modified Krifka's treatment of quantities in order to provide a more general treatment. For example:

> John bought two pints of milk.
> $\exists y[\mathrm{pint}(y) = 2 \wedge \exists x[\mathrm{milk}(x) \wedge y \sqsubseteq_o x] \wedge \mathrm{buy}(\mathrm{john}, y)]$

Krifka would treat "milk" in the sentence above as an Nbar when it is combined with the measure phrase but it seems to me that it is better analysed as a noun phrase so that the phrase "two pints of milk" can be treated in the same way as "two pints of that milk" and so on. The correct translation of 'of' in these cases seems to be in terms of $\sqsubseteq_o$ (compare Link's treatment of partitives). From the information that milk, like all nominal predicates, is fully distributive:

$$\begin{array}{l}\text{Nom(milk)}\\ \forall P[\text{Nom}(P) \Rightarrow \forall x[P(x) \Rightarrow \forall y[y \sqsubseteq_o x \Rightarrow P(y)]]]\end{array}$$

it follows directly that

John bought some milk.
$\exists y[\text{milk}(y) \wedge \text{buy}(\text{john}, y)]$

Plurals are also dealt with by measure functions, but these are dependent on the head noun and I make somewhat different assumptions about their properties (see the next section).

John bought two apples.
$\exists x[\text{apple}(x) \wedge \text{NU}(\text{apple})(x) = 2 \wedge \text{buy}(\text{john}, x)]$

I have tried to treat some other properties connected with quantities. For example:

Milk costs twenty pence per pint.
$$\begin{array}{l}\forall y[\text{milk}(y) \Rightarrow \exists z, x, w[\text{pence}(w) \wedge \text{NU}(\text{pence})(w) = 20 \wedge \\ \qquad\qquad\qquad\qquad \text{per}(w, x) = z \wedge \text{pint}(x) = 1 \wedge \text{cost}(y, z)]]\end{array}$$

This allows the deduction of:

Two pints of milk cost forty pence.
$$\begin{array}{l}\forall y[\text{pint}(y) = 2 \wedge \exists x[\text{milk}(x) \wedge y \sqsubseteq_o x] \wedge \\ \quad \exists z[\text{pence}(z) \wedge \text{NU}(\text{pence})(z) = 40 \wedge \text{cost}(y, z)]]\end{array}$$

if we have the extra information about cost and the meaning of per:

$$\begin{array}{l}\text{cost}(y, z) \wedge z = \text{per}(x, w) \wedge M(y) = n \wedge M(w) = m \wedge P(x) \wedge \\ \text{NU}(P)(x) = p \\ \Rightarrow \exists v[\text{cost}(y, v) \wedge P(v) \wedge \text{NU}(P)(v) = \text{NU}(P)(x).n/m]\end{array}$$
(where M is a measure function and n, m and p are reals).

4.2 Distributivity

As mentioned above all nominal predicates are fully distributive because of the decision to try and use the same predicate for both mass and count senses and to distinguish count senses by the specification of natural units. However, as discussed in Chapter 2, in order to deduce, from:

Some customers are students [1].
$\exists x[\text{customer}(x) \wedge \text{NU}(\text{customer})(x) > 0 \wedge \text{student}(x) \wedge$
$\quad \text{NU}(\text{student})(x) > 0]$

that

A customer is a student.
$\exists x[\text{customer}(x) \wedge \text{NU}(\text{customer})(x) = 1 \wedge \text{student}(x) \wedge$
$\quad \text{NU}(\text{student})(x) = 1]$

we need extra information about the properties of natural units:

If I is a predicate which holds for x iff x is an integer:
$\forall P, x, n[\text{NU}(P)(x) = n \Rightarrow I(n) \wedge n \geq 0]$
(mapping to positive integers)
$\forall P, x, y[\neg x o_o y \Rightarrow \text{NU}(P)(x) + \text{NU}(P)(y) = \text{NU}(P)(x \sqcup_o y)]]$
(additivity)
$\forall P, x[\text{NU}(P)(x) = n \wedge n > 0$
$\quad \Rightarrow \forall m[m \leq n \wedge m > 0 \wedge I(m) \Rightarrow \exists y[y \sqsubseteq_o x \wedge \text{NU}(P)(y) = m]]]$

and also that:

$\text{NU}(\text{student}) = \text{NU}(\text{human}) = \text{NU}(\text{customer})$

As I mentioned in the last chapter one source of distributivity on verbs is the floated quantifier "each" or the implicit each reading. For example:

Four students bought four apples.
$\exists y[\text{student}(y) \wedge \text{NU}(\text{student})(y) = 4 \wedge$
$\quad \exists x[\text{apple}(x) \wedge \text{NU}(\text{apple})(x) = 4 \wedge \text{buy}(y, x)]]$
(the group-group reading)
$\exists y[\text{student}(y) \wedge \text{NU}(\text{student})(y) = 4 \wedge$
$\quad \forall z[\text{student}(z) \wedge \text{NU}(\text{student})(z) = 1 \wedge z \sqsubseteq_o y$
$\qquad \Rightarrow \exists x[\text{apple}(x) \wedge \text{NU}(\text{apple})(x) = 4 \wedge \text{buy}(z, x)]]]$
(the implicit each reading)

has two readings, with the 'implicit each' reading corresponding to the single reading of:

Four students each bought four apples
$\exists y[\text{student}(y) \wedge \text{NU}(\text{student})(y) = 4 \wedge$
$\quad \forall z[\text{student}(z) \wedge \text{NU}(\text{student})(z) = 1 \wedge z \sqsubseteq_o y$
$\qquad \Rightarrow \exists x[\text{apple}(x) \wedge \text{NU}(\text{apple})(x) = 4 \wedge \text{buy}(z, x)]]]$
(actual each)

[1] I have ignored a kind reading for this sentence which is inappropriate here and possibly incorrect. It is equivalent to the desired reading of sentences like "Some wines are clarets."

Rather than use an adverbial modifier such as Link's D-operator I translate plural noun phrases as being ambiguous between a group interpretation where the predicate is applied to the entire group and a distributive interpretation where the predicate is applied to each unit of the group. These units depend on the noun phrase.

In the special case of group terms it is assumed that there will be some information in the underlying knowledge base as to what could constitute an individual member of the group.

The team each received a medal.
$$\begin{aligned}&\exists y[\mathrm{team}(y) \wedge \forall z[\mathrm{team}(z) \Rightarrow z \sqsubseteq_o y] \wedge \\ &\quad \forall x[x \sqsubseteq_o y \wedge \mathrm{IND}(\mathrm{team})(x) \\ &\qquad \Rightarrow \exists z[\mathrm{medal}(z) \wedge \mathrm{NU}(\mathrm{medal})(z) = 1 \wedge \mathrm{receive}(x, z)]]]\end{aligned}$$

In this case

$$\forall x[\mathrm{IND}(\mathrm{team})(x) \Rightarrow \mathrm{NU}(\mathrm{human})(x) = 1]$$

would seem to be a reasonable assumption.

It is necessary to translate 'and' so that it is ambiguous between $\oplus$ and a Boolean reading. The Boolean reading will be used when conjoining certain quantified phrases, for example:

Every customer and every student took two apples.
$$\begin{aligned}&\forall x[\mathrm{customer}(x) \wedge \mathrm{NU}(\mathrm{customer})(x) = 1 \\ &\quad \Rightarrow \exists y[\mathrm{apple}(y) \wedge \mathrm{NU}(\mathrm{apple})(y) = 2 \wedge \mathrm{take}(x, y)]] \wedge \\ &\forall z[\mathrm{student}(z) \wedge \mathrm{NU}(\mathrm{student})(z) = 1 \\ &\quad \Rightarrow \exists y[\mathrm{apple}(y) \wedge \mathrm{NU}(\mathrm{apple})(y) = 2 \wedge \mathrm{take}(z, y)]]\end{aligned}$$

However I also use the Boolean translation to give the implicit or explicit each reading, while the translation as $\oplus$ gives the group reading.

