Technical Report

UCAM-CL-TR-137
ISSN 1476-2986

Number 137

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Mass terms and plurals: from linguistic theory to natural language processing

F.V. Hasle

June 1988

15 JJ Thomson Avenue
Cambridge CB3 0FD
United Kingdom
phone +44 1223 763500

https://www.cl.cam.ac.uk/

Technical reports published by the University of Cambridge
Computer Laboratory are freely available via the Internet:

https://www.cl.cam.ac.uk/techreports/

ISSN 1476-2986

Abstract

Two linguistic theories within the tradition of formal semantics are investigated. One is concerned with *mass terms*, and the other with *plurals*.

Special attention is paid to the possibility of implementing the theories on a computer. With this goal in mind their basic ideas are examined, and the linguistic implications are discussed. In the process, various features of the theories are made formally precise. This leads to two formal systems, one for representing the meanings of sentences with mass terms, and another similar one for plurals. The systems are specified by machine-executable translation relations from fragments of natural language into logical representations.

The underlying model-theoretic semantics of each theory is partially axiomatised. From the axiomatisations all the paradigmatic inferences of each theory can be proved in a purely deductive manner. This is demonstrated by a number of rigorous proofs of natural language inferences.

Finally, some methodological issues are raised. Both theories recommend a particular approach within formal semantics for natural language. I explore the methodological views underlying the theories, and discuss whether the authors actually follow the methods which they recommend.

A note on the appendices

As a space-saving device, the appendices mentioned in this technical report are actually not included, with the exception of the grammar and lexicon of **appendix D**. The full appendices can be had from the author or from Steve Pulman, University of Cambridge Computer Laboratory.

Contents

Part I
Introduction

1 General Outline

This report is concerned with the linguistic analysis of mass terms and plurals. The approach taken to the problem domain is within the tradition of formal semantics for natural languages. Since all the grammars and also the inference systems of the report have been implemented on a computer, the investigation is also a contribution to the field of computational linguistics. . However, the investigation has been conducted under a specific perspective, which qualifies its status within the discipline of formal semantics. That perspective is the question of how to make "computational sense" out of linguistic theories of this tradition. I do no seek a *general* answer to that question, though; but throughout the investigation it is insisted that the fragments of language under scrutiny be made machine-executable, both w.r.t. their translation relations and their underlying semantic inference systems.

The methodological implications of this perspective will be addressed more closely below, but this much should be said right away: the goal of full implementation compels me to cut down on linguistic breadth. In fact my investigation essentially builds on only two recent theories within the problem domain; but as you will see there is enough work to be done simply in order to analyse these theories in depth and "prepare" them for computational implementation. It should be noted, however, that the theories in question by and large comprise the state of the art within the problem domain, and for that reason this apparent limitation is not as severe as it may seem.

It may be as well to give a general outline of the investigation now, even if that causes a bit of overlap with the actual examination of the two theories.

The theory which I examine first is the theory of Pelletier & Schubert on mass terms. This theory was stated rather sketchily in [PS84]. From the authors I obtained a more detailed account of the theory: "F. J. Pelletier and L. K. Schubert: Mass Expressions", which is to appear in "D. Gabbay and F. Guenthner: Handbook of Philosophical Logic", Vol. 4 ([PS86b]). My investigation is based on that as yet unpublished paper, from which various quotations are taken, and to which page references apply. This paper is enclosed as **Appendix H.**

In spite of the ingenuity of the theory it is not unfair to say that the formal system is only vaguely sketched. A careful examination makes it possible to "reconstruct" the intended formal system, with some necessary modifications to the theory. In a complementary way this reconstruction makes it possible to explain and assess various claims made in Pelletier & Schubert's paper—in particular claims w.r.t. which inferences the theory validates.

With this reconstruction of the theory it becomes possible to give a full implementation of its basic ideas, by the following two steps:

1. First I define a machine-executable translation relation from a fragment of natural language containing mass terms into a logical target language;

2. Then an axiomatic system with rules of inference is defined; this system makes it possible to give purely deductive proofs of those natural language inferences, which Pelletier & Schubert have stated as paradigmatic within their theory.

Furthermore, the axiomatisation has been implemented on computer by Thomas Forster, so all the inferences have also been tested mechanically; cf. p. 7.

The work on plurals closely follows G. Link's famous 1983 paper "The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach" ([Lin83]). That paper contains a framework for the treatment of plurals, a framework which itself gives rise to various questions. I implement the framework in a manner analogous to the implementation of the theory of Pelletier & Schubert, and in the process I try to resolve or at least discuss the questions arising from it.

Link has himself given a review of the ideas of his 83-paper in a later paper called "The Logic of Plurals, LP: Review of the basic Ideas" ([Lin86]). At the time of conducting my investigation into Link's 83 framework I was unaware of this paper (which was by then unpublished). However, as the title suggests, the 86-paper is basically a review of his earlier ideas, and do not alter those. A novelty in the 86-paper is a suggestion of lifting his logic for plurals into the Generalised Quantifier Framework. I should say that the prospect of combining the two can only make a thorough investigation into the basic ideas more important. Another novelty is a brief suggestion of how to deal with distributive transitive verb phrases in the plural. My investigation into such verb phrases has as its "point of reference" the 83-paper only. However, my work on this subject and Link's hints in his 86-paper are compatible; indeed, with hindsight the former can be seen as a thorough background examination of the latter. I shall address their relationship a bit closer in section 11.1.

Finally, the only other way in which the 86-paper deviates from the 83-paper is by suggesting various ways of broadening the linguistic scope of the earlier paper, but without altering any of the basic ideas. Again, I should say that the prospects of extending the fragment covered by the theory can only make an examination of its basic tenets more interesting.

2 Methodological Considerations

The investigation into the problem domain has been conducted with a specific goal of computational implementation. That goal has greatly influenced the method of investigation, and has also had other sorts of methodological implications. The easiest way to understand those implications is probably to consider the overall system resulting from the investigation.

The diagram of Fig. 1 (cf. p. 8) gives an overview of the overall system. Strictly speaking there are two systems of this form, one for each of the two theories in question[1].

The rectangular boxes are those parts which result directly from my implementation of the two theories. The first box contains a CFG + translation relation. This is a translation relation whose grammar part is a context-free grammar extended with syntactic features, and in a format which makes it executable by S. G. Pulman's parsing system (see below).

The second box contains my axiomatisations of the semantics of the two theories[2]. The axiomatisations make it possible to determine in a purely

[1] I have no doubt, though, that the two implementations can be integrated into one such system. In **Appendix D** I have sketched how that could be done for the translation relations. A similar amalgamation of the inference engines would then provide us with just one system, capable of translating natural language fragments with mass terms and plurals into suitable logical forms, and of carrying out natural language inferences involving those constructions.

[2] No proofs of soundness and completeness are given. Like most computational linguists

deductive manner which inferences are valid in the two theories.

The oval boxes contain software capable of executing the rectangular components of the system. The parsing system was implemented by S. G. Pulman. Given a CFG+[3] translation relation for a fragment of natural language, the system translates input sentences from this fragment. In the system here, the translations are of course designed to be logical representations of the meaning of the input sentences.

The inference engine which executes the axiomatic system was written by Thomas Forster. In fact, it was specifically developed on the basis of the axiomatisations of the two theories[4] in order to implement those on computer.

Technically, the inference engine mechanises a HOL (Higher Order Logic) theory. The HOL system was developed by Mike Gordon (cf. [Gor85] and [Gor87]).

Now this system in its entirety actually amounts to using a fragment of English as a *logic programming language*[5]. To show how I shall give a slightly simplified, but conceptually correct picture of how the system works on a natural language input sentence. For instance, it is possible to feed the system with a sentence like

"John and Mary see some colleges, so John sees a college";

the system first translates the sentence into an appropriate logical form, which is then passed on to the inference engine as a query[6]. The inference

---

I think that proving such meta-mathematical properties had probably best be left to proper mathematical logicians. But I may also add that the ultimate measure of soundness and completeness for these theories really is natural language itself. To that extent the soundness of the axioms and inference rules is really a matter of basic intuition more than of any concrete model theory, and the measure of completeness are empirical linguistic data.

[3] Which is, incidentally, Steve Pulman's term for translation relations executable by his parsing system. See also [Pul88].

[4] Of course, it would have preferable to have a *general* inference engine capable of carrying out *any* axiomatisation of a certain format—i.e. having a meta-theoretical framework. But that will have to await further research.

[5] Albeit in a restricted sense. I am thinking of the way in which databases are constructed and queried in a language like **Prolog**. But of course the system developed here does not possess the *general problem solving capability* of a logic programming language fully worth the name.

[6] This convention is obviously not quite faithful to the natural language sentence, which is *declarative*. It is merely a matter of convenience, though; it would be quite possible to

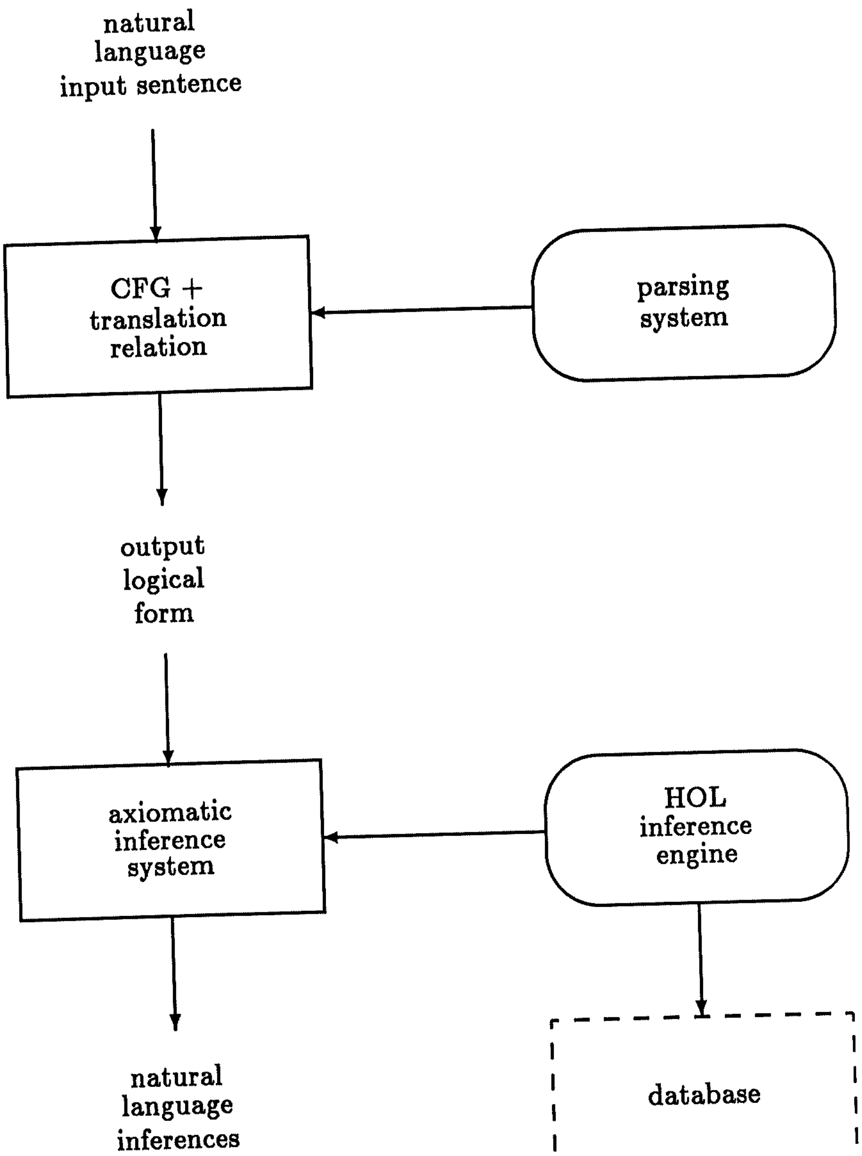

Figure 1: The Overall System

engine will establish the validity of the natural language sentence within the axiomatic inference system, and answer "yes"—thus confirming that the sentence is a theorem of the logical theory.

Equivalently, we could keep the information "John and Mary see some colleges" in a database[7]. Given the query "John sees a college" the inference engine would then access the database[8] and establish the truth of the query w.r.t. the axiomatic system *and* the database—and hence again the answer would be "yes".

Alternatively, if the theoremhood of the query cannot be established w.r.t. the current theory of the overall system, the inference engine will answer "no". I think that the affinity between this type of behaviour and that of a language like **Prolog** is quite striking.

The relationship between formal linguistic theories and logic programming languages has important consequences, which work both ways. Let me take the implications for linguistic theory and in particular the following investigation first.

The development of a system like the one in Figure 1 has been a main objective of the investigation right from the start[9]. This has put some severe but also fruitful constraints upon the investigation. The quest for the absolute precision necessary for computational implementation has compelled me to go into more depth with the underlying tenets of the two theories than they actually do themselves. This has led not only to full formal precision, but also to greater conceptual clarity. That is, the process of formalising rigorously brings forth various depths and also oddities of the theories of which the authors apparently (and sometimes only too obviously) have not

treat only interrogative sentences as queries, and treat declarative sentences as information to be stored in a database.

[7]We could actually keep the information in exactly that natural language form. That would be inefficient, of course, for it would then have to be translated whenever the inference engine needed it; but as a matter of principle it is important that we can use natural language directly in this manner.

[8]This procedure hasn't actually been tried, and for that reason the database is indicated by a dashed box; but clearly it would be possible to proceed as described here.

[9]See the project proposal to the United Kingdom Science and Engineering Research Council by Mike Gordon and Steve Pulman: "Logic and Natural Language Understanding", and my own project proposal to the Danish Research Council for the Humanities, "An Investigation into the Relationship between Philosophy of Language, Mathematical Logic, and Logic Programming" (Original title: "En undersøgelse af sammenhængen mellem sprogfilosofi, matematisk logik og logikprogrammering.") See **Appendix E**.

been aware themselves. Thus the need for computational implementation actually has significant repercussions on the development of linguistic theory. Of course, this point has been made by others before; I am merely saying that this investigation in my opinion substantiates the point.

But the priorities set by the objective of computational implementation also compels me to cut down on linguistic breadth. Compared with the two theories I investigate I have pruned the range of linguistic data to be accounted for; and another limitation is the very fact that I have based my investigation into the problem domain on only those two theories. Actually I make a point of staying as faithful to those as possible and not taking issue with their *linguistic* tenets except where absolutely necessary. The amount of work which is nonetheless required should explain why I do not go into a comprehensive comparison of available theories in the field.

However, as I remarked above, those two theories by and large represent state of the art[10]. Indeed I firmly trust that as regards computational implementations within the problem domain, the system developed here is at least as comprehensive as any other in the field.

There are three restrictions w.r.t. linguistic scope which ought to be pointed out separately:

1. I am dealing only with the *extensional* subfragments of the theories in question. These seem, however, to contain all the paradigmatic cases of the problem domain.

2. Group terms are excluded from the fragments studied here. Group terms are surely important for the problem domain, but they also seem to raise a host of new and partly independent problems (see section 3.3 below).

3. Generic readings are not considered[11]. Like group terms, these seem to raise a number of new and independent problems—and moreover, they may belong to the realm of pragmatics rather than semantics[12].

[10]With the exception of the work of Harry Bunt (cf. [Bun85]); this work is based on *Ensemble Theory* and represents a very different (and more complicated) approach to the problem domain. I also except some very recent and as yet unpublished papers, e.g. by Fred Landman (Cf. [Lan87]).

[11]Unless Pelletier & Schubert's "kind-readings" for mass terms, which will figure prominently in the treatment of those, are considered to be a sort of generic readings.

[12]Since I don't deal with those readings I can hardly embark on a substantial argumentation to the effect that they belong under pragmatics. But I may say this much: it seems that all the sentences "man is mortal", "a man is mortal", "the man is mortal", "every

As regards the limitations on other kinds of linguistic data, these are a mere matter of convenience; the system developed actually lays a firm foundation for incorporating the full range of syntactic constructions considered by the two theories.

Now let me go on to consider the potential impact of linguistic theory upon logic programming. In my opinion formal logic has its roots in the analysis of natural language, be it a linguistic or a more philosophical analysis. Certainly predicate calculus has at least some of its origins in the philosophy of Aristotle; and Gottlob Frege, who took those notions from the realm of philosophy into a full-fledged mathematical system, saw his own work as being philosophical as well as mathematical in kind. Logic programming languages in turn are based on those developments, although I am aware, of course, that this does not tell the *full* story about them.

If this very brief outline holds, it is only to be expected that the analysis of language can contribute to the development of logic programming and related fields within computer science. I believe that the system of Figure 1 does exactly that, even though it is, of course, on a very modest scale. What the system demonstrates is the fact that certain linguistic theories can form part of the foundation of a logic programming language. In the current context it is a language which facilitates the representations of and inferences with mass and plural entities. It is conceivable—in my opinion likely—that we shall meet cases of expert and artificial intelligence systems, where this will be useful.

It is most probable that for such concrete cases the system here is too inefficient. Instead some sort of standardised representation—rather than natural language itself—would presumably be needed. But the point here is that we have managed to derive and implement notions from natural language, such that we now have a clear idea of those notions and know that they can be used computationally. From that point on the further decision on whether and how to use them within computer science proper should naturally be left to computer science itself.

---

man is mortal", "men are mortal", "all men are mortal", "the men are mortal", can be interpreted as having a "generic reading"—in fact, one and the same "generic reading" for all the sentences in question. And that strongly suggests to my mind that that reading cannot be a semantic function of the determiners and nouns of those sentences—i.e. that that reading is not strictly a semantic part of the meaning of any of those sentences, but rather a phenomenon belonging to the realm of pragmatics.

Following the discussion here we may sum up the relationship between computer science and linguistic theory as follows: computer science helps in developing linguistic theories, since it increases demands for precision and also provides a means of testing the theories; and linguistic theory has the tradition and the formal apparatus for contributing to the development of concepts for logic programming languages and related areas.

3 The Problem Domain

The following introduction to the problem domain will be brief. In accordance with the priorities of this investigation I cannot undertake to give any comprehensive survey of the problem domain; I have to leave it to the reader himself to acquire such a broader perspective on the investigation, if he feels so inclined. This section will be exclusively concerned with those salient features of the problem domain which are of direct relevance to the investigation here. But it may be added that those features are the crucial ones within the problem domain, upon which the solution to other phenomena seem to depend—such as for instance floated quantifiers, backwards pronominalisation, numerals and amount terms, etc. etc. At any rate, the following remarks should be sufficient background for understanding the investigation to follow.

Plurals and mass terms are often dealt with together, because they do share some striking features. Those features in turn set both phenomena apart from the traditional domain of formal semantics and philosophical logic, which as far as nouns are concerned have dealt essentially with singular count nouns only[13]. I follow the tradition of dealing with mass terms and plurals together, if only to explain why their affinity to each other has been felt so strongly[14].

---

[13]Except for those spurious cases, where the noun phrase in question could be paraphrased into a singular count noun phrase; e.g. "all men are mortal" has been treated as equivalent to "every man is mortal".

[14]The fact that this investigation deals with eactly those two phenomena is also a bow to this tradition. But as you will see during the investigation, they also differ in important respects, and I for one have actually come to doubt the wisdom of trying to solve both by the same means.

3.1 Syntactic Properties

There are a number of syntactic properties which set mass nouns, plurals, and singular count nouns apart from each other, but the crucial ones are the following:

Pluralisation: Mass nouns are generally assumed not to occur in the plural, as opposed to count nouns:

- Books are expensive.
- *Golds are expensive.

When putative mass nouns nonetheless occur in the plural, it seems to be in a sort of "count sense" different from their normal sense:

- The wines of France are over-priced.

Determiners: Mass and count nouns differ w.r.t. the determiners with which they can co-occur; in particular the ability of co-occurring with the indefinite article "a" is often regarded as a criterion of being a count noun, whereas the absence of this ability is seen as a criterion of being a mass noun:

- John bought a book in Cambridge.
- *John bought a gold in Cambridge.

On the other hand, when a noun co-occurs with the determiners "much" and "little", it is generally assumed to be a mass noun:

- *John bought much book in Cambridge.
- John bought much gold in Cambridge.
- *John bought little book in Cambridge.
- John bought little gold in Cambridge.

When a noun is capable of occurring in both kinds of contexts, e.g.

- We had a chicken for dinner.
- We had chicken for dinner.

the first occurence may be classified as an occurrence "as a count noun" and the latter as an occurrence "as a mass noun"; or it can even be postulated that there are really *two* (homonymous) nouns involved, one count and the other mass.

Bare noun phrases: Bare mass nouns and plurals can function as noun phrases:

- Books are expensive.
- John bought books on his visit to Cambridge.
- Gold is expensive.
- John bought gold on his visit to Cambridge.

On the other hand, singular count nouns are apparently incapable of forming noun phrases by themselves:

- *Book is expensive.
- *John bought book on his visit to Cambridge.

Adjectival function: Mass and plural nouns can occur "as adjectives". Both can occur in predicative position:

- This ring is gold.
- John and Mary are students.

But singular count nouns cannot occur in this function:

- *This person is student.
- *John is student.

Furthermore, mass nouns—but not plurals—can also occur "as adjectives" in attributive position:

- This gold ring is expensive.

but no similar function is allowed for count nouns:

- *This student person is likeable.

3.2 Semantic Properties

There are a number of semantic concepts commonly used in discussions of this problem domain, which describe its most important semantic phenomena. My explanation of those concepts will be somewhat loose here; the finer details will be added as formal systems are developed later.

Cumulativity: This is used to describe the following general property of mass and plural entities: when two such entities with some shared property are fused or added together, the resulting new entity also has that property[15], as the following valid inferences illustrate:

- This piece of metal is gold and that piece of metal is gold, so the fusion of those two pieces is gold.
- John and Mary are students, and the freshmen are students, so John and Mary and the freshmen are students.

Actually, for plurals this property also applies at the "singular level":

- John is a student, and Mary is a student, so John and Mary are students.

Dissectiveness: This describes the feature that any part of a mass or plural entity has the same properties as that entity:

- The mug is gold, the mug has a handle, so that handle is gold.
- The undergraduates are students, John and Mary are undergraduates, so John and Mary are students.

Again, for plurals this property stretches down to the singular level:

- The undergraduates are students, John is an undergraduate, so John is a student.

On the other hand, mass nouns pose a particular problem here, sometimes called the **problem of minimal parts**: for obviously dissectiveness for mass nouns holds only down to some empirically determined lower limit; the individual electrons in a piece of gold are not themselves gold.

Homogeneous reference: This is simply a comprehensive notion for the two above properties: if a term is subject to both cumulativity and dissectiveness, it is said to have homogeneous reference (so mass terms and plurals in general have homogeneous reference—take "wine" and "the students", for instance.).

[15]Provided that the property in question is ***distributive***, cf. ***distributive reading*** below.

In connection with verb phrases mass and plural entities give rise to a distinction between *distributive* and *collective* readings:

Distributive reading: Certain verb phrases can or must be understood as applying not only to their subject phrase "in its entirety", but also to each of its individual parts:

- The students saw a college, John is a student, so John saw a college ("see a college" is unambiguously distributive).
- Wine is liquid, claret is a kind of wine, so claret is liquid (the property of "being liquid" is distributive among the subkinds of the subject phrase[16]).
- The students [each] built a college, John is a student, so John built a college ("build a college" can be read distributively).

Collective reading: Some verb phrases can or must be understood as applying to the subject phrase "in its entirety", but *not* to its individual parts:

- The students gathered, John is a student, *so John gathered ("gather" is unambiguously collective, so the inference is clearly not valid—indeed its consequence may be regarded as ungrammatical).
- The students built a college [together], John is a student, *so John built a college ("build" can be read collectively; on that reading the inference is invalid).
- Wine is widespread, claret is a kind of wine, *so claret is widespread ("is widespread" is unambiguously collective).

[16] This inference will perhaps become easier to understand in the course of the investigation into mass terms. The central idea here is that "claret" as a subkind of "wine" can be regarded as a *part* of the *kind* "wine". (Ultimately Pelletier & Schubert's theory deals with this inference as a case of *transitivity*, but pre-theoretically it is natural to regard it as a case of *distributivity*.)

3.3 A Note on Group Terms

Group terms such as "the committee" resemble plurals such as "the students" in that they have (at least in a sense) individual parts—e.g. the members of the committee. Yet "the committee" is not *identical* with the sum of its members, as opposed to "the students". Thus "the committee convenes" may be true even though some of its members are absent, whereas "the students convene" entails that each and every student is present.

Moreover, in the plural group terms sometimes are ambiguous in a way that "ordinary" plurals are not: "the committees convened" may be read distributively as well as collectively, as opposed to "the students convened". (See also p. 142.)

So although group terms exhibit some obvious structural analogies with plurals, they clearly also pose some problems of their own. In his 86-paper, Link himself remarks of group terms (the symbol $\leq_i$ stands for the relation "is an individual part of"):

> "Note, however, that most things that are intuitively called parts in our world ... are *not* i-parts [individual parts] in the sense of the $\leq_i$-structure. Not even group terms have i-parts, according to LP [Link's logic of plurals]; the $\leq_i$-structure is **only to take care of pluralities.** Thus, John is an i-part of the individual sum consisting of John and Mary, but a committee doesn't have i-parts in LP. This fact about group terms is certainly unsatisfactory from a semantical point of view, but formally, *committee* is singular, and therefore a committee is an atom ... **A special semantics of group terms should add more structure here.**
>
> [Lin86, page 2]

Thus the exclusion of group terms is at least partly a consequence of working within Link's framework.

3.4 The General Importance of the Problem Domain

Until very recently formal and philosophical logic alike have been preoccupied with atomic individuals, at least in so far as direct ways of expressing plural and mass entities have been lacking in those disciplines. However, such entities are pervasive in natural language: especially constructions in the plural seem to be as frequent and fundamental in the syntax of English as constructions in the singular. Furthermore, such constructions—and also constructions with mass nouns—belong in principle to that logical core of natural language, which can be described by extensional first order logic (or at least a significant number of those constructions do, and that suffices to make the point that we meet them already at a very fundamental level of the logic of natural language[17]).

I think these observations make it clear that a coherent solution to this problem domain will be a major step forward, both for the project of formal semantics for natural language and for logic in general. As regards computer science I have already tried to point out the potential relevance of solving these problems—although the systems developed here are, of course, still not a general solution to the problem domain.

4 Notation, Terminology, Etc.

The phrase "mass terms" has come to be the standard heading when discussing the problem domain involving mass nouns and possibly also other sorts of mass expressions. However, as Pelletier & Schubert have rightly pointed out[18], that heading is somewhat misleading. In Montague Grammar as well as in logic in general "term" has a fixed meaning, which is too narrow to cover the possible occurrences of mass nouns even in the restricted fragments studied here. For instance, we have already seen that a mass noun can occur as a predicate (cf. "adjectival function", p. 14). Pelletier & Schubert opt for the heading "mass expressions" instead. However, I stay faithful to tradition on this point except when the broader concept is strictly required—I then follow Pelletier & Schubert and speak of "mass

[17] To be a bit more specific about it: the examples of the previous section should make it clear that these phenomena basically occur in purely extensional contexts. As regards my later formalisations, these treat the denotations of mass terms and plural terms strictly as "first-order objects", but to that I must add that predicate modifiers and implicit higher order quantification nonetheless occur in those systems.

[18] Cf. [PS86b, page 1], footnote 1.

expressions".

Finally, before setting out on the investigation, a bit of practical information is required:

Quotations: In quotations I use **boldface** to emphasise parts which are not originally emphasised, but to which I have wanted to draw particular attention.

Text which is surrounded by square brackets has been interspersed by me to clarify or supplement the quotation.

Translation relations: The full machine-executable translation relations together with their lexicons can be found in appendices A, B, C, and D. Some of the rules of those translation relations contain syntactic features, which are not strictly necessary in order to provide grammars for the fragments under consideration. Such features have been incorporated in order to ensure compatibility with Steve Pulman's more general CFG+ translation relation for English. The extra features can hardly cause confusion for anyone studying those grammars; it should always be quite clear which features are important for the fragment in question.