Some customers and some students took two apples.
$$\begin{aligned}&\exists x[\mathrm{customer}(x) \wedge \mathrm{NU}(\mathrm{customer})(x) > 0 \wedge \\ &\quad \forall w[\mathrm{customer}(w) \wedge \mathrm{NU}(\mathrm{customer})(w) = 1 \wedge w \sqsubseteq_o x \\ &\qquad \Rightarrow \exists y[\mathrm{apple}(y) \wedge \mathrm{NU}(\mathrm{apple})(y) = 2 \wedge \mathrm{take}(w, y)]]] \wedge \\ &\exists x'[\mathrm{student}(x') \wedge \mathrm{NU}(\mathrm{student})(x') > 0 \wedge \\ &\quad \forall w'[\mathrm{student}(w') \wedge \mathrm{NU}(\mathrm{student})(w') = 1 \wedge w' \sqsubseteq_o x' \\ &\qquad \Rightarrow \exists y'[\mathrm{apple}(y') \wedge \mathrm{NU}(\mathrm{apple})(y') = 2 \wedge \mathrm{take}(w', y')]]]\end{aligned}$$
(Boolean translation)
$$\begin{aligned}&\exists x[\mathrm{customer}(x) \wedge \mathrm{NU}(\mathrm{customer})(x) > 0 \wedge \\ &\quad \exists y[\mathrm{student}(y) \wedge \mathrm{NU}(\mathrm{student})(y) > 0 \wedge \\ &\qquad \exists z[\mathrm{apple}(z) \wedge \mathrm{NU}(\mathrm{apple})(z) = 2 \wedge \mathrm{take}(x \oplus y, z)]]]\end{aligned}$$
(translation as isum)

In the grammar actually developed the implicit each reading is also given for intransitive verbs although it is subsumed by the group reading, since I could not think of a motivated way of blocking it. So the following readings are obtained:

> John and Mary sleep.
> $\text{sleep}(\text{john} \oplus \text{mary})$
> $\text{sleep}(\text{john}) \wedge \text{sleep}(\text{mary})$
>
> John and Mary convene.
> $\text{convene}(\text{john} \oplus \text{mary})$
> $\text{convene}(\text{john}) \wedge \text{convene}(\text{mary})$
>
> The committees convened.
> $\exists x[\text{committee}(x) \wedge \text{NU}(\text{committee})(x) > 0 \wedge$
> $\quad \forall y[\text{committee}(y) \wedge \text{NU}(\text{committee})(y) > 0 \Rightarrow y \sqsubseteq_o x] \wedge \text{convene}(x)]$
> $\exists x[\text{committee}(x) \wedge \text{NU}(\text{committee})(x) > 0 \wedge$
> $\quad \forall y[\text{committee}(y) \wedge \text{NU}(\text{committee})(y) > 0 \Rightarrow y \sqsubseteq_o x] \wedge$
> $\quad\quad \forall z[\text{committee}(z) \wedge \text{NU}(\text{committee})(z) = 1 \wedge z \sqsubseteq_o x \Rightarrow \text{convene}(z)]]$

The decision to make sentences like

> The students each convened.

grammatical was mentioned earlier. It was stated there the sentence will always be false because convene only takes recognised groups in its extension. In fact I think that there are some peculiar situations where convene may be true of a single individual; for example if all the other members of the committee are ill perhaps:

> ? John convened by himself.

is acceptable. However it would be simple to block this in the grammar, if desired, by use of a semantic plurality feature.

The distributive properties of predicates like 'die', 'gather', 'climb', 'heavy', 'wet' and so on will all arise as consequences of their meaning, as discussed in the last chapter. At least some of their effects with respect to distributivity can be specified in the current context, but what should be investigated is how these effects arise from their meanings. For example:

> $\forall x[\text{die}(x) \Rightarrow \forall y[y \sqsubseteq_o x \wedge \text{NU}(\text{organism})(y) \geq 1 \Rightarrow \text{die}(y)]]$
> $\forall x[\text{wet}(x) \Rightarrow \forall y[y \sqsubseteq_o x \Rightarrow \text{wet}(y)]]$

but these effects should arise because 'die' is a predicate that affects individual organisms and 'wet' affects stuff in general; they should not have to be stated as axioms.

4.3 Kinds and bare noun phrases

As I mentioned in the last chapter I do not treat all bare mass nouns as kinds. So the grammar gives:

Claret is wine.
$\forall y[\text{claret}(y) \Rightarrow \text{wine}(y)]$
$\text{Kind}(\text{claret}) = \text{Kind}(\text{wine})$

The second reading would be appropriate in sentences like:

Bordeaux is claret.

but isn't relevant to the discussion which follows and I will ignore it from now on. Obviously from the first reading, and the appropriate reading of:

Wine is liquid.
$\forall y[\text{wine}(y) \Rightarrow \text{liquid}(y)]$

it follows immediately that:

Claret is liquid.
$\forall y[\text{claret}(y) \Rightarrow \text{liquid}(y)]$

The universal reading is imposed because the verb phrase is stative. The translation for a non-stative verb would be existential:

John bought wine.
$\exists y[\text{wine}(y) \wedge \text{buy}(\text{john}, y)]$

However sentences such as:

Some claret is wine.
$\exists y[\text{claret}(y) \wedge \text{wine}(y)]$

are allowed. Since the grammar is conventional in that the quantifier is assumed to be known when the noun phrase is formed, these results are achieved in a way which is strictly non-compositional. The grammar could be modified to avoid this but the treatment of bare noun phrases suggested in the next chapter avoids the problem.

Bare noun phrases may also be interpreted as kinds. For example:

Claret is a wine.
$\text{Type}(\text{wine})(\text{Kind}(\text{claret})) \wedge \text{NU}(\text{Type}(\text{wine}))(\text{Kind}(\text{claret})) = 1$

Note that we end up with natural units for kinds as well as for objects. However in the kind domain we have the simplification that all natural units are equivalent, since kinds like Kind(claret) are taken to be atomic. So if we also know that:

Wine is a liquid.
Type(liquid)(Kind(wine)) $\wedge$ NU(Type(liquid))(Kind(wine)) = 1

we can deduce:

Claret is a liquid.
Type(liquid)(Kind(claret)) $\wedge$ NU(Type(liquid))(Kind(claret)) = 1

from the theorem:

$$\forall P, Q, R[\text{Type}(P)(\text{Kind}(Q)) \wedge \text{Type}(Q)(\text{Kind}(R)) \\ \Rightarrow \text{Type}(P)(\text{Kind}(R))]$$

Because of the connection between the kind and object domains given by:

$$\forall P, x[\text{Type}(P)(\text{Kind}(Q)) \Rightarrow \forall x[Q(x) \Rightarrow P(x)]]$$

we can also deduce:

Claret is liquid.
$\forall y[\text{claret}(y) \Rightarrow \text{liquid}(y)]$

However from

Claret is a wine.
Wine is liquid.

we can only deduce that

Claret is liquid.

but not

Claret is a liquid.
Claret is a kind of liquid.

or any other relationship with liquid in the kinds domain (contrast Hasle(1988) and Pelletier and Schubert (1986)). This seems reasonable; we have to at least know that Kind(liquid) is part of the kinds domain before we make any such inference, we do not want to deduce sentences like,

* Claret is a kind of wet.

The following examples should further indicate the coverage:

Wine is wine. (two readings, both analytic)
$\forall x[\text{wine}(x) \Rightarrow \text{wine}(x)]$
$\text{Kind}(\text{wine}) = \text{Kind}(\text{wine})$
(this is the form produced before removal of equalities)

All wine is wine.
$\forall x[\text{wine}(x) \Rightarrow \text{wine}(x)]$

Cows are mammals.
$\forall y[\text{cow}(y) \wedge \text{NU}(\text{cow})(y) = 1 \Rightarrow \text{mammal}(y) \wedge \text{NU}(\text{mammal})(y) > 0]$
$\text{Kind}(\text{cow}) = \text{Kind}(\text{mammal})$

Cows are a kind of mammal.
$\text{Type}(\text{mammal})(\text{Kind}(\text{cow})) \wedge \text{NU}(\text{Type}(\text{mammal}))(\text{Kind}(\text{cow})) = 1$

John is teaching students.
$\exists y[\text{student}(y) \wedge \text{NU}(\text{student})(y) > 0 \wedge \text{teach}(\text{john}, y)]$

Cheddar and edam and stilton are cheeses.
$\text{Type}(\text{cheese})((\text{Kind}(\text{cheddar}) \oplus \text{Kind}(\text{edam})) \oplus \text{Kind}(\text{stilton})) \wedge$
$\text{NU}(\text{Type}(\text{cheese}))((\text{Kind}(\text{cheddar}) \oplus \text{Kind}(\text{edam})) \oplus \text{Kind}(\text{stilton})) > 0$

Claret is a kind of wine.
$\text{Type}(\text{wine})(\text{Kind}(\text{claret})) \wedge \text{NU}(\text{Type}(\text{wine}))(\text{Kind}(\text{claret})) = 1$

4.4 Axioms

These are collected here for ease of reference.

4.4.1 Postulates for the semilattice of objects

These are taken from Krifka's paper.