In general the grammars cover wider fragments than have actually been discussed in the course of the investigation, first and foremost because the lexicons contain a number of determiners beyond those directly investigated. Such extra coverage should be seen only as suggestions of how to enlarge the fragments in the future.

With respect to proper names I follow a convention of Steve Pulman's: every word beginning with a capital letter is treated as a proper name by the overall system.

When discussing translation relations I use the following bits of notation:

A natural language expression in parentheses followed by a prime, e.g. (John and Mary)', (John and Mary die)', designates the translations of those expressions relative to the translation relation under consideration; for the notation to be meaningful the expressions must be unambiguous. On the other hand, I omit the traditional prime in the translations of verbs and nouns.

I use the symbol '$\triangleright$' to designate the relation of "translates unambiguously into"; so

$$\text{John and Mary die} \triangleright \phi$$

means that "John and Mary die" has ϕ as its one and only translation (in the translation relation currently under consideration). The symbol '$\trianglerighteq$' is used in connection with ambiguous sentences, e.g.

$$\text{John and Mary carry a piano} \trianglerighteq \phi$$

indicates that ϕ is one among more translations of the sentence in question.

Inference systems: The axioms of my inference systems are strictly speaking *axiom schemas*, but I follow tradition and simply call them axioms.

As for the style of proofs I have chosen to follow a widely used logic textbook, [Cop78]. This book states a relatively large number of inference rules for predicate logic, rather than limiting their number as one often strives to do in mathematical logic. Although less elegant as a mathematical system I think that a relatively large number of inference rules is more intuitively appealing when proving natural language inferences.

To some extent analogous remarks apply to my own axiomatisations—I have not tried to reduce the number of axioms for the sake of mathematical compactness of the systems. Among *all* the theorems of the systems I have selected those to be axioms, which I perceive to state fundamental and pre-theoretic intuition about the problem domain. The theorems which I prove are those which are either generally helpful in proving natural language inferences, or which otherwise express characteristic features of the systems being developed.

In the proofs I always indicate which inference rules of predicate logic are being used for each step; in this connection I use the following abbreviations: E.I. and E.G. for existential instantiation and generalisation, respectively, and U.I. and U.G. for universal instantiation and generalisation. Furthermore a number followed by a prime, e.g.

2', indicates a simplification[19] of the formula of the corresponding step (e.g. step 2) of the proof.

Other notational items: When discussing models I use the symbol $\|\cdot\|$ for the semantic valuation function of that model. Strictly speaking a sub-script should be used, such as $\|\cdot\|_{\mathcal{M}}$ when discussing a given model $\mathcal{M}$, but no confusion can arise in the contexts here, and so I venture to drop such sub-scripts. However, w.r.t. the usual symbol for truth in a model, I have to use such subscripts—e.g. as in $\models_{\mathcal{M}}$—in order to avoid confusing this concept with validity in general (simply $\models$). As usual the symbol '$\vdash$' indicates theoremhood (within a logical theory under consideration). I also use it in the individual steps of my proofs (which are always purely deductive).

When stating two-place relations in a model, I generally use $a, b, c, \ldots$ for the first-place members of the ordered pairs and $\alpha, \beta, \gamma, \ldots$ for the second-place members. To my mind, this sometimes makes it easier to understand the model. In logical forms involving a two-place predicate δ I alternate between using the notation $\delta(y)(x)$ and the "relational notation" $\delta(x, y)$, according to whichever is clearer in the context.

With respect to parentheses I sometimes insert extra pairs for the sake of perspicuity.

[19] According to the inference rule called *simplification*: $\phi \wedge \psi \vdash \phi$.

Part II

The Theory of F. J. Pelletier and L. K. Schubert on Mass Expressions

5 Introduction

One of the most recent theories within the problem domain of *mass terms* is presented by F. J. Pelletier and L. K. Schubert in [PS86b]. Both of these two researchers have been working within this domain for a number of years—especially F. J. Pelletier, who has for a long time been acknowledged as one of the foremost authorities on the subject. The theory incorporates the experience and the progress within the domain for the last few years, and so there can be little doubt that this theory will emerge as one of the most important contributions for the foreseeable future.

In fact two theories are presented in [PS86b], the so-called *p-theory* and the *s-theory*[20]. However, these two theories are for all practical purposes very close to being identical; they share all significant features which distinguish any one of them from other theories in this field. Notably they both allow *all common nouns to occur as both mass and count nouns*: that is, sentences such as "there is man all over the floor" and "a gold is precious" are not seen to be ungrammatical. Instead, "man" in the first sentence simply receives a "suitable mass interpretation", whereas "gold" in the second sentence receives a "suitable count interpretation".

The two theories differ, however, in one respect: the *s-theory* is a *syntactic lexical approach*, whereas the *p-theory* is a *semantic occurrence approach*[21]. Very briefly, the difference is that the *syntactic lexical approach* equips common nouns with features such as *+mass* and *+count*, and obtains the "stretched uses" such as those of "man" and "gold" in the above sentences by auxiliary syntactic rules. In contrast, the *semantic occurrence approach* makes no (syntactic) mass-count distinction at all. Instead a noun

[20]No explanation of these names for the two theories is given. However, previous work by Pelletier makes it likely that *p-theory* stands for *pelletier-theory*, and by analogy we may infer that *s-theory* stands for *schubert-theory*.

[21]Pelletier & Schubert use the term *syntactic expression approach* rather than *syntactic lexical approach*; the latter seems to me to be more accurate, though. For a full explanation of the notions *syntactic lexical approach* and *semantic occurrence approach*, see [PS84].

receives a *mass-* or *count*-interpretation *in the semantics* according to how it occurs in the translation; for instance, when a singular common noun occurs as a bare noun phrase, it will receive a *mass-interpretation*, such as "man" in the above sentence. So, in the *p-theory* there are no basic categories *a priori* marked as *mass* or *count*. From a theoretical point of view, the *p-theory* seems to be the "purest", and it is this one I am going to investigate in the following. However, I always refer to the theory as being by "Pelletier & Schubert", since their two theories differ only in this less important aspect.

The investigation about to be embarked upon could suitably be called a "reconstruction" of the theory of mass expressions put forth by Pelletier & Schubert in [PS86b]. Although the theory which is eventually developed here is essentially based on Pelletier & Schubert's ideas, it is not improper to call it a "reconstruction"; for their own exposition in [PS86b] is so riddled with imprecision that it doesn't do the ingenuity of their ideas any justice.

Sometimes my comments and results contradict those of [PS86b], but I have to say now that the reason for such discrepancies is to be found in the imprecisions of that paper. Some of those imprecisions will of necessity be addressed in the following account, and these cases should provide sufficient justification for my critical remarks here. I may add that there is also a certain slipperiness about the terminology in [PS86b]—thus the crucial concept of "being a *kind* of" progressively undergoes slight changes in meaning throughout the paper.

I shall not elaborate further on that subject, though. I have only made these points in order to forestall some avoidable confusion on part of the reader, and also to make it more understandable why my own account proceeds the way it does.

It is not my intention merely to paraphrase [PS86b] into a more understandable exposition, nor for that matter to repeat here the gradual and motivated introduction of the various concepts in that paper. I presuppose [PS86b] in the sense that much of the linguistic discussion of and motivation for the theory must still be found there—my task here is to *explicate* the notions of [PS86b] and to *implement* those notions in accordance with the general goals of my investigation.

To achieve these goals I proceed with a top-down approach: first, I explicate the central notions of [PS86b] (according to my reconstruction of those); next, I briefly describe their formal content; then I go on to discuss Pelletier & Schubert's translation relation, which brings those concepts into use; and then I am finally in a position to discuss those interpretations of individual sentences, which they claim their system gives rise to, and the inferences which follow from those. Not surprisingly this investigation shows that certain revisions are necessary, and with these revisions in mind I then go on to give an implementation of the crucial parts of the theory.

6 A Reconstructive Exposition of the Theory

6.1 The Special Symbols and Their Intuitive Interpretation

The core of the theory is given by four special operators, introduced in order to be able to account properly for the semantics of mass expressions.

In the following, let P be a predicate, e.g. *wine* or *man* (the translations of the common nouns "wine" and "man", respectively). Then the operators are informally defined as follows:

> μ: This is a name-forming operator, applying to predicates and yielding names. In practice it is restricted by the translation relation to apply only to predicates which are translations of common nouns. The type of μ is $\langle\langle e,t\rangle,e\rangle$[22].
>
> If M translates a basic common noun, then $(\mu\ M)$ names the most general kind of M; e.g. $(\mu$ *wine*$)$ names the kind or the substance "wine". $(\mu$ *(cheap wine)*$')$ names "the most general kind of cheap wine" (having for instance "cheap red wine" as a subkind).
>
> For more formal details, see the section below.

The next operators β, γ, and δ differ in type from the μ-operator, since they are all predicate modifiers. In the spirit of Montague Grammar (modulo intensionality) they are of type $\langle\langle e,t\rangle,\langle e,t\rangle\rangle$:

> β: $(\beta\ P)$ denotes the ***conventionally recognised*** subkinds of P; thus, $(\beta$ *wine*$)(\mu$ *claret*$)$—corresponding to "claret is a conventionally recognised subkind of wine"— is true (in the "actual" world); $(\beta$ *wine*$)(\mu$ *(cheap wine)*$)$ is false, for although "cheap wine" is a kind of wine it is not a conventionally recognised subkind; in particular it is worth noting

[22] Pelletier & Schubert do not bother to spell out the typing system implicit in their theory, however.

that on this account "wine is a wine," which translates into $(\beta\ wine)(\mu\ wine)$, is false.

γ: $(\gamma\ P)$ designates conventionally recognised portions or servings of P. Suppose that b denotes a bottle of beer; then $(\gamma\ beer)(b)$ is true. A counter-intuitive but significant point about γ can be seen from the fact that the formula $(\gamma\ man)(john)$ is true, given that "John" is a (conventionally recognised portion of) man. In fact, "John is a man" has as its only translation in the translation relation of [PS86b]

$$[(\beta\ man)(john) \vee (\gamma\ man)(john)]$$

which is true because $(\gamma\ man)(john)$ is true. Predicates of the form $(\gamma\ P)$ correspond for most purposes exactly to the traditional reading of P as denoting a set of *individuals*. It is only for the very rare *conventional serving*-readings[23] that we need to think of portions and servings at all.

δ: $(\delta\ P)$ denotes a quantity of the stuff P, or an object coinciding with such a quantity. Thus "John drank sm water" (where "sm" is the unstressed "some") can be represented by

$$\exists x[(\delta\ water)(x) \wedge drank(john, x)]$$

which will be true just in case John drank a quantity of water (but possibly without this quantity being a quantity of a conventionally recognised kind of water). Furthermore, the representation of "John is sm man"

$$(\delta\ man)(john)$$

is true in the somewhat special sense of John being an object coinciding with an actual quantity of "the stuff *man*".

[23] Which according to Pelletier & Schubert are necessary to account for sentences like "this is a beer", where "this" denotes a bottle of beer.

As regards predicates, there is one more central notion, but this time having to do with the absence of the above operators:

Comprehensive extension:

A predicate P occurring unmodified by β, γ, or δ, is said to have *comprehensive extension*. Then, P is true of whatever $(\beta\ P)$, $(\gamma\ P)$, and $(\delta\ P)$ are true of, plus such kinds (and possibly such portions) which are not conventionally recognised, but are nonetheless kinds (or portions); for instance,

$wine(\mu\ wine)$, "wine is wine",
$wine(\mu\ (cheap\ wine))$, "cheap wine is wine",

are true.

6.2 The Underlying Formal System: The Semilattice of Kinds

Given a common noun M, the theory views the kinds of M as forming an upper semilattice with M at the top. That is, M "itself" is taken to be the most general kind of M.

A semilattice is an algebraic structure with the following properties:

domain: there is a partially ordered domain A.

operators: there is one operator '$\vee$' (read "*union*").

predicates: there is one relational predicate '$\preceq$', which is transitive and reflexive. Given that A is partially ordered this defines a further predicate '$=$' such that $[a = b] \leftrightarrow [(a \preceq b) \wedge (b \preceq a)]$.

axioms: 1. $a \preceq (a \vee b)$
2. $b \preceq (a \vee b)$

rule: $[(a \preceq c) \wedge (b \preceq c)] \rightarrow [(a \vee b) \preceq c]$.
This rule guarantees that any two members a and b of A have a unique least upper bound (namely $(a \vee b)$).

Applying this to "the kinds of M", the following observations seem to be relevant:

1. The set A of kinds of M is partially ordered.

2. There is a "formal union-operator" '$\vee$'; there is no obvious intuition corresponding to '$\vee$', but tentatively you may think of it as "mixed with" (in the case of wines). For instance, "sauterne mixed with claret" is a kind of wine.

3. The predicate '$\preceq$' may be interpreted as "is a kind of". Thus the relation "is a kind of" is in the theory of Pelletier & Schubert seen as having the following important properties[24]:

   reflexivity: for any kind of wine a, a is a kind of a. In particular, "wine is a kind of wine" is true[25].

   transitivity: *a is a kind of b* and *b is a kind of c* implies *a is a kind of c*.

To sum this up, the common noun "wine" gives rise to an upper semi-lattice of all the kinds of "wine". The predicate *wine* will be true of any member of the semilattice, and in particular it will be true of the "union" of any two members of the semilattice (again, think of "sauterne mixed with claret").

The precise meaning of the operator μ above can now be given as follows: for any common noun M, $(\mu\ M)$ names the topmost element of the semilattice of M-kinds. That is, $(\mu\ M)$ is the "formal union" of all the members of the lattice. We may read this as "the most general kind of M" or "the substance M" [26].

---

[24] I may add that the same properties naturally are to hold for conventional kinds, too (although Pelletier & Schubert never explicitly say so—but they certainly rely on this being the case in their inferences).

[25] Not to be confused with the false "wine is a wine", which is translated into $(\beta\ wine)(\mu\ wine)$—that is, that sentence is interpreted as "wine is a conventionally recognised subkind of wine".

[26] It seems to me to be somewhat unclear how they would actually impose this structure upon the "kinds of M" in their semantics (when one goes on to give a full formal system). For we have to do with a "first-orderish semantics" with only one "quantificational" domain, and that domain is not in general an upper semi-lattice of the sort described here, to be sure. Therefore, it would probably be necessary to express the above properties of M-kinds by means of meaning postulates. See also comment 1, p. 33.

6.3 The Grammar and Translation Relation

The grammar and the translation relation are given[27] by the following *extended context-free phrase-structure grammar* rules ("**S&T** *i*" means "syntax-rule *i* and translation rule *i*"):

S&T 1. S $\Longrightarrow$ NP VP / $vp'(np')$

S&T 2. VP $\Longrightarrow$ [V +be] PRED / $pred'$

S&T 3. PRED $\Longrightarrow$ N / n'

S&T 4. PRED $\Longrightarrow$ [DET +a] N / $\lambda x[(\beta\ n')(x) \vee (\gamma\ n')(x)]$

S&T 5. PRED $\Longrightarrow$ NP / $\lambda x[x = np']$

S&T 6. PRED $\Longrightarrow$ ADJP / $adjp'$

S&T 7. NP $\Longrightarrow$ N/ $(\mu\ n')$

S&T 8. NP $\Longrightarrow$ DET N / $\langle det'\ n'\rangle$

S&T 9. [N +ADJP] $\Longrightarrow$ [ADJP +INT] N/ $\lambda x[adjp'(x) \wedge n'(x)]$

S&T 10. [N +ADJP] $\Longrightarrow$ [ADJP -INT] N/ $adjp'(n')$

T 11. sm' = $\lambda P\ \langle\exists;\ (\delta\ P)\rangle$

T 12. some' = $\lambda P\ \langle\exists;\ \lambda x[(\gamma\ P)(x) \vee (\beta\ P)(x)]\rangle$

T 13. all' = $\lambda P\ \langle\forall;\ P\rangle$

[27] See [PS86b, pages 82 and 84].

6.3.1 Comments

1. This grammar allows for complex syntactic nodes, e.g. [N +ADJP] indicates a noun which has the feature of being premodified by an adjective phrase.

2. The diamond bracket expression in T. 8—⟨*det' n'*⟩—indicates that the quantifier *det'* is restricted by the predicate *n'*, and that it will be assigned a scope in a "post-parsing phase"[28].

3. The determiners to be covered by the theory are explicitly enumerated in [PS86b, page 83]; they are: "this", "all", "some", "sm", "much", "little", "each", "every", and the numeral quantifiers. As for the determiner "a", see immediately below.

4. The rule **S&T 4** is a special rule for the determiner "a", which plays a crucial role in introducing *kind*-readings. Terms like "a wine" and "a man" by this rule translate into $\lambda x[(\beta\ wine)(x) \vee (\gamma\ wine)(x)]$ and $\lambda x[(\beta\ man)(x) \vee (\gamma\ man)(x)]$, respectively. Since there is no (syntactic) distinction between mass- and count-nouns in the *p-theory*, all nouns are dealt with by the same syntactic rules and hence follow the same "pattern" of translation. However, in most (or all) situations we need the "kind"-reading ($\beta\ wine$) for "a wine"[29] and the "portion"-reading (i.e. individual reading) ($\gamma\ man$) for "a man".

   Unfortunately, the existence of a particular rule for the indefinite article "a" causes some unclarity w.r.t. Pelletier & Schubert's intentions for "kind"-readings.

   Of course, "a" is also a quantifier, and in fact it is mentioned as such on a par with "some" in [PS86b, page 84]. But it would be to rash to conclude that Pelletier & Schubert have simply forgotten to mention "a" among their determiners, for if "a" is indeed a determiner, rule 4 would seem to become superfluous—save perhaps for expository purposes? At any rate, if one includes "a" among the determiners in the

[28]Nothing more is said about this phase—neither w.r.t. how it is expected to work nor how it should be implemented. However, in [PS86b], the reader is referred to [PS82] for more details on this subject. Unfortunately, the full details cannot be found there, either.

[29]See the further discussion of the inferences involving "a wine", "a claret" and "a liquid" in section 6.5.2.

lexicon, all sentences with the indefinite article become syntactically ambiguous in the present framework[30].

5. The predicate "is a kind of"—i.e. '$\preceq$'—does not occur in the translations given by Pelletier & Schubert's translation relation (and it will not be easy to make it occur, either). I think this is a weakness of the system. Consider, for instance, the translation of "wine is wine":

$$\textit{wine}(\mu\ \textit{wine})$$

The comprehensive predicate ***wine*** does not tell us that this is to be interpreted as "wine is a kind of wine"—the only way to know that is by considering the expression (μ *wine*), which is necessarily (by the definition of μ) a ***kind***. However, the logical target language actually allows us to express the intended reading more precisely by

$$(\mu\ \textit{wine}) \preceq (\mu\ \textit{wine})$$

whose analyticity is much more self-evident.

Now of course the two translations ***are*** equivalent within the present framework, but the first translation is the less precise one and a potential source of confusion; indeed this becomes clear when at a crucial point Pelletier & Schubert themselves forget or fail to notice that expressions of the form (μ P) always denote ***kinds*** (cf. footnote 40).

6.4 Meaning Postulates

The meaning postulates of [PS86b] are not given formally—and in fact most of them are not even stated in general terms, but rather they are indicated by way of example. Nonetheless, in listing their meaning postulates I shall stay very close to their original wording in [PS86b][31]. In most cases I shall add to the actual formulation a hint at how it may be generalised formally. (As far as I can see most inferences discussed by Pelletier & Schubert can be covered by a suitable formalisation of the following six meaning postulates, but for a complete theory of mass terms more meaning postulates will be needed).

---

[30] This does not have any semantic consequences, though; the translations will be identical on the different derivations. See also the discussion of sentence (2) in section 6.5 below.

[31] Meaning postulate MP-1 is found in [PS86b, page 79]; the rest of the meaning postulates are found in [PS86b, pages 86–87].

MP-0. "If x is a kind of M and a threw x at b [for M, take "snow"], then there is a y which is a quantity of M such that a threw y at b." [And similarly for verbs similar to "throw"].
"If a drinks a kind of M, then a drinks a quantity of M." [And similarly for verbs similar to "drink"].

So, for instance "John drank a wine" implies "there is a quantity of wine such that John drank it"; that is,

$$\models \exists x[(\beta\ \mathit{wine})(x) \wedge \mathit{drank}(j,x)] \rightarrow \exists y[(\delta\ \mathit{wine})(y) \wedge \mathit{drank}(j,y)]$$

MP-1. "If a predicate is not exclusively a kind predicate, and if it holds of a kind M, then it (1) holds of some (perhaps all) subkinds of M (recall that the theory takes kinds of M to form an upper semilattice), and (2) holds of at least some (perhaps all) quantities of any such subkind."

For instance, *liquid* is a predicate with comprehensive extension, and holds of the kind (μ *wine*); and, since *liquid* is true also of things other than kinds, e.g. "this puddle of wine" (which simply denotes a physical object), it is not exclusively a kind predicate. Therefore, by MP-1 it

(1) holds of (at least) some subkinds of "wine"
(2) holds of (at least) some quantities of such subkinds.

To sum this up in a slightly more formal way: given that *liquid* has the properties discussed here, we have for instance

$$(1)\quad \models\ \mathit{liquid}(\mu\ \mathit{wine}) \rightarrow \exists x[x \preceq (\mu\ \mathit{wine}) \wedge \mathit{liquid}(x)]$$
$$(2)\quad \models\ \mathit{liquid}(\mu\ \mathit{wine}) \wedge (\mu\ \mathit{claret}) \preceq (\mu\ \mathit{wine}) \rightarrow \exists x[(\delta\ \mathit{claret})(x) \wedge \mathit{liquid}(x)]$$

MP-2. "Sometimes there are relationships between the different types of predicates. For example "is a liquid" is true only of kinds while "is liquid" is true of both [kinds and objects]. The theory takes the view that in such cases the kind property entails the comprehensive property."

So, for instance, "water is a liquid" entails "water is liquid". Formally,

$$\models\ [(\beta\ \mathit{liquid})(\mu\ \mathit{water}) \vee (\gamma\ \mathit{liquid})(\mu\ \mathit{water})] \rightarrow \mathit{liquid}(\mu\ \mathit{water})$$

This rule is worth noting for the simple reason that it is heavily relied upon later. However, it is rather odd that it should be stated as a meaning postulate—this fact is discussed in detail below (cf. comment (3), page 33).

MP-3. "When the predicate in question is one of the comprehensive ones, the theory takes the view that if the kind has it, then at least some quantity of the subject has it."
Thus, "water is liquid" entails "sm water is liquid", or formally,

$$\models \mathit{liquid}(\mu\ \mathit{water}) \rightarrow \exists x[(\delta\ \mathit{water})(x) \wedge \mathit{liquid}(x)]$$

MP-4. "The theory takes the position that stative VP's of the comprehensive sort, such as "is liquid", "is wine", etc., induce a universal reading on the sentence."

Thus, "water is liquid" is equivalent to "all water is liquid"; formally,

$$\models \mathit{liquid}(\mu\ \mathit{water}) \leftrightarrow \forall x[\mathit{water}(x) \rightarrow \mathit{liquid}(x)]$$

As a consequence of this "water is liquid" also entails "all quantities of water are liquid".

MP-5. "Non-stative VP's of the comprehensive sort, such as "is dripping from the faucet" or "is lying on my desk", are assumed to induce an existential reading on the sentence."

Thus "water is dripping from the faucet" implies "sm water is dripping from the faucet". Formally (ignoring the "dripping from the faucet" part of the sentence)

$$\models \mathit{dripping}(\mu\ \mathit{water}) \rightarrow \exists x[(\delta\ \mathit{water})(x) \wedge \mathit{dripping}(x)]$$

6.4.1 Comments

1. It will be necessary to add meaning postulates for quite a few constructions. Just to take one example, it will be necessary to add a meaning postulate (or perhaps rather to modify MP-4) such that

   "water is abundant"

   will not be equivalent to

   "all water is abundant"

   and will not entail

   "all quantities of water are abundant".

   Actually, MP-4 is rather controversial. Most authors—for instance Parsons in his famous paper [Par79b]—see constructions of the form "water is abundant" as *elliptical*. On his account the sentence "water is abundant" might equally well be interpreted as "some water is abundant", "most water is abundant", etc. etc.

   It is also worth noting that inferences like "water is abundant, so unpolluted water is abundant" are not intuitively valid—or to take an example still closer to Pelletier & Schubert, "water is liquid, so frozen water is liquid".

   The problem is, however, that MP-4 is really a consequence of the assumption of an "underlying" semi-lattice structure, and cannot be eliminated as long as that assumption is upheld without modification. An obvious course of modification would be to state the "semi-lattice properties" that we want as meaning postulates and simply leave out the rest of those properties[32].

2. MP-0 does not come into use in the fragment of English given by the grammar of Pelletier & Schubert. I merely include it here to add to the overall understanding of the theory.

3. As mentioned above some clarification is needed regarding MP-2.

   First, it must be added to the wording of MP-2 that "is a liquid" is also at least potentially true of conventional portions—its translation being

$$\lambda x[(\beta\ liquid)(x) \vee (\gamma\ liquid)(x)]$$

[32] Essentially this is what is done in my axiomatisation of the system.

Second, MP-2 is somewhat surprising because it should not be necessary to state it at all, its content being already given by the definition of *comprehensive extension*[33] (see page 26).

4. MP-3 and MP-5 resemble each other so much that it may be confusing at first glance (in fact, the example which I give to illustrate MP-5 is used in [PS86b] to illustrate MP-3). However, although they do overlap in some cases, MP-3 is in general restricted to cases, where the subject phrase has a translation of the form (μ M)—this restriction does not apply to MP-5; and MP-5 is restricted to cases where the predicate translates some non-stative VP—which does not apply to MP-3.

5. The meaning postulates MP-1 to MP-5 make it sufficiently clear that we are currently dealing with "concrete" mass expressions only. In general it does not seem to make sense to speak of a "quantity" (or equivalently, a "portion") of "love", "software", etc. Thus an inference like "John loves software, so John loves some portion of software" can hardly be said to be valid[34].

6.5 Applications of the Theory

Now we are at last in a position to see what the theory of Pelletier & Schubert actually accomplishes with respect to the semantics of mass expressions. On

[33]MP-2 would only be non-trivial if it is taken to mean (in the case of our example) that "water is a liquid" indeed entails that

(a) either there are other water-kinds than those conventionally recognised, and at least some of these kinds are liquid

(b) or there are quantities of the kind water such that they are liquid.

To put it more technically, MP-2 is only a meaning postulate in its own right if it is interpreted as

necessesarily, for all predicates P in $\{liquid, wine, \ldots, \ldots\}$,
$\exists x[(\beta\ P)(x) \vee (\gamma\ P)(x)] \rightarrow \exists y[P(y) \wedge \neg(\beta\ P)(y) \wedge \neg(\gamma\ P)(y)]$

However, that seems to be an implausible interpretation and it is certainly not that way the meaning postulate is used in the inferences discussed later on.

[34]In fact it may be possible to distinguish between *abstract* and *concrete* mass nouns by reversing e.g. MP-1 on "throw" and similar verbs into a test: for instance, since

"John threw software at Mary"

does not imply

"There was a quantity of software such that John threw it at Mary"

we may conclude that "software" is an *abstract* mass noun.

the last pages of [PS86b], they list a number of sentences and inferences, making certain claims about the way their theory deals with those. Most of their claims are correct, and the tedious work of reconstructing their theory is now rewarded by being able to spell out in rigorous detail exactly why that is so in each case.

6.5.1 Some Sample Sentences

Before turning to the individual inferences it will be useful to discuss the logical representations and interpretations of some sample sentences—these interpretations arise from Pelletier & Schubert's framework as developed so far, and are also very briefly discussed in [PS86b, page 88]. I shall try to elucidate somewhat further the logical properties of those interpretations:

(1) Wine is wine.