$\forall x, y, z[x \sqcup_o y = z \Rightarrow O(x) \wedge O(y) \wedge O(z)]$ (restriction to O)
$\forall x, y[O(x) \wedge O(y) \Rightarrow \exists z[x \sqcup_o y = z]]$ (completeness)
$\forall x, y[x \sqcup_o y = y \sqcup_o x]$ (commutativity)
$\forall x[x \sqcup_o x = x]$ (idempotency)
$\forall x, y, z[x \sqcup_o [y \sqcup_o z] = [x \sqcup_o y] \sqcup_o z]$ (associativity)
$\forall x, y[x \sqsubseteq_o y \iff x \sqcup_o y = y]$ (part)
$\neg\exists x \forall y[x \sqsubseteq_o y]$ (no 0-element)
$\forall x, y[x \sqsubset_o y \iff x \sqsubseteq_o y \wedge \neg x = y]$ (proper part)
$\forall x, y[x o_o y \iff \exists z[z \sqsubseteq_o x \wedge z \sqsubseteq_o y]]$ (overlap)
$\forall x, y[x \sqsubset_o y \iff \exists x'[\neg x o_o x' \wedge x \sqcup_o x' = y]]$ (complementarity)
$\forall x, P[FU_o(P) = x \iff \forall x'[P(x') \Rightarrow x' \sqsubseteq_o x] \wedge$
$\forall x''[\forall x'[P(x') \Rightarrow x' \sqsubseteq_o x''] \Rightarrow x \sqsubseteq_o x'']]$ (fusion operator)

4.4.2 Properties of measure functions

Again from Krifka's paper:

> R is a predicate which holds for x iff x is a real number.
> AAM_o is defined so that:
> $\forall m[AAM_o(m) \Rightarrow \forall x[R(m(x)) \wedge m(x) \geq 0]]$
> (mapping to positive real numbers)
> $\forall m[AAM_o(m) \Rightarrow \forall x, y[\neg x o_o y \Rightarrow m(x) + m(y) = m(x \sqcup_o y)]]$
> (additivity)
> $\forall m[AAM_o(m) \Rightarrow \forall x, y[\neg m(x) = 0 \wedge y \sqsubseteq_o x$
> $\quad \Rightarrow \exists n[n > 0 \wedge n.m(y) \geq m(x)]]]$
> (Archimedean property)

As mentioned earlier I have modified Krifka's notion of a natural unit so that it only takes integer values, so only the additivity property holds. For natural unit measure functions we have:

> If I is a predicate which holds for x iff x is an integer:
> $\forall P, x, n[\mathrm{NU}(P)(x) = n \Rightarrow I(n) \wedge n \geq 0]$
> (mapping to positive integers)
> $\forall P, x, y[\neg x o_o y \Rightarrow \mathrm{NU}(P)(x) + \mathrm{NU}(P)(y) = \mathrm{NU}(P)(x \sqcup_o y)]]$
> (additivity)
> $\forall P, x[\mathrm{NU}(P)(x) = n \wedge n > 0 \Rightarrow \forall m[m \leq n \wedge m > 0 \wedge I(m)$
> $\quad \Rightarrow \exists y[y \sqsubseteq_o x \wedge \mathrm{NU}(P)(y) = m]]]$

4.4.3 Kinds

The kind domain is assumed to make use of the same operators as those given above for the domain of objects. The further structure imposed by the kind hierarchy is formalised by the following axioms:

$$\forall x, y, z[x <_k y \wedge y <_k z \Rightarrow x <_k z]$$
$$\forall P, x[(\mathrm{Type}(P))(x) \iff x <_k \mathrm{Kind}(P)]$$

The following theorem is that used in all the examples:

$$\forall P, Q, R[\mathrm{Type}(P)(\mathrm{Kind}(Q)) \wedge \mathrm{Type}(Q)(\mathrm{Kind}(R)) \\ \Rightarrow \mathrm{Type}(P)(\mathrm{Kind}(R))]$$

The connection between the domain of kinds and the domain of objects assumed in the examples is given by:

$$\forall P, x[\mathrm{Type}(P)(\mathrm{Kind}(Q)) \Rightarrow \forall x[Q(x) \Rightarrow P(x)]]$$

However I would expect a fuller treatment to modify this.

Chapter 5

Problems

This chapter considers some of the problems which arise for my somewhat simplistic treatment. A general principle which I have tried to follow is that if some gadgetry that complicates a theory does not appear to account for the observations for which it was purportedly introduced, or does not account for clearly related phenomena, then it should be abandoned and the simpler theory investigated before trying to introduce new techniques to account for the original observations[1]. So, although the first two sections of this chapter are fairly concise, the sections on groups and generics are necessarily discursive.

5.1 Adverbial phrases which affect distributivity

The grammar introduced in Chapter 4 provides a treatment of each. A discussion of "all" as a floated quantifier can be found in Roberts(1987). I have mostly ignored the issue of involvement in this report; in the context of my current grammar I would suggest that it is possible to treat it in terms of the interpretation of the group-group reading. The most plausible treatment of "all" would be to regard it as an adverbial modifier which stipulates full involvement, but I will not discuss the mechanics of this here, because a motivated treatment of involvement is outside the scope of this work.

I would also regard "together", "in pairs" and so on as adverbial modifiers but these restrict the possibilities for the groups which will be involved in individual events. For example, my earlier interpretation of lift(x,y) in terms of a predicate ${}^{C'}$lift which was true of an individual lifting event was:

$$\begin{aligned} \text{lift}(x,y) \;\;\Rightarrow\;\; & x = FU_o(\lambda z[\exists u[{}^{C'}\text{lift}(z,u) \wedge z \sqsubseteq_o x \wedge u \sqsubseteq_o y]]) \wedge \\ & y = FU_o(\lambda v[\exists w[{}^{C'}\text{lift}(w,v) \wedge v \sqsubseteq_o y \wedge w \sqsubseteq_o x]]) \end{aligned}$$

If "together" was applied to the subject for example,

Four boys together lifted four rocks.

then we could add the condition that the subject in the lifting events has to be the entire group.

[1]The difficulty of course is that my techniques for removing parts of theories are probably more analogous to wielding a blunt instrument than Occam's razor.

5.2 The count/mass distinction

The previous chapter describes a grammar where, following Pelletier, there is no distinction between the mass and count senses of a noun. For example 'apple' contained in its extension apple stuff as well as individual apples; the burden of the distinction between the two was carried by the natural unit specification in the count case. (Pelletier does not actually distinguish between the two at all.) As mentioned earlier, this is not completely satisfactory, although it is actually surprising how far this treatment can be taken. The first problem is that 'the apple' would always be translated as:

$$\exists x[\text{apple}(x) \wedge \forall y[\text{apple}(y) \Rightarrow y \sqsubseteq_o x]]$$

which is too general because as stated it could refer to more than one individual apple. This could be fixed; a more general problem is that the NU device is not really suited to carrying all the burden of quantisation, especially since we have to assume equivalence of natural units for living organisms, for example. It is also difficult to find suitable representations for nouns like 'puddle' and 'cube' which do not carry any real connotations about matter as opposed to physical shape.

I would therefore prefer to adopt an approach where count and mass senses of predicates are distinguished. I have tested the following proposal on a small grammar but I have not had time to integrate it into my full grammar. There is no need to give separate lexical entries for the two senses of words; instead predicate modifiers can be employed to force the count sense in the same way as the natural unit specification is currently introduced. Here I will use the notation ${}^{Count}P$ to represent the count sense of P. We have the following axiom:

$$\forall x, P[{}^{Count}P(x) \Rightarrow P(x)]$$

By having a predicate ${}^{Count}P$ which is true of single individuals of P a definition of natural units can be provided.

$$\forall x, P[NU(P)(x) = |\{y : y \sqsubseteq_o x \wedge {}^{Count}P(y)\}|]$$

The value of $NU(P)(x)$ is equal to the cardinality of the set of all entities which are material parts of x and which are also ${}^{Count}P$. Plurals will be translated as ${}^{\star Count}P$ reintroducing the use of a closure operator.

Ordinary count predicates will have a mass counterpart which will be more or less difficult to refer to due to the factors mentioned in Chapter 2. However I do not think there is any point trying to treat the relative difficulty during the parsing process. Predicates such as 'cube' also technically are regarded as having a mass sense but this will only denote matter which is currently in the shape of a cube, and therefore no sentence which results from the application of the Universal Grinder such as, There was cube all over the floor.

can ever be true. Mass nouns will also have a count sense which can be assumed to be accessed by certain partitive constructions. Mass nouns with obvious units, such as furniture, will have these units as the extension of their count sense; unless this is specified by reference to "an item of furniture" for example, any reference to furniture will be to its mass sense, but as mentioned in Chapter 2 there will be a strong default that any reference to an object with obvious units implies that the reference is to some number of whole units. The translation of phrases like "the man" in terms of a matter sense alone could then be retained and this default relied upon to force the count interpretation except in very strange contexts.

5.3 Groups and group terms

In this section I go back to the discussion of groups which I began in Chapter 2 and attempt to show that group terms may be represented as non-atomic individuals, and that a separate group level in the semantics is not needed for a purely extensional treatment.

In the grammar in Chapter 4 group nouns such as 'team' refer to individuals in the same way as any other nouns. These individuals are not atomic; we need to be able to refer to their iparts.

The team each received a medal[2].

$$\begin{aligned}&\exists y[\mathrm{team}(y) \wedge \forall z[\mathrm{team}(z) \Rightarrow z \sqsubseteq_o y] \wedge \\ &\quad \forall x[x \sqsubseteq_o y \wedge \mathrm{IND}(\mathrm{team})(x) \\ &\qquad \Rightarrow \exists z[\mathrm{medal}(z) \wedge \mathrm{NU}(\mathrm{medal})(z) = 1 \wedge \mathrm{receive}(x, z)]]]\end{aligned}$$

where $\forall x[\mathrm{IND}(\mathrm{team})(x) \Rightarrow \mathrm{NU}(\mathrm{human})(x) = 1]$ is assumed to be derivable from information in the knowledge base.

Landman produces an argument to suggest that group terms such as committee cannot be isums of their members. Paraphrased, this runs:

> Consider a situation where John is a member of committee A who is under threat of expulsion. Committee A holds a meeting to decide on John's expulsion without inviting John. In this situation both the following are true:
>
> John and Committee A do not meet in room A.
> Committee A meets in room A.
>
> But "John and Committee A" is "John $\oplus$ Committee A" which is equal to "Committee A"

[2]Roberts(1987) finds this sort of sentence dubious, but allows "each" following the direct object. Both seem perfectly acceptable to me.

Note, however, that we get exactly the same problem if John is a fellow and we replace "Committee A" by "the fellows" in the sentences above. The same argument then leads to the conclusion that plurals cannot be treated as isums.

It seems to me that the problem really is that in the current context John is *not* a member of Committee A. In the situation described it seems possible to say:

All the members of Committee A voted for John's expulsion.

even though John didn't vote for his own expulsion. I shall consider this further below, and argue that it may be part of the general problem of non-extensionality discussed by Landman, but first I want to review the desirability of introducing a separate level of groups as is done in Link(1984) and Landman(1987). I will mainly consider Landman's notion of groups rather than Link's because Link introduces groups to treat a rather limited class of sentences, and does not discuss their integration with the rest of the theory in detail.

The problems caused by the introduction of groups, particularly the type proliferation, were briefly mentioned in Chapter 3. There is also the problem of deciding what constitutes a group; in Landman's interpretation this comes very close to the problem of disambiguating down to individual events that was also discussed earlier. So the arguments for groups are of considerable importance, because to introduce them greatly complicates the theory. But all the examples on which Landman and Link rely to show the necessity for a notion of a group which is distinct from an isum are very similar, in that they involve noun phrases conjoined by 'and' and verb phrases which are either reciprocals or seem to me to be very similar.

In fact Landman introduces three arguments for postulating a separate group level, but does not regard the first two as conclusive, so I will just mention them briefly here. The first is 'upward closure', even if all the members of Committee A are judges the sentence:

Committee A are judges.

seems peculiar. I think that this is may just be a syntactic agreement problem; the sentence seems slightly odd to me but not untrue. Similar sentences seem perfectly acceptable:

That group of men are judges.

The second argument concerns involvement but, as Landman concedes, there are ways of treating this without involving groups. It isn't clear to me that the introduction of groups really helps.

Landman's main argument concerns the sentence:

The cards below seven and the cards from seven up are separated.

If we treat 'and' as iplus then this sentence becomes equivalent to, for example,

The cards below ten and the cards from ten up are separated

because the isum refers to all the cards. On the other hand we cannot simply treat 'and' as Boolean conjunction because that would give us:

The cards below seven are separated and the cards from seven up are separated.

which is not the reading we want (although it is a possible reading).

Some of Link's examples run along similar lines.

The Leitches and the Latches hate each other.

has an interpretation where the hate applies to the relationship between the two groups, not to all the individual relationships or to the relationships internal to each group[3]. Again we can get the second and the third interpretations by considering 'hate each other' applied to either the isum of the two groups or to each group separately. But in both these examples we can provide a reading which seems to correspond to that required without invoking a separate group level.

$$\text{separated}(\sigma x(x < 7), \sigma y(y \geq 7)) \wedge \text{separated}(\sigma x(x \geq 7), \sigma y(y < 7))$$
$$\text{hate}(\sigma x \text{Leitch}(x), \sigma y \text{Latch}(y)) \wedge \text{hate}(\sigma x \text{Latch}(x), \sigma y \text{Leitch}(y))$$

So in both cases the motivation for the group reading is just to keep the two noun phases distinct within the conjunction.

Now it seems to me that verbs phrases like 'separated', 'divided' and so on behave very similarly to reciprocals. The reading that we are interested in for the sentence:

The cards below seven and the cards from seven up are separated.

is the same as the single reading for:

The cards below seven and the cards from seven up are separated from each other.

[3]Link denies the existence of this last reading but Landman admits it and it seems plausible to me, especially in context.

But it seems that all the sentences that Landman produces to support the claim for a separate group level are of this type. It seems plausible that a syntactic solution might be found which allows the representations given above to be derived.

One device that might be used is to interpret 'and' initially in such a way as to preserve the distinction between the two noun phrases and their ordering. When an ordinary verb phrase was applied to a noun phrase formed in such a way, the interpretation would be equivalent to isum. If a reciprocal or pseudo-reciprocal verb phrase was applied the conjoined noun phrase could optionally be interpreted as providing two arguments to the verb phrase. This may be a dubious move, but something of the sort could also be used in order to interpret:

> Sandy and Kim are husband and wife.

Link mentions that some notion of ordered pair may be needed in addition to the symmetrical isum operator (Link 1983) (he sketches a possible treatment in Link 1984).

Link also uses groups to give a representation for the type of construction which he refers to as *hydra*. For example,

> The landlords and the tenants who fight against each other will
> all come together tomorrow in an attempt to find a compromise.

The device mentioned above will cope with these constructions if I make the same assumptions about the syntax as Link does (which essentially involves treating hydra as a peculiar case of Nbar conjunction).

It seems to me that some version of this technique is the only modification to Link's original treatment of 'and' that is actually needed for the treatment of these examples, and causes far fewer complications than the introduction of a separate group level into the logic.

Landman mentions that Link's notion of groups, unlike his own, allows some non-extensional aspects to be represented. The committees problem is an example; if committee A necessarily has the same members as committee B then it is still possible that:

> Committee A paid an official visit to South Africa.

is true, but that:

> Committee B paid an official visit to South Africa.

is false[4]. Landman argues however that exactly the same sort of problem can

[4]In fact Landman's discussion is slightly misleading; Link does not seem to have intended any such use for the notion of group as introduced in Link(1984). Committees are treated as atomic entities, and there is no reason for committees with the same members to be equated.

arise with non-group denoting terms. For example, suppose that the judge is also the hangman. Then

> The judge is on strike.

does not imply

> The hangman is on strike.

(Landman, 1987, p 83). The individual concerned is on strike in his capacity as judge but this doesn't necessarily imply that he is on strike in his capacity as hangman. But the committees problem seems very similar; another of Landman's examples is if the chairman of committee A is the chairman of committee B then it is still possible that:

> The chairman of committee A paid an official visit to South Africa.

is true, but that:

> The chairman of committee B paid an official visit to South Africa.

is false. "Committee" seems to give rise to these problems because of its connotations with respect to officialdom, not because it is a group denoting term. If this sort of problem is not specific to group denoting nouns there is no point in providing a solution for groups alone.

I would claim that the example of John and Committee A discussed earlier is also similar; the problem still seems to arise if we explicitly talk about membership; the only solution therefore is to say that *in this context* we cannot substitute any sort of representation of the individuals who are formally members for "Committee A" or even for "the members of Committee A".

For further discussion of extensionality, intensionality and intentionality, see Landman; the only point that I am trying to make here is that since such examples seem to require that we either invoke non-extensional entities or non-extensional properties for true individuals, plurals and group terms alike, there is no advantage in introducing a treatment for groups alone.

So the conclusions from this section are that the device of a separate group level which complicates the theory so considerably, may be avoidable and that, in purely extensional contexts, groups may be identified with the isums of their members.

5.4 Generic Sentences

The earlier chapters introduced Pelletier and Schubert's treatment of generics and the somewhat unprincipled treatment that I adopted in my grammar, which was designed to show that a simplistic treatment can account for a significant range of sentences. In this section I suggest how my approach could be improved, without introducing the complications of theories such as those of G. Carlson, Chierchia and Pelletier and Schubert themselves. Much of this draws on the discussion in Pelletier and Schubert (1988) but the conclusions I draw are somewhat different.

The first question is what constitutes a generic sentence? There doesn't actually seem to be any very good answer to this. Pelletier and Schubert refer to the 'fundamental intuition' that generic sentences refer to kinds and non-generic sentences refer to instances of kinds. They give a nice sequence of examples which conclusively show that generic sentences can involve any proportion of the entities which may be (eventually) introduced by a 'generic' noun phrase.

Snakes are reptiles.	All snakes.
Telephone books are thick books.	Most telephone books.
Guppies give live birth.	All of the females which give birth.
Italians are good skiers.	Some proportion which is greater than that for other countries.
Frenchmen eat horsemeat.	Possibly very few.

Later on they introduce a formal distinction on the basis of necessary and contingent facts, but they accept that this does not account for sentences such as the last two of the series. They show that sentences like:

> Wildlife is being destroyed.

may involve different extents of generality; this sentence could refer to a specific, very limited, spatio-temporal location or be completely general, or to anything in between, depending on context.

The problems that arise from Pelletier and Schubert's attempts to treat all bare mass nouns as kinds have been discussed in Chapter 2. Although the paper that I am considering here makes the justification for doing this much clearer, the problems of retrieving the non-kind readings (and possibly restricting the kind readings) still remain. Intuitively I disagree with their assumption that most predicates can be true of both kinds and objects, and their notion of kinds seems far too abstract. As stated earlier, I am assuming a much more restricted view of kinds and I try to only introduce them in cases where the name of a substance or a class of objects seems to be the only possible interpretation of the bare noun phrase. I do not think any of the sentences listed above refer to kinds in this sense.

So my 'fundamental intuition' is somewhat different; I think that statements are all generic to differing extents: the extent to which a statement is generic is the extent of the occasions in which it may validly be used in making inferences.