By the grammar and translation relation, (1) has two translations[35], namely

> **(1.a)** $wine(\mu\ wine)$
> **(1.b)** $(\mu\ wine) = (\mu\ wine)$

(1.a) is analytically true—this is a trivial consequence of the definition of *comprehensive extension*: (μ *wine*) is the topmost member of *the kinds of wine*, which are in turn a subset of the comprehensive extension predicate *wine*, so (μ *wine*) is *a fortiori* a member of *wine*.

Likewise, (1.b) is obviously analytically true. These results are in accordance with [PS86b], which describes (1) as having "two readings, both

---

[35]These are produced by successive applications of the translation rules as it is seen in their derivations:

(1.a): S $\Longrightarrow$ NP VP (S.1), NP VP $\Longrightarrow$ N VP (S.7),
N VP $\Longrightarrow$ wine VP (Lexicon), wine VP $\Longrightarrow$ wine [V +be] PRED (S.2),
wine [V +be] PRED $\Longrightarrow$ wine is PRED (Lexicon),
wine is PRED $\Longrightarrow$ wine is N (S.3), wine is N $\Longrightarrow$ wine is wine

(1.b) S $\Longrightarrow$ NP VP (S.1), NP VP $\Longrightarrow$ N VP (S.7),
N VP $\Longrightarrow$ wine VP (Lexicon), wine VP $\Longrightarrow$ wine [V +be] PRED (S.2),
wine [V +be] PRED $\Longrightarrow$ wine is PRED (Lexicon),
wine is PRED $\Longrightarrow$ wine is NP (S.5), wine is NP $\Longrightarrow$ wine is N (S.7),
wine is N $\Longrightarrow$ wine is wine.

I shall not give the derivations for the rest of the sentences discussed, but these are quite easy to construct, provided that the constructor is willing to simulate the "post-parsing phase".

analytic".

(2) Wine is a wine.

Sentence (2) has only one translation,

$$(\textbf{2.a})\ [(\beta\ wine)(\mu\ wine) \vee (\gamma\ wine)(\mu\ wine)]$$

which is false, given that "wine" is neither a conventionally recognised subkind nor a conventionally recognised serving of wine. (But bear in mind that "wine is a kind of wine"[36], i.e. $(\mu\ wine) \preceq (\mu\ wine)$ is true in the semilattice of wine-kinds.)

In fact (2) has two different derivations if "a" is included among the determiners. Provided that the determiner "a" has the same translation as the determiner "some", and that the post-parsing phase does its job properly, the translations according to the two derivations will be identical[37].

(3) Chilled wine is a wine.

Provided that "chilled" is a non-intersective adjective in the lexicon, (3) has only one analysis tree and hence only one translation,

$$(\textbf{3.a})\ (\beta\ wine)(\mu\ (chilled\ wine)) \vee (\gamma\ wine)(\mu\ (chilled\ wine))$$

Given that "chilled wine" is neither a conventionally recognised subkind of wine, nor a conventionally recognised serving of wine, (3) must turn out to be false.

(4) All wine is wine.

One derivation, yielding the translation

$$\forall x[wine(x) \rightarrow wine(x)]$$

which is analytically true.

[36] Unfortunately this sentence is not included in the fragment covered by the grammar, so I have to stipulate its translation myself. However, my translation is implied by Pelletier & Schubert's discussion of the semilattice of kinds in general.

[37] If there really are two be two different derivation patterns for these cases, then those patterns must certainly yield semantically equivalent results.

(5) All wine is a wine.

Pelletier & Schubert claim that their theory would characterise (5) as semantically anomalous. To examine what may be meant by that claim, let us consider its (one and only) translation:

$$\textbf{(5.a)}\ \forall x[wine(x) \rightarrow [(\beta\ wine)(x) \vee (\gamma\ wine)(x)]]$$

Obviously, (5) must turn out to be false on this analysis, given the definition of *comprehensive extension* and our knowledge of the real world (which implicitly serves as a model for our sample sentences). In the real world there exist kinds of wine which are not conventionally recognised, and furthermore there exist quantities of wine—(δ *wine*)—which are neither servings nor conventionally recognised portions—so (5.a) is clearly false with respect to the real world.

It is less obvious why (5) should be called "semantically anomalous". The reason for saying so is, no doubt, that (5.a.) entails

$$\textbf{(5.b)}\ \forall x[(\delta\ wine)(x) \rightarrow [(\beta\ wine)(x) \vee (\gamma\ wine)(x)]]$$

But since δ, β, and γ technically are just ordinary predicate modifiers, (5.b.) should be no more anomalous than, say, "all small men are big men". So I conclude that at least in the present framework it is more natural to characterise this sentence as simply being inherently false rather than semantically anomalous.

6.5.2 Some Exemplary Inferences

Pelletier & Schubert state eight inferences ([PS86b, page 89]), which are meant to highlight the properties of their system. With the reconstruction of their theory given in the previous pages we can now scrutinise the claims they make about the various inferences. Before doing so, however, there remains one last but rather radical piece of reconstruction to be done:

In order to get the discussion of the eight inferences right, it will be necessary to revise the translation of predicates of the form "is a wine" from

$$\lambda x[(\beta\ wine)(x) \vee (\gamma\ wine)(x)]$$

into just

$$\lambda x[(\beta\ wine)(x)]$$

When considering inferences which involve such predicates, Pelletier & Schubert take it for granted that we can avail ourselves of the properties of the semilattice of wine-kinds; but unfortunately that presupposition is unwarranted with their translation relation. To see this, consider a sentence like "a is a wine": if "a is a wine" translates into $(\beta\ wine)(a)$, then the truth of this sentence will guarantee that a is in the semi-lattice of kinds. On the other hand, the truth of

$$[(\beta\ wine)(a) \vee (\gamma\ wine)(a)]$$

carries no such guarantee[38].

I shall not discuss the further implications of this revision here[39], but I do suggest that it is bound to have rather radical semantic consequences for Pelletier & Schubert's theory.

With this modification of the translation relation in mind we can now turn to the inferences.

(1) Claret is a wine, wine is a liquid, so claret is a liquid.

Formally, this inference can be rendered by

$$\models \ (\beta\ wine)(\mu\ claret)$$
$$\models \ (\beta\ liquid)(\mu\ wine)$$
$$\models \ (\beta\ liquid)(\mu\ claret)$$

This is valid because we are in a semilattice of kinds—in fact, a sub-semilattice of conventionally recognised kinds—and the predicate "is a (conventionally recognised) kind of" is transitive.

Thus the theory can account for the validity of this normally recalcitrant type of inference.

(2) Claret is a wine, wine is a liquid, so claret is liquid.

It has already been established immediately above that the two premises entail "claret is a liquid". By MP-2, "claret is a liquid" entails "claret is liquid". Therefore, (2) is a valid inference.

---

[38]That formula might be true simply in virtue of $(\gamma\ wine)(a)$ being true; then a is *not* in the semi-lattice of kinds, but rather a is an individual object representing a conventionally recognised serving of wine (compare with "this bottle of beer is a beer", which is true in exactly that sense).

[39]See section 8.2.1 for a bit further discussion of the implicated problems.

(3) Claret is a wine, wine is liquid, so claret is a liquid.

Pelletier & Schubert claim of (3) that it is valid because "the first premise ... asserts that Claret is a kind (of wine) ... [and] we are told that wine is liquid, so ... claret is a liquid-kind" ([PS86b, page 89]). However, "claret is a liquid-kind" asserts less than "claret is a liquid", and indeed the conclusion does not follow. And of course it shouldn't—for it may be that wine is a kind of liquid and yet not a conventionally recognised kind of liquid, in which case claret may or may not be a conventionally recognised kind of liquid. Thus, (3) is not a valid inference.

(4) Claret is a wine, wine is liquid, so claret is liquid.

By MP-4, "wine is liquid" is equivalent to "all wine is liquid". Thus, the inference is rendered formally by

$$\models \ (\beta\ wine)(\mu\ claret)$$
$$\models \ wine(\mu\ claret)$$
$$\models \ \forall x[wine(x) \rightarrow liquid(x)]$$
$$\models \ liquid(\mu\ claret)$$

which is valid.

(5) Claret is wine, wine is a liquid, so claret is a liquid.

In this case, the first premise does not tell anything about the status of claret with respect to its being conventionally recognised. For this reason (5) is invalid[40] (this can perhaps better be perceived by substituting "cheap wine" for "claret" in (5)).

(6) Claret is wine, wine is a liquid, so claret is liquid.

By MP-2, "wine is a liquid" entails "wine is liquid". By MP-4, "wine is liquid" is equivalent to "all wine is liquid", and "claret is wine" is equivalent to "all claret is wine". Formally we have

[40]Pelletier & Schubert remark that "Here we are not given in the first premise that claret is a kind – for all we are told, it might be a name or description of a quantity or object..." ([PS86b, page 89]). But this is certainly wrong, since "claret" in this context translates into (*μ claret*), which by definition names the *kind* claret. And this is also the intention of Pelletier & Schubert: "When a bare mass noun phrase (or indeed other bare noun phrases, although we shall not dwell on them here) is used as a subject...it is taken to name the kind." ([PS86b, page 75]).

$$\models \forall x[claret(x) \rightarrow wine(x)]$$
$$\models \forall x[wine(x) \rightarrow liquid(x)]$$
$$\models \forall x[claret(x) \rightarrow liquid(x)]$$

which is a valid logical inference.

(7) Claret is wine, wine is liquid, so claret is a liquid.

Since (5) has already been shown to be invalid, (7) is *a fortiori* also invalid.

(8) Claret is wine, wine is liquid, so claret is liquid.

Of course, (8) is valid—this follows trivially by applying MP-4 to both premises, which produces the same inference as in (6) above.

7 A Partial Implementation of the Theory

7.1 A Revised Grammar and Translation Relation

We can now turn to the implementation of the system. This will proceed in two phases: first a tiny translation relation from a fragment of natural language into the target logical language is given. The fragment will contain those crucial syntactic constructions from which inferences are to be made. The next phase will be the axiomatisation of those aspects of the theory which we need for the inferences.

The system given in the following sections is extremely limited—it only works for a lexicon which is restricted to contain singular mass nouns only. However, it does implement that crucial core of the theory which has been thoroughly discussed so far. Once this core is "in place" it will be much easier to extend the system to other cases[41]. At any rate, the system manages to give a reconstruction of the theory in the sense that it makes it possible to follow the desirable inferences rigorously in each and every detail: from their natural language premises into their translations, and from those logical forms into a deductive proof of the desired consequence—which is in turn a translation of the intended natural language consequence.

[41]Several suggestions of how to do this are already given in the comments below. In **Appendix D** you will find a CFG+ grammar which makes use of all the notions developed here, but which further incorporates count nouns and plural noun phrases with both kinds of nouns.

A compiled and tested version of the grammar below can be found in **Appendix A**.

7.1.1 Some Important Translations

1. a'=some'= $\lambda P \lambda Q \exists x[(\beta\ P)(x) \wedge Q(x)]$

2. be'= $[\lambda \mathcal{P} \lambda x \mathcal{P}(\lambda y\ [y = x])]$
   (This is simply the [PTQ]-translation of "be", minus intensionality. Here, $\mathcal{P}$ is of type $\langle\langle e,t\rangle,t\rangle$, i.e. $\mathcal{P}$ ranges over terms).

7.1.2 The Rules

1. S $\Longrightarrow$ NP PRED / $[np'(pred')]$
2. NP $\Longrightarrow$ NBAR / $\lambda P[P(\mu\,(nbar'))]$
3. NP $\Longrightarrow$ DET NBAR / $[det'(nbar')]$
4. NBAR $\Longrightarrow$ N / n'
5. NBAR $\Longrightarrow$ ADJ NBAR / $\lambda x[adj'(x) \wedge nbar'(x)]$
6. PRED $\Longrightarrow$ VP / vp'
7. VP $\Longrightarrow$ V-BE ADJP / $adjp'$
8. ADJP $\Longrightarrow$ ADJ / adj'
9. VP $\Longrightarrow$ V-BE NP / $[v\text{-}be'(np')]$
10. PRED $\Longrightarrow$ V-BE NBAR / $nbar'$

7.1.3 Comments

This grammar differs from the one given by Pelletier & Schubert in several respects. The following points are important:

1. The post-parsing phase has been eliminated, so it is possible with this grammar to follow every step of a translation rigorously.

2. Following the tradition from Montague, I apply *np'*(*pred'*) rather than *pred'*(*np'*). The rather crucial rule 2 reflects this by translating a singular bare noun (phrase) N into $\lambda P[P(\mu\ N')]$. (For a note on bare plural noun phrases, see below).

3. The β-operator is introduced by the translations of the determiners "a" and "some"[42]. In an extended fragment, other determiners will have to do the same thing, ***when they combine with mass nouns***. Thus these translations in fact anticipate that nouns—and indeed determiners—***will*** be equipped with features such as *+mass* and *-mass*. Of course, this is a major departure from the *p-theory*, and it is motivated by the fact that we need "pure" *kind*-translations of the form $(\beta\ P)$ to get the inferences right. I shall return to this point in section 8.

   As a "complementary consequence" the translation relation does not employ the γ-operator of Pelletier & Schubert; but when the fragment is extended to incorporate ***count nouns*** such as "man", "bottle" etc., γ can of course be introduced—simply by letting a determiner subcategorised as *-mass* employ that operator where its *+mass* counterpart introduces β.

Apart from these differences between my grammar and that of Pelletier & Schubert there are a few more general points to be made for the above grammar:

adjectives: The grammar rules which deal with adjectives are here restricted to *intersective* adjectives only.

As for the question of whether there are any "truly intersective" adjectives at all, notice that we could easily "parameterise" rule 5 into having a translation of the form

$$\lambda x[\mathit{adj}'\text{-relative-to-}\mathit{nbar}'(x) \wedge \mathit{nbar}'(x)]$$

Then an expression like "cheap claret" would (at the "NBAR-level") be rendered by

$$\lambda x[\mathit{cheap}\text{-relative-to-}\mathit{claret}(x) \wedge \mathit{claret}(x)]$$

[42]The translation of "a"/"some" is perhaps surprising in only applying β to the "first incoming" predicate—but this has to be so to work properly for the semantics, since the next predicate may be a *comprehensive extension* predicate, as in "a liquid is wine"; see the translations of this sentence on p. 43.

In this connection let me point out that expressions of the form $(\mu\ (\lambda x[\phi(x)]))$ are of course well-formed (although we see no examples of them in [PS86b]). Thus "cheap claret" as a bare noun phrase translates into

$$(\mu\ (\lambda x[\textit{cheap}\text{-relative-to-}\textit{claret}(x) \wedge \textit{claret}(x)))$$

(given the suggested parameterisation of intersective adjectives).

plural noun phrases: When plurals and in particular bare plural noun phrases are involved rule 2 will of course have to be modified. For instance we could have a "dual rule" using syntactic features[43]; e.g.

(2.a) NP[agr singular] $\Longrightarrow$ NBAR[agr singular] / $\lambda P[P(\mu\ (\textit{nbar}'))]$
(2.b) NP[agr plural] $\Longrightarrow$ NBAR[agr plural] / τ

where τ stands for the intended interpretation of the plural noun phrase.

the be-of-predication: While rule 9 together with the translation of "be" provides the traditional *be-of-identity*-reading, rule 10 takes care of the *be-of-predication*—which in our context in turn yields the *comprehensive extension*-reading of predicates such as "is wine".

It is worth noting that a sentence like "a liquid is wine" is ambiguous between the following two translations[44]:

a liquid is wine
$\trianglerighteq_1$ $\exists x[(\beta\ \textit{liquid})(x) \wedge \textit{wine}(x)]$
$\trianglerighteq_2$ $\exists x[(\beta\ \textit{liquid})(x) \wedge x = (\mu\ \textit{wine})]$

On translation 2, the sentence is equivalent to the unambiguous "wine is a liquid". I think that this is as it should be.

[43] In **Appendix D** a complete solution to this particular problem can be found.

[44] For that reason it is actually somewhat misleading to simply subsume a predicate like "is wine" under the *comprehensive extension*-predicates, as Pelletier & Schubert sometimes do; that neglects the fact that such predicates also have *be-of-identity*-readings. (I have nonetheless followed them in this in the discussion of the meaning-postulates in order to avoid too many forward references.)

7.2 Implementing the Required Semilattice-Properties

In the following, let $\mathbf{P}_A$[45] denote the set of phrases involving mass nouns such as "claret", "wine", "liquid", "rich wine", "rich old wine" etc. Let $\mathrm{ME}_{\langle a \rangle}$ denote the corresponding set of meaningful expressions of type $\langle a \rangle$ in the target language of logical forms. Then those properties of Pelletier & Schubert's theory which are relevant to the inferences 1–8 are implemented by the following definitions and axioms[46]:

Definition 1 (General Kind)
For all P in $\mathrm{ME}_{\langle a \rangle}$*, for all x:*
$(\varsigma\ P)(x)$ *if and only if* x **is a kind of** P.

Example:
$(\varsigma\ wine)(x)$ **if and only if** x is a kind of wine.

Comment:
This definition introduces a new symbol ς, to be read "***is a kind of***". The reason for introducing ς is to be able to explicitly impose (some of) the assumed "semi-lattice properties" on "kinds"[47].

Definition 2 (Comprehensive Extension)
For all P in $\mathrm{ME}_{\langle a \rangle}$: $\forall x[P(x) \leftrightarrow_{\mathrm{def}} (\gamma\ P)(x) \vee (\delta\ P)(x) \vee (\varsigma\ P)(x)]$

Example:
x is liquid **if and only if** x is a portion of liquid **or** x is a quantity of liquid **or** x is a kind of liquid.

Comment:
This definition is not entirely uncontroversial[48], since it restricts the *comprehensive extension* of P to be a mere union of portions (γ), quantities (δ), and kinds (ς) of P. In turn, the ontology would seem to be restricted to comprise these sorts of entities only.

---

[45] In the present fragment we can let $A = CN$. But when count nouns are included too we shall have to differentiate between a category of basic count nouns, say CCN, and a category of basic mass nouns, say MCN. Then $A = MCN$.

[46] Or more precisely, *axiom schemas*.

[47] In Pelletier & Schubert's theory those properties are simply assumed to hold, without any suggestion of how to implement this assumption.

[48] As it has been pointed out in the discussion of Pelletier & Schubert's MP-2, it seems that Pelletier & Schubert foresee that it may be necessary to introduce other sorts of entities than those mentioned here. On the other hand, the explicit statement of their ontology in [PS86b, pages 73–74] is compatible with Definition 2.

However, in the present context this definition is rather helpful, and I see no reason to believe that other types of entities are needed as possible denotations of expressions in $ME_{\langle a \rangle}$.

Notice that the ***conventionally recognised subkinds of*** P —i.e. $(\beta\ P)$— must be included in $(\varsigma\ P)$ (which is ensured by axiom 1 immediately below).

Axiom 1 *For all P in* $ME_{\langle a \rangle}$: $\forall x[(\beta\ P)(x) \rightarrow (\varsigma\ P)(x)]$

Example:
if x is a conventional kind of wine, **then** x is a kind of wine.

Comment:
Axiom 1 simply states that if x is a conventional kind of P, then x is a kind of P. Since most of the "semi-lattice properties" for kinds will be stated in terms of axioms etc. which involve ς, this axiom enables us to apply these properties with respect to conventional kinds too. (To appreciate the importance of this, see for instance the proof of inference 2 on p. 47.)

Axiom 2 (Transitivity for General Kinds)
For all P, Q in $ME_{\langle a \rangle}$: $\forall x[(\varsigma\ P)(x) \wedge (\varsigma\ Q)(\mu\ P) \rightarrow (\varsigma\ Q)(x)]$

Example:
if x is a kind of wine **and** wine is a kind of liquid, **then** x is a kind of liquid.

Comment:
This axiom implements the transitivity-property of the semilattice of kinds.

Axiom 3 (Transitivity for Conventional Kinds)
For all P, Q in $ME_{\langle a \rangle}$: $\forall x[(\beta\ P)(x) \wedge (\beta\ Q)(\mu\ P) \rightarrow (\beta\ Q)(x)]$

Example:
if x is a conventional kind of wine **and** wine is a conventional kind of liquid, **then** x is a conventional kind of liquid.

Comment:
This axiom implements the transitivity-property of the semilattice of kinds with respect to a sub-lattice of conventional kinds.

Axiom 4 (Reflexivity) *For all P in* $\mathrm{ME}_{\langle a\rangle}$*:* $(\varsigma\ P)(\mu\ P)$

Example:
wine is a kind of wine.

Comment:
This axiom implements the reflexivity-property of the semilattice of kinds.

Axiom 5 *For all P, Q in* $\mathrm{ME}_{\langle a\rangle}$*:* $Q(\mu\ P) \leftrightarrow (\varsigma\ Q)(\mu\ P)$

Example:
claret is wine if and only if claret is a kind of wine.

Comment:
By its definition $(\mu\ P)$ always denotes a *kind of P*. Therefore $(\mu\ P)$ is to be in the *comprehensive extension of Q* if and only if $(\mu\ P)$ is in *the kinds of Q*.

Theorem 1 *For all P in* $\mathrm{ME}_{\langle a\rangle}$*:* $P(\mu\ P)$

Proof 1 *Immediately from Axiom 5 and Axiom 4*

Theorem 2 *For all P, Q in* $\mathrm{ME}_{\langle a\rangle}$*:* $[(\mu\ Q) = (\mu\ P)] \rightarrow Q(\mu\ P)$

Proof 2

$\vdash\ Q(\mu\ Q)$	theorem 1
$\vdash\ (\mu\ P) = (\mu\ Q)$	premise
$\vdash\ Q(\mu\ P)$	Leibnitz' Law
$\vdash\ [(\mu\ Q) = (\mu\ P)] \rightarrow Q(\mu\ P)$	Deduction Theorem

7.3 The Natural Language Inferences

1. Claret is a wine, wine is a liquid, so claret is a liquid.

Proof:

$\vdash$	$(\beta\ wine)(\mu\ claret)$	premise
$\vdash$	$(\beta\ liquid)(\mu\ wine)$	premise
$\vdash$	$(\beta\ wine)(\mu\ claret) \wedge (\beta\ liquid)(\mu\ wine)$	conjunction
$\vdash$	$(\beta\ wine)(\mu\ claret) \wedge (\beta\ liquid)(\mu\ wine) \rightarrow$ $(\beta\ liquid)(\mu\ claret)$	axiom 3
$\vdash$	$(\beta\ liquid)(\mu\ claret)$	detachment

2. Claret is a wine, wine is a liquid, so claret is liquid.

Proof:

	$\vdots$	(as above)
$\vdash$	$(\beta\ liquid)(\mu\ claret)$	
$\vdash$	$(\beta\ liquid)(\mu\ claret) \rightarrow (\varsigma\ liquid)(\mu\ claret)$	Axiom 1
$\vdash$	$(\varsigma\ liquid)(\mu\ claret)$	detachment
$\vdash$	$(\varsigma\ liquid)(\mu\ claret) \leftrightarrow liquid(\mu\ claret)$	Axiom 5
$\vdash$	$liquid(\mu\ CLARET)$	replacement

(The consequence "claret is liquid" is ambiguous between the readings corresponding to the translations

"claret is liquid"
$\trianglerighteq_1$ $liquid(\mu\ claret)$
$\trianglerighteq_2$ $(\mu\ claret) = (\mu\ liquid)$

What has been proved here is of course only the validity of the inference with respect to one of these readings. In general I shall only examine more than one reading of the consequence, if it has more than one reading which makes the inference valid.)

3. Claret is a wine, wine is liquid, so claret is a *kind of*[49] liquid.

Since the premise "wine is liquid" is ambiguous, there are two inferences to consider, depending on the translation of this sentence:

(a) "wine is liquid" $\trianglerighteq$ $liquid(\mu\ wine)$

Proof:

$\vdash\ liquid(\mu\ wine)$	premise
$\vdash\ (\varsigma\ liquid)(\mu\ wine)$	Axiom 5 + replacement
$\vdash\ (\beta\ wine)(\mu\ claret)$	premise
$\vdash\ (\varsigma\ wine)(\mu\ claret)$	Axiom 1 + detachment
$\vdash\ (\varsigma\ wine)(\mu\ claret) \wedge (\varsigma\ liquid)(\mu\ wine)$	conjunction
$\vdash\ (\varsigma\ liquid)(\mu\ claret)$	Axiom 2 + detachment

(b) "wine is liquid" $\trianglerighteq$ $(\mu\ wine) = (\mu\ liquid)$

Proof:

$\vdash\ (\mu\ wine) = (\mu\ liquid)$	premise
$\vdash\ [(\mu\ wine) = (\mu\ liquid)] \rightarrow liquid(\mu\ wine)$	Theorem 2
$\vdash\ liquid(\mu\ wine)$	detachment
$\vdots$	(as above)
$\vdash\ (\varsigma\ liquid)(\mu\ claret)$	

4. Claret is a wine, wine is liquid, so claret is liquid.

Proof:

$\vdots$	(as above)
$\vdash\ (\varsigma\ liquid)(\mu\ claret)$	
$\vdash\ (\varsigma\ liquid)(\mu\ claret) \leftrightarrow liquid(\mu\ claret)$	Axiom 5
$\vdash\ liquid(\mu\ claret)$	replacement

(We can disregard the ambiguity of "wine is liquid", since it has already been shown above that each of the possible readings carries to the same conclusion, when it is conjoined with the premise "claret is a wine".)

[49]In order to get this inference right I have inserted the *kind of*-part. Pelletier & Schubert claim that it is valid in the stronger sense without this qualification, i.e. that "claret is a liquid" follows from the premises. This claim is incorrect, both with respect to their own description of their system, and with respect to my implementation of it. (Cf. the discussion of this inference in section 6.5.2). However, there is a weaker "version" of the inference which *is* valid, namely the one which I am studying here. For these reasons I have considered only the validity of the inference as modified by me.

5. Claret is wine, wine is a liquid, so claret is a liquid.

This inference is invalid, as also observed by Pelletier & Schubert.

6. Claret is wine, wine is a liquid, so claret is liquid.

Here the inference once again involves ambiguity of a *premise*, so we shall have to consider two cases:

(a) "claret is wine" $\trianglerighteq$ $wine(\mu\ claret)$

Proof:

$\vdash\ wine(\mu\ claret)$	premise
$\vdash\ (\varsigma\ wine)(\mu\ claret)$	Axiom 5 + replacement
$\vdash\ (\beta\ liquid)(\mu\ wine)$	premise
$\vdash\ (\varsigma\ liquid)(\mu\ wine)$	Axiom 1 + detachment
$\vdash\ (\varsigma\ wine)(\mu\ claret) \wedge (\varsigma\ liquid)(\mu\ wine)$	conjunction
$\vdash\ (\varsigma\ liquid)(\mu\ claret)$	Axiom 2 + detachment
$\vdash\ liquid(\mu\ claret)$	Axiom 5 + replacement

(b) "claret is wine" $\trianglerighteq$ $(\mu\ wine) = (\mu\ claret)$

Proof:

$\vdash\ (\mu\ claret) = (\mu\ wine)$	premise
$\vdash\ wine(\mu\ claret)$	Theorem 2 + detachment
$\vdots$	(as above)
$\vdash\ liquid(\mu\ claret)$	

It is worth noting that the inference actually is stronger than observed by Pelletier & Schubert. In both 6(a) and 6(b) we have as an interim result

$$\vdash (\varsigma\ liquid)(\mu\ claret)$$

—which means that from the premises we can infer not only that "claret is liquid", but also that "claret is a kind of liquid" (in fact the latter entails the former).

I have pointed out earlier that Pelletier & Schubert fail to give any translation for "is a kind of". That may be the reason why the stronger version of (6) is not mentioned in their paper. Whichever way it is, we can add to the valid inferences

6': Claret is wine, wine is a liquid, so claret is a kind of liquid.

7. Claret is wine, wine is liquid, so claret is a liquid.

This inference is also invalid.