Wine is drunk during religious ceremonies.

does not seem to me to necessarily mean anything more restrictive than:

Some wine is drunk during some religious ceremonies.

but it still indicates a general fact about the world, not about a specific occasion. Habitual sentences, like:

John drinks wine.

indicate a general property of a particular entity and are thus, in some sense, less generic.

John drank wine on Tuesday.

is a statement which is still further limited and is therefore less likely to be applicable to an arbitrary situation, even one involving John. It seems to me that the genericity of sentences in which there is no specific quantification is not directly a property of the logical form of the isolated sentence. Obviously the closer the approach to universal quantification the more certain an inference about a particular entity may be. If a statement obviously refers to a limited spatio-temporal location, it is unlikely to be as useful in making predictions as one which doesn't limit its applicability in that way. If a statement is about a necessary property rather than a contingent one a similar effect will apply.

I think this viewpoint about genericity, as a primarily pragmatic phenomenon connected with inference, is supported by the observation that some of the "generic" sentences discussed above appear to be similar to the sentences that workers in non-monotonic logic feel that they should represent. The motivation for this is not necessarily representation of natural language, but rather representation of information that must be used to make inferences which seem not only plausible, but necessary for an agent to function in the world. This indicates that the correct representation of generic sentences may only be possible under the assumption of an underlying non-monotonic inference system.

Most work on generic sentences has concentrated on bare noun phrases. The problem is that any interpretation in terms of quantification is highly context dependent. As we have just seen bare noun phrase sentences can refer to any non-zero proportion of entities (if we confine our attention to real entities at least). There appears to be no way of reliably quantifying such sentences other than existentially during the sort of parsing process assumed here.

Consider the utterance of

> Frenchmen eat horsemeat.

Now I might say this in response to:

> Surely nobody eats horses!

even if I believe only a small proportion of Frenchmen eat them. This is perhaps more generic with respect to horsemeat than with respect to Frenchmen. But if, having the same belief, I make this statement in response to:

> What do you think Pierre would like for dinner?

I am being at least highly misleading. I would go so far as to say that the sentence would be false in this context, because it would be interpreted as a statement about Pierre's likely dietary preferences.

Because of this I think an attempt, within the current context, to provide a treatment of generics which will validly allow inferences like:

> Snow is white. This is snow. $\Rightarrow$ This is white

is doomed to failure. We cannot say, as Pelletier and Schubert attempt to do by use of meaning postulates, that predicates such as white are always necessary properties of (contextually dependent) most of the entities involved.

> Poodles are white.

is a perfectly reasonable utterance even though 'whiteness' is not a necessary property of poodles, and under certain circumstances it can be reasonable even if fewer than half of all poodles are white.

There are some predicates which always seem to induce a universal quantification however.

> Dogs are mammals

does imply that

> All dogs are mammals.

But considering the meaning of mammal (and, to some extent, dog) universality can be deduced from the information that "some dogs are mammals". If I know that some member of a species is a member of a biological class, then I know that all members of that species are members of that class, and of course the same effect will be true for less formal taxonomies. I think that this may be formally explicable in terms of interaction with the hierarchy of kinds. An assertion like “dogs are mammals”, where the kind mammal does exist in the hierarchy of kinds, might lead to the assertion that the kind dog is a subkind of the kind mammal. This does not mean that the sentence should be interpreted initially as being about kinds however, because exactly the same result has to be achieved from interpreting “all dogs are mammals”, where even Pelletier and Schubert would accept that the initial translation is in terms of objects rather than kinds.

So far I have been careful to only cite examples which can clearly be true of ordinary objects and not just ‘kinds’. There is clearly a class of predicates which can only be true of kinds, as distinct from objects, for example “has atomic number 79” or “was invented by Fleming”. I discussed this in Chapter 2 and suggested that in most cases we could regard kind predicates and non-kind predicates as distinct. Pelletier and Schubert’s argument against this is the acceptability of some sentences which conjoin kind and non-kind predicates. In some of their examples I would argue that both predicates are non-kind. In general I think that sentences where kind and non-kind predicates are conjoined may not be fully acceptable:

? Claret is a type of wine and has been spilt all over this carpet.

Some of Pelletier and Schubert’s examples do sound better, but it seems possible to treat such sentences by type-raising, if necessary, by making the ‘kind’ predicates take appropriate objects (the supremum of the semilattice, see below) and transform them to kinds (or even to copy one of Pelletier and Schubert’s tricks and give a disjunctive expression for the semantics of the noun phrase).

So there are some cases for which we have to introduce kinds at some level, but it seems counterintuitive to introduce kinds for all bare mass nouns and plurals, as Pelletier and Schubert do. Some of the characteristics which we have to account for specially if we introduce kinds, are paralleled in the cases which cannot involve kinds. Consider:

Storks have a favourite nesting area.
The storks in the park have a favourite nesting area.

We do not need to treat the bare plural in the first sentence as a kind; we can perfectly well treat it as the plural individual comprising all the individual storks (ie as the supremum. In fact this would be exactly the same as ‘the storks’ but obviously the way in which the bare plural may be limited by context is different.) The problem of making sure that we can get the different scopes for the first sentence, which is discussed at length in Pelletier and Schubert’s paper, seems to me exactly the same as that of giving two scopes for the second sentence (the “implicit each” reading, which has been described in previous chapters).

Furthermore we also see different degrees of involvement with ordinary plurals.

> The reporters asked questions after the press conference.

can be true if only a few of the reporters asked questions. The extent of the similarity between these cases and those of the bare noun phrases is not obvious to me; but in both cases context seems to be important and there do not seem to be any consistent statements that can be made based on information that could be available during the type of parsing process that I (and Pelletier and Schubert) assume.

So my tentative modification to the grammar presented in the last chapter is to translate a bare noun phrase, for example "storks" as:

$$\lambda P \exists x \mathrm{stork}(x) \wedge \mathrm{SIGP}(x, \mathrm{stork}) \wedge P(x)$$

where $\mathrm{SIGP}(x, Q)$ means x is a contextually determined significant portion of the supremum of Q. For example if P were "are nesting on my chimney" then $\mathrm{SIGP}(x, \mathrm{stork})$ would be true for any number of storks, so the effect would just be of ordinary existential quantification, whereas "build their nests on chimneys" would have to be true of most of the nest-building part of the population, if the statement were said in the context of giving information about storks. (If the context was information about chimneys, as in an answer to

> What is that big pile of sticks on the chimney? It looks like a nest.

then the statement could be true even if most stork nests are built somewhere else.)

> Storks are birds.
> $\exists x \mathrm{stork}(x) \wedge \mathrm{SIGP}(x, \mathrm{stork}) \wedge \mathrm{bird}(x)$

would require that x was equal to the supremum of stork (this is determined by the taxonomic status of 'bird', not by meaning postulates on the individual word).

The modification needed to the grammar presented in the previous chapter is rather small; instead of giving the ambiguous universally/existentially quantified readings of bare noun phrases there is a reading which refers to a contextually dependent proportion of the supremum. It may be possible to make kind predicates take this as an argument and transform it into a kind, otherwise there will also be a kind reading of the bare noun phrase. Obviously a lot more work on the semantics is needed before this could become a reasonable theory. The SIGP device is obviously not well motivated, but is no worse than Pelletier and Schubert's "contextually dependent most". The relationship between kinds and objects is not clear (again it is not clear in Pelletier and Schubert's work either). (Accounts of kinds based on property theory currently seem to be the most promising.)

Chapter 6

Conclusions

The coverage of my main grammar has been illustrated in Chapter 4. It includes all the examples discussed in Hasle's report, while integrating the treatment of mass nouns and plurals. My grammar includes quantity terms, and the treatment of distributivity covers a much wider range than the earlier work did. Although the work on group terms and generics is very sketchy, I think it serves to indicate that the grammar might be extended fairly naturally. Some of the extensions mentioned in Chapter 5 have been tested on smaller grammar fragments; since the main grammar evolved rather than being planned it would be preferable to rewrite it from scratch if these extensions were incorporated.

Although the logical forms illustrated in Chapter 4 were all produced automatically, in general the proofs outlined have not been. Mike Gordon has written tactics in HOL to perform some proofs. In the case of the measure functions in particular, further work would have to be done on axiomatisation before all the desirable inferences could be produced. In other cases, where I have indicated inferences in the text, I have checked by hand that they do follow from the axiomatisation, but I have not included formal proofs because they did not seem especially revealing.

In many ways my grammar puts more pressure on the inference system than Hasle's did. I hope I have justified this adequately. In order to achieve some inferences I have assumed certain assertions existed in an underlying knowledge base; something which Hasle went to great lengths to avoid. Again I have tried to indicate in such contexts the need for such information and why I believe that it is non-lexical; but to avoid it altogether seems an entirely artificial restriction.

My grammar is an amalgamation of ideas from several different sources. I would have greatly preferred to adopt a single theory and to attempt to implement it but I think I have shown why neither of the theories originally considered were entirely suitable. What I have been attempting to do is to simplify existing theories. The main grammar and axiomatisation which I have presented are inadequate in ways which I have mentioned and no doubt in other ways as well. What I have tried to achieve is a simple flawed theory rather than a complex flawed theory. My hope is that by looking at a wide range of aspects of failure of a simple account before we introduce new gadgets, we will be able to get a significant increase in coverage for each addition.

From this viewpoint there are a few observations that I think are reasonable. There seems to be a general tendency to regard singular count nouns as the 'standard' case for formal linguistic theories. I think that we end up with more robust theories if singular nouns and verbs are regarded as a special case of plurals, where the number happens to be known. If singular noun phrases are considered we tend to be able to interpret clauses in terms of singular events but when plural nouns are introduced disambiguation to individual events may be impossible. Furthermore we cannot start off by treating count nouns in a formalism which requires atomicity and then expect it to be possible to remove that property when we come to deal with mass terms. It seems to me better to regard atomicity as a property which can be added, to a semilattice or to an ensemble, for example.

In general in this theory I identify entities with their constituents; it seems a reasonable basic notion of an entity and gives us the correct inferences about material properties. Obviously this does break down and different aspects of this breakdown can be seen in two examples considered in the body of this report:

> This ring is new but the gold of which it is made is old.
> Committee A and Committee B necessarily have the same members, but Committee A paid an official visit to South Africa and Committee B did not.

It is possible that introduction of a more general notion of constituency than Link's would allow such sentences to be treated, while still allowing the components of an entity to be accessed. As I mentioned earlier this would have to apply to plurals as well as to group terms. However there are also Landman's examples:

> The judge is necessarily the hangman, the judge is on strike but the hangman is not.
> The chairman of committee A is necessarily the chairman of committee B. The chairman of committee A paid an official visit to South Africa but the chairman of committee B did not.

Landman uses this to argue that there is a general problem of non-substitutability, which we cannot expect a purely extensional theory to deal with, but which don't seem to be covered by the standard treatment of intensionality either. If Landman's ideas are accepted, and predicates can be restricted so that they are true only of particular aspects of entities, it seems reasonable to expect that the constituency examples mentioned above would be covered.

So we may as well equate entities with their constituents, since this gives us the correct inferences in a large proportion of cases and does not preclude a more sophisticated treatment, along the lines suggested above.

References

Bach, E.(1986) 'The Algebra of Events', *Linguistics and Philosophy, vol.9,* 5–16

Bunt, H.(1985) *Mass Terms and Model-theoretic semantics,* Cambridge Studies in Linguistics 42, Cambridge University Press

Dowty, R. and Brodie, B.(1984) 'The Semantics of "Floated" Quantifiers in a Transformationless Grammar', *Proceedings of the 3rd West Coast Conference on Formal Linguistics,* Stanford University, pp.75–90

Gordon, M.J.C.(1987) 'HOL: A Proof Generating System for Higher-Order Logic' in G. Birtwistle and P.A. Subrahmanyam (eds.), *VLSI Specification, Verification and Synthesis,* Kluwer

Gould, S.J.(1987) 'A Most Ingenious Paradox' in *The Flamingo's Smile,* Penguin

Hasle, P.(1988) *Mass Terms and Plurals: from Linguistic Theory to Natural Language Processing,* Technical Report 137, Computer Laboratory, University of Cambridge

Krifka, M.(1986) *Nominalreferenz und Zeitkonstitution. Zur Semantik von Massentermen, Pluraltermen und Aktionsarten,* Doctoral dissertation, University of Munich

Krifka, M.(1987) 'Nominal Reference and Temporal Constitution: Towards a Semantics of Quantity', *Proceedings of the 6th Amsterdam Colloquium,* University of Amsterdam, pp.153–173

Landman(1987) *Groups,* Unpublished paper, Department of Linguistics, University of Massachusetts, Amherst

Link, G.(1983) 'The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach' in Bäuerle, Schwarze and von Stechow (eds.), *Meaning, Use and Interpretation of Language,* de Gruyter, Berlin, pp.302–323

Link, G.(1984) 'Hydras: On the Logic of Relative Clause Constructions with Multiple Heads' in Landman and Veltman (eds.), *Varieties of Formal Semantics, GRASS 3,* Foris, Dordrecht, pp.245–257

Link, G.(1986) 'The Logic of Plurals, LP: Review of the Basic Ideas.' in Gärdenfors (eds.), *Generalized Quantifiers: Studies in Linguistics and Philosophy.,* Reidel, Dordrecht, 1988

Link, G.(1984 (draft), forthcoming) 'Plural' in Wunderlich and von Stechow (eds.), *Handbook of Semantics*

Link, G.(1987) 'Algebraic Semantics of Event Structures', *Proceedings of the 6th Amsterdam Colloquium,* University of Amsterdam, pp.243–262

Pelletier, F.J., and Schubert, L. K.(1986,forthcoming) 'Mass Expressions.' in Gabbay and Guenthner (eds.), *Handbook of Philosophical Logic, Vol 4,* Reidel, Dordrecht

Pelletier, F.J., and Schubert, L. K.(1988) 'Problems in the Representation of the Logical Form of Generics, Plurals, and Mass Nouns.' in LePore (eds.), *New Directions in Semantics,* Academic Press, London, pp.385–451

Pulman, S.G.(forthcoming) *Computational Linguistics,* Cambridge University Press

Roberts, C.(1987) *Modal Subordination, Anaphora, and Distributivity,* Doctoral Dissertation, Department of Linguistics, University of Massachusetts, Amherst

Scha, R.J.H.(1983) *Logical Foundations for Question Answering,* PhD thesis

Appendix A

A simplified version of the grammar fragment

The following is an extract from the grammar described in Chapter 4. I have simplified it to illustrate more clearly the parts which are relevant to the current project. I have therefore removed some grammar rules and also removed references to features such as "gap" which do not interact in any interesting way with the treatment of mass nouns and plurals. Most of these features were taken directly from the grammar written by Steve Pulman on which my fragment is based. As mentioned in the main text, my grammar evolved rather than being designed, and contains several infelicities, some of which are pointed out below. The listing below consists of three parts; the first is a list of the relevant semantic features with the possible values they can take and their default value, the second part is the list of grammar rules, and the third is a representative extract from the lexicon.

A.1 Features

`agr[] DEFAULT Xx`
Syntactic plurality

`sempl[+,-] DEFAULT Xx`
Semantic plurality — set for plurals (except those like "scissors") and for group nouns (like "team").

`count[+,-] DEFAULT Xx`
Set for count senses — most singular nouns are regarded as potentially being either mass or count.

`cum[+,-] DEFAULT Xx`
Cumulativity — set for mass senses and plurals.

`measure[+,-] DEFAULT -`
eg "litre".

`measured[+,-] DEFAULT -`
eg "a litre of water"

`dist[+,-] DEFAULT Xx`

This is set when a noun phrase is given a distributive reading, in order to allow what I have referred to as the "implicit each" reading.

`distq[+,-] DEFAULT Xx`
This is set for a quantificational determiner such as "all" and NPs containing such a determiner.

`num[+,-] DEFAULT Xx`
Only relevant to determiners, it is used to indicate whether the determiner can occur with a number. (Possibly only relevant to "one", contrast "* each one man" with "the one man").

`comp[+,-] DEFAULT Xx`
This is simply used when conjoining NPs to avoid multiple analyses for phrases like "John and Mary and Bill".

`kind[+,-] DEFAULT -`
This is "+" for Nbars and NPs which are interpreted as kinds.

`takesk[+, -] DEFAULT -`
Verbs and verb phrases are regarded as being specified (normally lexically) as taking either kinds or ordinary objects.

`kindsubj[+, -] DEFAULT -`
As above, but specifies whether the subject must or must not be a kind.

`ambgen[+, -] DEFAULT Xx`
This stands for ambiguously generic and, as mentioned in the text, is not very well motivated. It is only set for bare noun phrases — those that are universally quantified are "+" — those which are existentially quantified are "-".

`prefgen[+, -] DEFAULT Xx`
Distinguishes statives from non-statives and must have the same value as "ambgen" in rules where an NP subject is combined with a VP.
See `S->NP+VP`, `S->NP+VP+dist` and `S->NP+fquant+VP`.

`nform[-, knd] DEFAULT -`
this takes the value "knd" only for nouns such as "type" and "kind" which are regarded as turning a non-kind noun into a kind.
See the rule `Nbar->kind+of+Nbar`.