8. Claret is wine, wine is liquid, so claret is liquid.

There are four cases to consider, depending on the translations of the premises (the respective combinations of translations will be obvious from the proofs):

(a)

Proof:

$\vdash$ *wine*(μ *claret*)	premise
$\vdash$ (ς *wine*)(μ *claret*)	Axiom 5 + replacement
$\vdash$ *liquid*(μ *wine*)	premise
$\vdash$ (ς *liquid*)(μ *wine*)	Axiom 5 + replacement
$\vdash$ (ς *wine*)(μ *claret*) $\wedge$ (ς *liquid*)(μ *wine*)	conjunction
$\vdash$ (ς *liquid*)(μ *claret*)	Axiom 2 + detachment
$\vdash$ *liquid*(μ *claret*)	Axiom 5 + replacement

(b)

Proof:

$\vdash$ (μ *claret*) = (μ *wine*)	premise
$\vdash$ *wine*(μ *claret*)	Theorem 2 + detachment
$\vdots$	(the rest as (a))
$\vdash$ *liquid*(μ *claret*)	

(c)

Proof:

$\vdash$ (μ *liquid*) = (μ *wine*)	premise
$\vdash$ *liquid*(μ *wine*)	Theorem 2 + detachment
$\vdots$	(the rest as (a))
$\vdash$ *liquid*(μ *claret*)	

(d)

Proof:

$\vdash$ (μ *claret*) = (μ *wine*)	premise
$\vdash$ (μ *wine*) = (μ *liquid*)	premise
$\vdash$ (μ *claret*) = (μ *liquid*)	Leibnitz' Law

8 Conclusions on the Theory of Mass Terms

It has become a widely held opinion that formal semantics for natural language should not concern itself with "real" metaphysics but rather with the metaphysics of concrete languages only. For instance, one should study the metaphysics (or ontology) semantically inherent in English "as it is" without being diverted by qualms about the philosophical acceptability of such metaphysics. In a sense, such a view represents a strictly linguistic and empirical conception of the project of formal semantics for natural language as opposed to a more philosophical conception. This view implies that the semantics of natural language should—ideally at least—be inferred from a large number of empirical observations and be in accordance with such data. I shall call this conception of the project of formal semantics for natural language the *empiricist conception*[50].

It is not too hard to see the attractions of this conception of formal semantics. First, it sets a standard for serious semantic discussion: in view of the successes of that project since the advent of [PTQ], it is no longer enough to vent vague philosophical ideas in the realm of formal semantics; to be taken seriously one must present well-defined grammars, translation relations[51] etc. Second, the view implies the autonomy of formal semantics viz-a-viz philosophy: the former is a scientific empirical discipline not answerable to the latter, so one need not be distracted in one's work by philosophical objections—indeed, at least in principle one needs not know

[50] Unfortunately, this term may be slightly misleading. (Like any other reasonably short term I could think of.) The intention is to stress the importance given to empiricial linguistic data by this methodological conception—as opposed to certain more philosophically motivated approaches. But the latter need not be "un-empirical" or incompatible with philosophical empiricism. Indeed that conception which I have called the *empiricist conception* may be more open to platonistic elements in the semantics of language, simply because language actually seems to imply such notions, whether one likes it or not. An example of this will be seen in the discussion of G. Link's methodological views in section 11.2.

[51] I believe that this is what Pelletier & Schubert have in mind with the concluding remark of their paper:

> "The time for studies of mass expressions with only casual reference to the syntax and semantics of language is past. Only systematic attempts to account for large classes of mass expressions within formal syntactic-semantic-pragmatic frameworks can hope to resolve the remaining issues."
>
> [PS86b, page 91]

anything about philosophy[52].

So, on this conception one poses oneself a harder task in one respect and rids oneself of a task in another respect.

One very important consequence of the view is that *regimentation* becomes unacceptable—it simply ceases to be a legitimate part of the methods available to the formal semanticist. Of course, it it still legitimate to study restricted fragments, but within the chosen fragment all linguistic phenomena must be accounted for, even such uses of language which are very far-fetched (but still possible in some conceiveable circumstances).

8.1 Pelletier & Schubert on Methodology

Pelletier & Schubert obviously adhere to the *empiricist conception* of formal semantics. Let me try to trace their views in some depth. In [Pel79][53] Pelletier sets out to investigate the relation between the linguistic mass/count distinction and the corresponding philosophical sortal/non-sortal distinction (where *sortal* approximately corresponds to *count*, and *non-sortal* to *mass*). In the course of his investigation Pelletier remarks that

> "Part of the problem here seems to be that we want to cling to the grammatical distinction because it stands some chance of being clearly made ..."
>
> [Pel79, page 9]

This is, I believe, a very representative remark of the *empiricist conception* of the project of formal semantics. It leads on to the almost programmatic conclusion that

> "Thus the goal should now be to show how to "generate" the [philosophical] sortal/non-sortal distinction from the [linguistic] mass/count distinction."
>
> [Pel79, page 9]

[52] Although a few formal semanticists seem to be almost scornful of philosophical considerations applied to the description of language, the view does not in general imply any such attitude, nor, for that matter, that philosophy and formal semantics are irrelevant to one another. Thus for instance Pelletier & Schubert are clearly interested in philosophical consequences of formal semantics and make specific statements about their relation to each other, as you will see soon.

[53] The views there are consistent with the views set forth in [PS86b], so although Schubert is not co-author of that paper I think it must be all right to start there; in fact one is also referred in [PS86b, page 4] to that paper for some philosophical background.

These views are closely reflected in [PS86b]. That paper opens with the following echo of the above programme, with particular respect to the problem domain of course:

> "Problems associated with mass expressions can be divided into the following general areas: (1) distinguishing a class of mass expressions, (2) describing the syntax of this class, (3) describing the formal semantics of this class, (4) explicating the ontology such a class of expressions presupposes, and (5) accounting for various epistemological issues involving our perception of the ontology."
>
> [PS86b, page 1]

Obviously these points are enumerated in a sort of methodological order: each point methodologically precedes the next point, and so we may take it that points (4)-(5) are to be "generated" from points (1)-(3). I think this programme clearly falls under the broader programme of the *empiricist conception* for formal semantics. Various other remarks throughout the paper indicate the same ideas, culminating in the explicit statement that

> "The ontological underpinnings of both theories [the *p-theory* and the *s-theory*] are that "reality" contains two sorts of items: (1) "ordinary objects" such as rings, sofas, puddles ... (2) "kinds of stuff" ... and "kinds of portions"... **We wish to make no special metaphysical claims about the relationships that might hold between "ordinary objects" and "kinds"—instead we content ourselves with describing how such an ontology leads to a simple and natural description of various of the facts concerning mass ... expressions."**
>
> [PS86b, pages 73–74]

Of course, there is no denial that there is an ontology presupposed—but the motivation for that ontology is not to be found in any "special metaphysical claims" but rather in the way in which it leads to an adequate description of linguistic facts. In other words, they are concerned with "the ontology of English" rather than "real ontology".

These are the methodological views underlying Pelletier & Schubert's attempt at developing a theory of mass terms. Below, I shall (cautiously) try to assess the practical impact of those views upon the development of their theory.

8.2 Achievements and Limitations of the Theory

The theory of Pelletier & Schubert on mass terms (or more correctly, mass expressions) is in my opinion a real step forward within the problem domain. Just how much has been achieved by the theory is nicely highlighted, I think, by comparing the results so far with the following remarks by G. Link:

> "Substances are abstract entities and cannot be defined in terms of their concrete manifestations. The question, then, is of what kind the connections are that are intuitively felt between substances and their quantities. Take water, for instance. A quantity is water if it displays the internal structure of water, that is H_2O. But this relation is not a logical one. Or else we might look for substance properties which carry over to the quantities of the substance in question. Water is a liquid and yet, all concrete water might be frozen. So we have to go over to dispositional properties, getting more and more involved into our knowledge of the physical world ... What I am getting at is that *nominal mass terms do not seem to have a proper logic*".
>
> [Lin83, pages 305–306]

I shall not here take issue with any details of Link's argument. But as regards his conclusion that nominal mass terms do not seem to have a proper logic, I think that that is exactly what the theory of Pelletier & Schubert provides. True, there are various limitations to their theory. The fragment discussed so far is obviously a very limited one. But, in the best of Montague-tradition the problems which they have addressed are of a paradigmatic and fundamental nature, and once those problems have been solved it is very much easier to give a full theory for a substantial fragment of language involving mass expressions[54].

Of course, that does not mean to say that there are no qualitative problems left. I shall mention three points:

- The theory hardly addresses the problems of *abstract mass nouns*. It may be that these can be subsumed under the general framework of the theory, but if that is so it certainly remains to be shown[55].

[54]My statement of the meaning postulates in section 6.4, the comments on the revised grammar of section 7.1, and the fuller translation relation of **Appendix D** should all serve to substantiate this point.

[55]Unless, of course, one assumes that abstract mass nouns do not pose any problems of their own—i.e., that the *abstract-concrete* distinction is spurious w.r.t. mass terms. But that can hardly be the case; see also my remark on the meaning postulates on page 34.

- It remains to be more thoroughly scrutinised just how much of the power of the underlying semi-lattice we actually need. As I have pointed out meaning postulate 4 leads to unacceptable consequences, but that meaning postulate is in turn a consequence of assuming an underlying semi-lattice. On the other hand, if we acknowledge that the assumption of an underlying semi-lattice is too strong, we shall have to sort out which "semi-lattice-properties" we *do* want, and how to implement those.

- The treatment of nominal mass terms by means of the μ-operator relies on making a radical semantic abstraction which in my opinion takes us far beyond psychological reality. A sentence like "John threw snow at Mary" is always interpreted as meaning "John threw *the kind* snow at Mary"[56]—which is, I believe, an obvious abstraction from what speakers mean and hearers understand by that sentence. There may be nothing wrong in making such an abstraction, as long as it leads to all and only the intuitively valid inferences. But it does raise some questions in relation to the methodology which Pelletier & Schubert have professed; I shall address this issue below.

With these and perhaps a few more exceptions it can be said that the theory of Pelletier & Schubert does manage to account for a number of classical recalcitrant problems with mass terms, and gives a solid foundation for future research in this field.

On the other hand, we have seen that the theory (in particular the *p-theory*) gets into trouble when it comes to providing translations for expressions of the form "is a M", where M is what would traditionally be called a mass noun. I believe that this is a problem which is connected with the methodological views of Pelletier & Schubert, and I shall try to make this connection clear in the final section. First the problem itself should be examined more closely, however.

8.2.1 The Problem of Translating Common Nouns

It will be recalled that the translation pattern for all common nouns N preceded by the indefinite article "a" was

$$\lambda x[(\beta\ N)(x) \vee (\gamma\ N)(x)]$$

[56] Its translation being stipulated as **Threw-at**$'[j, (\mu \text{ Snow}'), m]$ (cf. [PS86b, page 79])

This translation pattern is characteristic of both theories, because it embodies their fundamental position that no common noun is *a priori* excluded from having a mass as well as a count sense.

As we have seen, the pattern leads to various technical problems w.r.t. getting certain intuitively valid inferences; but apart from that one might also ask whether its underlying assumption really holds.

If we discuss actual language usage in the terms provided by Pelletier & Schubert's framework, it would seem to follow that traditional mass nouns in most cases will have the *conventionally recognised subkind*-reading; and those usually called count nouns will tend to have the *conventionally recognised portion*-reading—i.e. the traditional "individual reading". But of course there are overlapping cases, as we have seen with the sentence "this is a beer", which may be true on both kinds of reading. However, it would hardly be unfair to say that "is a beer" is atypical among predicative mass terms in having a clear-cut *conventionally recognised portion*-reading[57]. Consider instead "this is a water". It seems to me that the *portion*-reading can only become possible in a rather special context—for instance a waiter bringing glasses of varying content to a table and uttering that sentence to a person there who ordered a glass of water. So the *portion*-reading for "is a water" seems to be heavily context-laden. Consider then a sentence like "this is a liquid". In this case it seems to me that we simply cannot get the portion-reading any more—that there can be no such thing as a "conventionally recognised portion of liquid" by the very generalising nature of the noun "liquid".

In fact it causes Pelletier & Schubert themselves great difficulty in practice to hang on to the position that every noun may have a *mass*- as well as a *count*-sense (in this context, corresponding to a *kind*- as well as a *portion*-sense). In spite of the fact that this position is stated several times throughout [PS86b][58], and the fact that it is reflected by the above trans-

---

[57]Perhaps because of the high incidence of bottles and cans of beer in our society, and their importance in various social rituals.

[58]Consider, for instance, the following remarks (which occur in the context of discussing a *syntactic lexical* approach, but whose general implications are nonetheless clear):

> "*Every* noun—even *hole* and *pore*—sometimes occurs in the noun phrases that we would intuitively call +mass. And every noun sometimes occurs in the noun phrases that we would intuitively call +count. It simply seems that **there is no construction to be ruled out** by treating mass/count as features of lexical entries."
>
> [PS86b, page 26]

And by implication, there would be no sense to be ruled out.

lation rule, Pelletier & Schubert nonetheless wind up speaking of certain predicates being true in only one sense:

> "Linguistically, that is semantically, we take there to be three distinct types of predicates: (a) those which apply only to "kinds", e.g. *is a substance*, *is a wine*, *is a kind of wine*, *is scarce*, *is abundant*, (b) those which apply only to "objects", e.g. *is a quantity of gold*, *is a puddle*, and (c) those which can apply to both "kinds" and "objects". In this last group we have in mind mass predicates such as *is wine*, *is furniture*, *is food* and the like."
> [PS86b, page 74]

As we have seen, Pelletier & Schubert in fact rely on this being the case in their discussion of natural language inferences—that is, they rely on "is a liquid", "is a wine", etc. being always true or false in the *conventional kind*-sense.

Before I go on to draw conclusions from all this, let me say a few words about how to rectify the overall system in order to make the intuitively valid inferences formally valid. My own solution to this problem was in effect to introduce syntactic features of the form $\pm$*mass*[59] to govern the translation procedure: if the feature value of a noun N is *+mass*, then the translation of "is a N" is

$$\lambda x[(\beta\ N)(x)]$$

which in turn guarantees that all the "semi-lattice properties" are available. If the feature value is *-mass*, the translation must be $\lambda x[(\gamma\ N)(x)]$. And if a noun is deemed capable of having both senses, it is simply given dual lexical entries.