A.2 Grammar Rules

```
Rule% S->NP+VP
S[agr X1] -->
     NP[agr X1, kind X4, ambgen X5, dist X6, distq X6]
     VP[agr X1, takesk X4, prefgen X5]

Semantics%  (NP VP)
```

This will give the "group-group" reading for noun phrases which have group-denoting determiners and the quantificational reading for noun phrases which have quantificational determiners. In the first case both dist and distq will be `"-"`, in the second they will both be `"+"`. The features `"kind"` and `"takesk"` ensure that in sentences such as "Claret is a wine", "Claret is scarce" and "Claret is a kind of liquid", "claret" unambiguously denotes a wine. With stative verb phrases which do not take kinds, such as "is liquid" the generic reading would be selected, because `"prefgen"` will be `"+"`. With other verbs phrases such as "is dripping on the carpet" the reading corresponding to "some claret" would be selected. Bare plurals will behave in a parallel way.

```
Rule% S->NP+VP+dist
S[agr plur] -->
     NP[agr plur, kind X4, ambgen X5, dist +, distq -]
     VP[agr plur, takesk X4, prefgen X5]

Semantics%  (NP VP)
```

This gives the "implicit each" reading. It is restricted to sentences where the NP is ambiguously distributive (ie the distributivity does not arise from the quantifier). So this gives the reading of "four boys lifted four rocks" where "four boys" has wide scope over "four rocks". There is no such reading for "the team lifted four rocks" and so the rule is restricted to syntactically plural NPs.

```
Rule% S->NP+fquant+VP
S[agr X1] -->
     NP[agr X1, kind X4, ambgen X5, dist X6, distq X8, sempl +]
     fquant[dist X6, distq X8]
     VP[agr X1, takesk X4, prefgen X5]

Semantics%  (NP VP)
```

This is a very crude rule for floated quantifiers, which just means "each" as far as this fragment is concerned. Because "each" as a floated quantifier is `[ dist + distq - ]` this acts in a very similar way to the previous rule. However it allows: "the team each received four pounds" because it refers to the semantic plurality feature and not the syntactic one.

```
Rule% NP->NP+termjoin+NP
NP[agr plur, sempl +, comp +, kind X4, ambgen X1, dist -, distq -]
  --> NP[kind X4, ambgen X1, dist -, distq -] termjoin
      NP[comp -, kind X4, ambgen X1, dist -, distq -]

Semantics% ((termjoin NP2) NP3)
```

This is for conjunction to form isums. (See the lexical entry for "and"). It is only possible for NPs with a group interpretation. It can apply to kinds as well as non-kind NPs. "comp" is simply to avoid multiple analyses for phrases like "John and Mary and Bill".

```
Rule% NP->NP+conj+NP
NP[agr X3, sempl +, comp +, kind X4, ambgen X1, dist +, distq X7]
  --> NP[kind X4, ambgen X1, dist +, distq X7] conj
      NP[comp -, kind X4, ambgen X1, dist +, distq X7]

Semantics% ((conj NP2) NP3)
```

This is for intersective or Boolean conjunction, eg "Every man and every woman slept". (See the lexical entry for "and"). It is also used to conjoin distributive NPs which give the implicit or explicit each sentences. "The men and the women (each) lifted a rock". It seems that the result of conjoining two singular NPs can have singular or plural agreement. This rule can only be applied to distributive conjuncts, but this may be too restrictive. Also it doesn't cope with some combinations of NPs which it should allow, eg "every woman and one man".

```
Rule% NP->Det+Nbar
NP[agr X, kind X4, ambgen X7, dist X5, distq X2, sempl X8] -->
      Det[agr X, count X1, num X3, dist X5, cum X6, distq X2]
      Nbar[agr X, count X1, kind X4, cum X6, sempl X8]

Semantics%
 (and (and (count Nbar +) (agr Nbar plur)) (num Det +))
                  => (Det (L (_x)
                             (and (Nbar _x)
                                  (greater ((NU Nbar) _x) zero))))
                  => (Det Nbar)
```

This is the simplest of the four rules for introduction of determiners. The NP will be distributive iff the determiner is quantificational rather than group denoting. The specification of plurality in terms of natural units occurs here if appropriate — note that this doesn't exclude the possibility that, for example, "the cows" could be used to refer to a single cow — this decision is discussed in the main text. "ambgen" should never be set. However the NPs which result from this rule could be interpreted as kinds, in "the poodle is a breed of dog" for example.

```
Rule% NP->Det+Nbar+dist
NP[agr X, type X0, kind X4, ambgen X6, dist +, distq -, sempl +]
  --> Det[agr X, type X0, count +, num X3, dist -]
```

```
      Nbar[agr X, count +, measure -, measured -, kind X4, sempl +]

Semantics%
(and (agr Nbar plur) (num Det +))
      => (L (_Q)
           ((Det (L (_x)
                    (and (Nbar _x)
                         (greater ((NU Nbar) _x) zero))))
            (L (_w)
               (A (_y) (if (and (and (Nbar _y)
                                     (equal ((NU Nbar) _y) one))
                                (ipart _y _w))
                           (_Q _y))))))
(agr Nbar plur)
      => (L (_Q)
           ((Det Nbar)
            (L (_w)
               (A (_y) (if (and (and (Nbar _y)
                                     (equal ((NU Nbar) _y) one))
                                (ipart _y _w))
                           (_Q _y))))))
       => (L (_Q)
            ((Det Nbar)
             (L (_w) (A (_y) (if (and (ipart _y _w)
                                      ((IND Nbar) _y))
                                 (_Q _y))))))
```

This rule gives a distributive reading to semantically plural NPs with group denoting determiners.

```
Rule% NP->Det+Num+Nbar
NP[agr X, kind X4, ambgen X6, dist X5, distq X7, sempl X8] -->
  Det[agr X, count X1, num +, dist X5, distq X7] Num[agr X]
  Nbar[agr X, count X1, measure X2, measured X3, kind X4, sempl X8]

Semantics%
           (or (measure Nbar +)
               (measured Nbar +))
                     => (Det (L (_w) ((Nbar _w) Num)))
                     => (Det (L (_x)
                                (and (Nbar _x)
                                     (equal ((NU Nbar) _x) Num))))
```

The rule for forming measure phrases is different because we don't want to bring in natural units. This rule is equivalent to NP->Det+Nbar — that is it is the non-distributive version

```
Rule% NP->Det+Num+Nbar+dist
NP[agr X, kind X4, ambgen X6, dist +, distq -, sempl +] -->
 Det[agr X, count X1, num +, dist -] Num[agr X]
 Nbar[agr X, count X1, measure -, measured -, kind X4, sempl +]

Semantics%
(L (_Q) ((Det (L (_x)
               (and (Nbar _x)
                    (equal ((NU Nbar) _x) Num))))
         (L (_w)
            (A (_y) (if (and (and (Nbar _y)
                                  (equal ((NU Nbar) _y) one))
                             (ipart _y _w))
                        (_Q _y))))))
```

I have assumed that we don't want to form a distributive version of measure phrases but this may be wrong. I'm not sure whether examples like "Three pounds of apples each filled a paper bag" are OK or not. (The distribution seems to me to be over the measured quantity and not the individual apples, but it seems a very dubious sentence.)

```
Rule% NP->Nbar+exists
NP[agr X, kind X5, ambgen -, dist -, distq -, sempl X8] -->
      Nbar[agr X, cum +,measure -, measured -, kind X5, sempl X8]

Semantics%  (and (count Nbar +) (agr Nbar plur))
            =>  (L (_Q)
                   (E (_y)
                      (and
                        (and (Nbar _y)
                             (greater ((NU Nbar) _y) zero))
                        (_Q _y))))

            =>  (L (_Q)
                   (E (_y)(and (Nbar _y)(_Q _y))))
```

This rule forms an existentially quantified bare NP. The value of "ambgen" is set to "-". For example "men" in "Men were lifting a rock".

```
Rule% NP->Num+Nbar+exists
NP[agr X, kind X5, ambgen -, dist -, distq -, sempl X8] -->
   Num[agr X]
   Nbar[agr X, measure X1, measured X2, kind X5, count +, sempl X8]
```

```
Semantics%

(or (measure Nbar +) (measured Nbar +))
                 =>  (L (_Q)
                        (E (_y) (and ((Nbar _y) Num) (_Q _y))))
                 =>  (L (_Q)
                        (E (_y)
                           (and (and (Nbar _y)
                                     (equal ((NU Nbar) _y) Num))
                                (_Q _y))))
```

This forms an existentially quantified, numerically specified, bare NP which is group denoting. For example "three men" in the group reading of "Three men lifted a rock".

```
Rule% NP->Num+Nbar+exists+dist
NP[agr X, kind X5, ambgen -, dist +, distq -, sempl X8] -->
   Num[agr X]
   Nbar[agr X, measure -, measured -, kind X5, count +, sempl X8]

Semantics%
(L (_Q)
    ((L (_P)
        (E (_y)
           (and (and (Nbar _y)
                     (equal ((NU Nbar) _y) Num))
                (_P _y))))
      (L (_w)
         (A (_z) (if (and (and (Nbar _z)
                               (equal ((NU Nbar) _z) one))
                          (ipart _z _w))
                     (_Q _z))))))
```

This forms an existentially quantified, numerically specified, bare NP which is distributive. For example, "three men" in "Three men (each) lifted a rock."

```
Rule% NP->Nbar+generic
NP[agr X, kind X5, ambgen +, dist +, distq +, sempl X8] -->
   Nbar[agr X, cum +, kind X5, measure -, measured -, sempl X8]

Semantics%  (and (count Nbar +) (agr Nbar plur))
                  =>  (L (_Q)
                         (A (_y)
```

```
                          (if
                            (and (Nbar _y)
                                 (equal ((NU Nbar) _y) one))
                            (_Q _y))))

                 =>  (L (_Q)
                        (A (_y)(if (Nbar _y)(_Q _y))))
```

Universally quantified bare NP (in effect always distributive). This give NPs like "cows" in "Cows are animals".

```
Rule% NP->Num+Nbar+generic
NP[agr X, kind X5, ambgen +, dist +, distq +, sempl X8] -->
 Num[agr X]
 Nbar[agr X, kind X5, measure X1, measured X2, count +, sempl X8]

Semantics%
(or (measure Nbar +) (measured Nbar +))
                =>  (L (_Q)
                       (A (_y) (if ((Nbar _y) Num) (_Q _y))))
                =>  (L (_Q)
                       (A (_y)
                          (if (and (Nbar _y)
                                   (equal ((NU Nbar) _y) Num))
                              (_Q _y))))
```

Universally quantified numerically specified bare NP. There are very few examples where this sort of construction might be needed (it's the "where any three are gathered together" sort of sense). One might use it for "Three men could lift a piano" where it seems we mean any group of three men.

```
Rule% NP->Nbar+kind
NP[agr X, kind +, ambgen X1, dist X3, distq -, sempl -] -->
      Nbar[agr X, kind -, cum +]

Semantics% (L (_Q) (_Q (Kind Nbar)))
```

Bare NPs can also denote kinds, eg "claret" in "claret is a kind of wine".

```
Rule% NP->Name
NP[agr X, dist X1, distq -, sempl X8] -->
  Name[agr X, sempl X8]

Semantics% (L (_P)(_P Name))
```

Names like "John" etc.

```
Rule% Nbar->N
Nbar[agr X count X0 measure X1 cum X2 kind X4 nform X5 sempl X6]
--> N[agr X count X0 measure X1 cum X2 kind X4 nform X5 sempl X6]

Semantics% N

Rule% Nbar->Nbar+PP
Nbar[agr X count X3 cum X4 kind X5 measured X2 sempl X8] -->
    Nbar[agr X measure X2 count X3 cum X4 kind X5 sempl X8]
    PP [measured X2]

Semantics% (L (_x) (L (_m) (and ((Nbar2 _x) _m) (PP _x))))
```

The only PPs in the grammar are in constructions like "three pounds of apples" and those like "a type of cheese". This takes care of the first sort. See lexical entry for "of".

```
Rule% Nbar->kind+of+Nbar
Nbar[agr X count + cum X4 kind + sempl X8] -->
    Nbar[nform knd agr X cum X4 sempl X8] P[pform of measured X5]
    Nbar[agr X1 count X3 kind -]

Semantics% (Type Nbar3)
```

Handles "a kind of wine" etc. See lexical entries for "type" and for "of".

```
Rule% Nbar->kindNbar
Nbar[agr X count + cum - kind + sempl X1] -->
    Nbar[agr X count X3 cum X4 kind - sempl X1]

Semantics% (Type Nbar2)
```

This allows formation of bare kind Nbars and therefore also NPs, eg "claret" in "Claret is a wine".

```
Rule% PP->P+NP
PP[measured X ]  --> P [measured X]
   NP[kind -, ambgen -, dist X6, distq X6]

Semantics% (P NP)
```

Features will need altering — the rule is pretty bogus because "of" is the only preposition.

The VP formation rules below are not much affected by the NP grammar.

Note that the object of an NP cannot be interpreted distributively. There is no reading of "Two boys lifted two rocks" where there are four boys (two for each rock), for example.

```
Rule% VP->V+NP
VP[agr X, takesk X5, prefgen X6] -->
   V[agr X, takes np, prefgen X6, takesk X4, kindsubj X5]
  NP[kind X4, ambgen -, dist X7, distq X7]

Semantics% (V NP)

Rule% VP->Vbe+NP
VP[agr X, takesk X5, prefgen X6] -->
   V[agr X, aux +, takes np, prefgen X6, takesk X5]
  NP[kind X5, ambgen -, dist X7, distq X7]

Semantics% (V NP)

Rule% VP->V
VP[agr X, takesk X4, prefgen X2] -->
        V[agr X, takes -, takesk X4, prefgen X2]

Semantics% V

Rule% VP->V+VP
VP[agr X, takesk X5, prefgen X6]  -->
   V[aux +, sai X4, agr X, vform X1, takes X3]
   VP[gap X2, vform X3, takesk X5, prefgen X6]

Semantics% (L (_x) (V (VP2 _x)))

Rule% Vaux->null

V[aux +, agr X1, takes X2, vform tnsd] -->

Semantics% (L (_P) _P)
```

A.3 Lexicon

Nouns

```
(john (Name [ agr sing sempl - ] john))

(student (N [ agr sing sempl - ]  student))
```

```
(students (N [ agr plur sempl + count + cum +]  student))

(apple (N [ agr sing sempl - ] apple))
(apples (N [ agr plur sempl + count + cum + ] apple))
```

Although "apple" is much more likely to have a mass sense there is no distinction between it and "student" in the lexicon.

```
(puddle (N [ agr sing sempl - count + cum - ] puddle))
(puddles (N [ agr plur sempl + count + cum + ] puddle))
```

Puddle is blocked from having a mass sense, because it seems only to refer to form, rather than stuff.

```
(team (N [ agr sing sempl + ] team))
(teams (N [ agr plur sempl + count + cum +] team))

(wine (N [ agr sing sempl - count - cum +] wine))
(wines (N [ agr plur sempl + count + cum + kind +] (Type wine)))
```

Wines must refer to kinds in the plural.

```
(litre (N [ agr sing sempl - count + cum - measure +]
          (lambda (_x) (lambda (_n) (equal (litre _x) _n)))))
(litres (N [ agr plur sempl + count + cum + measure + ]
          (lambda (_x) (lambda (_n) (equal (litre _x) _n)))))

(type (N [agr sing sempl - count + cum - nform knd] type))
(types (N [agr plur sempl + count + cum + nform knd] type))
```

See special grammar rule `Nbar->kind+of+Nbar`.

Numbers

```
(one (Num [ agr sing ] one))
(two (Num [ agr plur ] two))
```

Conjunction

Only "and", which actually might just as well be syncategorematic in the current grammar.

```
(and (conj (lambda (_np1)
              (lambda (_np2)
                 (lambda (_P) (and (_np1 _P) (_np2 _P))))))
     (termjoin (lambda (_np1)
```

```
                (lambda (_np2)
                    (lambda (_P)
                       (_np1 (lambda (_z)
                          (_np2 (lambda (_y)
                                 (_P (iplus _z _y))))))))))
```

Adjective

```
(scarce (Adj [ takesk + ] scarce))
```

Verbs

```
(is (V [ aux + agr sing takes ing vform tnsd ] (lambda (_P) _P))
    (V [ aux + agr sing takes np vform tnsd prefgen + takesk X1]
       (lambda (_np)(lambda (_x)
                         (_np (lambda (_y) (equal _x _y)))))))

(sleep (V [ vform inf aux - takes - ]  sleep)
       (V [ vform tnsd agr plur aux - takes - ] sleep))
(sleeps (V [ vform tnsd aux - takes - agr sing ]  sleep))
(slept (V [ vform tnsd aux - takes - prefgen - ] sleep))

(convene (V [ vform inf aux - takes - ]  convene)
       (V [ vform tnsd agr plur aux - takes - ] convene))
(convened (V [ vform tnsd aux - takes - prefgen - ] convene))
```

Note that there is no difference between convene and sleep — distributivity is not regarded as lexical at all.

```
(costs (V [ vform tnsd aux - takes np agr sing prefgen +]
        (lambda (_np)
            (lambda (_x)
               (_np (lambda (_y)(cost _x _y)))))) )
```

Determiners

```
(some (Det [ agr X type n count X1 num + dist - distq -]
    (lambda (_P)(lambda (_Q)(some (_x)(and (_P _x)(_Q _x)))))))
(the (Det [ agr X type n count X1 num +  dist - distq -]
      (lambda (_P)(lambda (_Q)
                 (some (_x)
                   (and (and (_P _x)
                             (all (_y)
                               (if (_P _y) (ipart _y _x))))
                        (_Q _x)))))))
```

Determiners like "some" and "the" are group denoting.

```
(every (Det [ agr sing type n count + num - dist + distq +]
         (lambda (_P)
           (lambda (_Q)
             (all (_x)(if (and (_P _x) (equal ((NU _P) _x) one))
                          (_Q _x)))))))
```

Quantifiers like "every" are unambiguously distributive and are therefore marked `dist + distq +`.

```
(each (Det [ agr sing type n count + num - dist + distq +]
         (lambda (_P)
           (lambda (_Q)
             (all (_x)(if (and (_P _x) (equal ((NU _P) _x) one))
                          (_Q _x))))))
      (fquant [ dist + distq - ] each))
```

The floating quantifier sense is, in effect, introduced syncategorimatically (for no particularly good reason).

```
(all (Det [ type n count X cum + num + dist + distq +]
     (lambda (_P)(lambda (_Q)(all (_x)(if (_P _x)(_Q _x)))))))

(a (Det [ agr sing type n count + num - dist X distq -]
     (lambda (_P)
      (lambda (_Q)
         (some (_y)
            (and (and (_P _y)(equal ((NU _P) _y) one))
                 (_Q _y)))))))
```

Note the specification of natural units.

Prepositions

(Actually only "of" as it occurs in partitives.)

```
(of (P [measured + pform of] (lambda (_np)
           (lambda (_x) (_np (lambda (_y) (ipart _x _y)))))))
```