This solution of course violates the spirit of the *p-theory* as a semantic occurrence approach. But the only alternative to introducing syntactic features[60] would be to state a meaning postulate like

[59]They did not occur explicitly in the grammar of section 7.1 or that of **Appendix A**, but they were presupposed in the sense that those grammars were only expected to work for lexicons where all basic nouns were mass nouns. In **Appendix D**, where the fragment incorporates both mass- and count-nouns, these syntactic features occur explicitly.

[60]Or rather, the only solution short of revising our intuition about the six valid inferences. Such a solution need not be as devious as it immediately seems—there may be cases where the process of formalisation gives legitimate causes for revising initial intuitive judgments. But I don't think this is one of those cases, and at any rate I am trying to follow Pelletier & Schubert as faithfully as possible, so I shall not elaborate further on this point.

$$\forall x[[(\beta\ P)(x) \lor (\gamma\ P)(x)] \rightarrow (\beta\ P)(x)]$$
where P translates *wine*, *liquid*, ..., ...

and a corresponding meaning postulate for exclusive count nouns.

Although this solution would save us from introducing syntactic features—that is, it would save the principle of a semantic occurrence approach—it is obviously quite unreasonable. First the translation relation would spit out lots and lots of forms like $\lambda x[(\beta\ P)(x) \lor (\gamma\ P)(x)]$, and then a large number of those forms would be "run" through the meaning postulates to simplify them into $\lambda x[(\beta\ P)(x)]$ or $\lambda x[(\gamma\ P)(x)]$.

In [PS86b] Pelletier & Schubert raise the theoretical issue of which sort of approach is preferable: a syntactic lexical approach or a semantic occurrence approach. I should say that these observations point strongly to the syntactic approach as the preferable one[61].

8.3 Conclusion

In consequence of the discussion so far there are two points to be made.

We have seen that the theory[62] of Pelletier & Schubert runs into problems when trying to provide senses for common nouns in certain contexts. In my opinion those problems are a consequence of the *empiricist conception* approach: all conceiveable readings—including a few barely conceiveable ones—*must* be accounted for, regardless of the costs. The perceived obligation to be able to meet even the most far-fetched of uses has apparently overridden all other philosophical, logical, or for that matter computational considerations.

On the other hand, if one accepts from the outset a certain degree of regimentation, then I believe that the "serving situations" for most predicative mass terms could safely be ignored. For instance, no *portion*-reading would be required for "is a gold" or "is a wine". The few serious exceptions to this such as "is a beer" can then be handled by dual lexical entries (as previously suggested).

[61]For the sake of completeness I add that the syntactic features introduced here are what Pelletier & Schubert's would call "a mere reflection of the semantics", and in their opinion that should be avoided even in a syntactic lexical approach. However, they are less artificial than the "meaning postulate solution", and so we may just have to accept that such "reflex" features are necessary for getting the semantics of natural language right.

[62]In particular the *p-theory*, but to some extent also the *s-theory*, in so far it shares the position that *all* nouns have *mass-* as well as *count*-senses.

Of course, it would go to far to turn these observations into a general rejection of the *empiricist conception*; but I do think that that conception is underlying Pelletier & Schubert's approach, and that it is their administration of this methodology which leads them into problems.

If we accept that the methodological programme of Pelletier & Schubert plays a role in the shortcomings of their theory, it is all the more remarkable that they hardly really carry it through anyway. I have already observed that their treatment of bare noun phrases as denoting "the most general kind" is a departure from basic intuition. Indeed the whole treatment of mass senses of nouns is a radical abstraction: in their system the "immediate" sense is almost always a *kind*-sense[63], as in (β *wine*) or (μ *wine*). From those senses the more ordinary intuitive senses such as *quantity*, *stuff* etc. can then be "derived", for example by using meaning postulates. In my opinion this abstraction is a clear example of imposing a philosophically or logically motivated structure upon language rather than the other way round. Not that there is anything wrong with that: it is this very abstraction which makes their theory truly ingenious and makes it possible to develop their logic for mass expressions in general. But I should say that it takes us a far way form the programme of generating philosophical concepts from strictly empirical linguistic observations.

So in the final analysis it seems to me that what Pelletier & Schubert have accomplished is actually not so much a linguistic theory of mass terms as a theory of "reasoning about kinds and their interrelations with quantities and objects". As such it is more of a philosophical theory—with the very important qualification that it is computationally implementable. True, the theory does take its initial clues from scrutinising natural language, as it would have to do in order to become a proper theory of reasoning about any subject. But once having settled upon its central notions it becomes ever more abstract and separated from empirical linguistics—running its own course, as it were.

Apparently this implies that Pelletier & Schubert are actually in discrepancy with their own statements on methodology. Apart from that I can see no harm in proceeding the way they do as I have just described it—on the contrary it seems to me to be the most advisable and productive course of formal investigation into natural language for the foreseeable future.

[63]The exception being translations of the form (δ M), for instance where the noun "M" co-occurs with the determiner "sm" as in "sm M".

Part III
The Theory of G. Link on Plural

9 The Basic Theory: Common Nouns and Intransitive Verbs

9.1 Introduction

In [Lin83], G. Link laid out an intriguing set of new ideas for dealing with mass terms and plurals. His paper has since dominated further work in this field, although there has been a tendency towards concentrating upon that part of it which has to do with plurals. Actually there can be little doubt that his ideas do work better for plurals than for mass terms, but this is an issue which I shall not consider further here. (I may remind you, though, that in [Lin83] it was assumed that nominal mass terms do not have a proper logic, whereas I suggested that the theory of Pelletier & Schubert provided exactly that—see section 8.2). In the following I shall ignore those parts of Link's paper which deal with mass terms. Instead I shall try to give a full implementation of his basic ideas with respect to plurals. In doing so I shall try to be as faithful as possible to those ideas.

In the current section I shall completely ignore intransitive verb phrases built from a transitive verb and an object phrase. Instead, I shall deal with basic intransitive verbs only (with the exception of the verb "be"). Transitive verbs will be taken care of in the next section. The reason for this sharp division is that transitive verbs pose some problems of their own, problems which are apparently not realised in [Lin83].

9.2 Basic Concepts and Symbols

In order to explain the basic notions let there be a micro-universe consisting of only three entities, say "John", "Mary", and "George". More precisely, let

$$A = \{\| j \|, \| g \|, \| m \|\}$$

and let us further assume that

$$\begin{array}{lcl} John & \rhd & j \\ Mary & \rhd & m \\ George & \rhd & g \end{array}$$

Then, how do we provide denotations for a composite (and plural) noun phrase like "John and Mary"? Link's answer to this problem is to introduce a sum-operator (for which I shall use the symbol $\uplus$), which from two individuals form a new composite individual. Thus for instance

$$\| (John\ and\ Mary)' \| = \| j \| \uplus \| m \|$$

It is crucially important to understand that in Link's framework such new individuals are of the same type as the non-composite (atomic) individuals—in terms of the types of [PTQ] they are of type $\langle e \rangle$. And they are as much individuals in their own right as are "John" and "Mary" by themselves.

So, the total domain of individuals in the above micro-universe is actually larger than just A. Let us call this domain E. Then

$$\begin{aligned} E = \ & \{\| j \|, \| m \|, \| g \|, \| j \| \uplus \| m \|, \\ & \| m \| \uplus \| g \|, \| j \| \uplus \| g \|, \\ & \| j \| \uplus \| m \| \uplus \| g \|\} \end{aligned}$$

The set A is the set of *atomic individuals* and is a subset of E. Technically, we impose a semi-lattice structure on E, where $\uplus$ is the usual join-operator, and $\leq$ is the symbol of the partial order induced on E by $\uplus$. We have the Boolean equation

$$\| a \| \leq \| b \| \ \text{ iff } \ \| a \| \uplus \| b \| = \| b \|$$

So far we have been looking at strictly semantic symbols and concepts. But in order for these to be of use in a logical target language they must have some syntactic counterparts. The semantic symbol $\uplus$ will have the syntactic counterpart $\oplus$, then, and the semantic relation symbol $\leq$ will have Π as its counterpart. For $x \oplus y$ we read *the i-sum of x and y* or, equivalently, *the individual sum of x and y*. For $x \,\Pi\, y$ we read *x is an i-part of y* or *x is an individual part of y*.

Equipped with these simple notions we can now go on to provide denotations for common nouns in the plural. Consider for instance the common noun *man*, and suppose with respect to our micro-universe that

$$\| (man') \| = \{\| j \|, \| g \|\}$$

On the basis of the semi-lattice structure, Link introduces a new symbol (*), which operates on predicates. If P is a predicate, then $^*P$ is the semi-lattice closure of that predicate. The simple rules of the closure are

1. For all $a \in \| P \|$, $a \in \| {}^{*}P \|$.

2. If $x, y \in \| {}^{*}P \|$, then $x \uplus y \in \| {}^{*}P \|$.

3. Nothing else is in $\| {}^{*}P \|$.

Thus, for instance, E above is actually the semi-lattice closure of A.

Given the semantic value of *man* above, we can now quite easily form the semantic value of the closure of *man*, viz

$$\| {}^{*}man \| = \{\| j \|, \| g \|, \| j \| \uplus \| g \|\}$$

This type of semantic construction is called the *plural predicate of P* by Link. However, obviously it includes "singular" (i.e. *atomic*) entities as well as genuinely plural entities, and so we also need an operator which gives us the latter only. This is provided by a further operator (${}^{\bullet}$), which is simply defined to be (for all predicates P)

$$\| {}^{\bullet}P \| = \| {}^{*}P \| \setminus A$$

Thus in our present universe we have

$$\| {}^{\bullet}man \| = \| {}^{*}man \| \setminus A$$

that is

$$\| {}^{\bullet}man \| = \{\| j \| \uplus \| g \|\}$$

The predicate ${}^{\bullet}P$ is called the *proper plural predicate of P*.

Now we cannot say with complete generality that the semantic value of a plural common noun N is to be $\| {}^{\bullet}N' \|$. The semantic value depends on the context and in particular on the determiners, with which the noun co-occurs. (This will become clear in due course.) However, the gadgetry given so far provides us with the essentials of a logical language—call it **LP**—within which we can give fairly natural logical representations of a significant number of plural constructions. Some examples of this will be given below. Let me add that the closure-operator and the plural-operator apply not only to translations of common nouns, but to the translations of a number of basic intransitive verbs as well[64].

[64]Technically these operators can of course be applied to *any* one-place predicate symbol within **LP**. But in practice the translation relation will secure that they are applied to only those one-place predicates of **LP**, which are translations of *distributive* nouns or verbs.

I hope that this brief description of Link's system will give sufficient background for understanding that which is to follow. Naturally, I refer to [Lin83] for the full details, but I have here tried to extricate all those features which bear directly on the following account of plural constructions. So let me now go on to show how these constructs are used by Link in the description of natural language. First I shall give a few examples of logical representations of natural language sentences, and some natural language inferences; then I shall go on to discuss briefly Link's translation relation from natural language into **LP**.

9.3 Some Representations and Inferences

Before looking at the representations let me recapitulate the distinction between *collective* and *distributive* verbs. A verb like "die" is said to be *distributive*, because it distributes over the "members" of its subject term. To take a very simple example, "John and Mary die" is semantically equivalent to "John dies and Mary dies"[65]. However, for a significant number of verbs we have no similar result: for instance "John and Mary convene" is certainly not equivalent to "John convenes and Mary convenes". The latter sentence is inherently false, and perhaps even ungrammatical. Verbs such as "convene" are said to be *collective*.

The algebraic properties of the semi-lattice structure—and in particular, the closure- and the plural-operators—allow us to represent this distinction in a very direct manner. At the same time, plural constructions are dealt with.

Consider, then, the following sentences and their representations:

1. some child dies
   $\rhd \exists x[\mathit{child}(x) \wedge \mathit{die}(x)]$

2. some children die
   $\rhd \exists x[{}^{*}\mathit{child}(x) \wedge {}^{*}\mathit{die}(x)]$

3. John and Mary die
   $\rhd {}^{*}\mathit{die}(j \oplus m)$

4. some children convene
   $\rhd \exists x[{}^{*}\mathit{child}(x) \wedge \mathit{convene}(x)]$

---

[65]Or rather, most authors (including myself) consider sentences of this pattern to be equivalent. But that point can be disputed, and as you will you see shortly, they do not emerge as equivalent within Link's system.

5. John and Mary convene
   $\triangleright$ $convene(j \oplus m)$

6. John and Mary do not die
   $\triangleright \neg{}^{\bullet}die(j \oplus m)$

Sentence (1) is exactly as we would expect. In sentence (2) the plural-operator has come into use. The translation of (2) will be true if and only if there are at least two separate entities such that each of them is a child who dies. Notice that in this truth-condition distributivity is inherent. (The inferences below will illustrate this further.) In the translation of sentence (3), both the plural- and the sum-operator come into use. So its truth-condition will be that there are two separate entities, namely "John" and "Mary", and that each of them dies.

Actually this truth-condition for (3) is not entirely uncontroversial. Pragmatically speaking we would no doubt as hearers of an utterance of (3) presuppose that John and Mary are different. But there are plural predicates which bear no such presupposition, for instance "are identical". The analogous translation of "John and Mary are identical" into ${}^{\bullet}identical(j \oplus m)$ is simply non-sensical. But even if we expect such plural predicates—i.e. those without a presupposition of "proper semantic plurality"—to be limited in number and amenable to some form of special treatment, it must be noted that on the translation of (3) above that sentence is *not* equivalent to

(3') John dies and Mary dies.

On its Linkian translation (3) implies (3'), but not conversely.

In (4), the plurality of the subject term is expressed simultaneously with the collectivity of the verb phrase. The truth condition is that there are at least two entities such that each of them is a child, and together (or collectively) these entities perform the act of convening. Thus it is expressed at the same time that each of the entities involved is a child, but that no inference to the effect that each of them convenes can be made. In (5), we see a similar structure with the sum-operator; no inference with respect to the individual "constituents" of the subject term is warranted.

Finally, the translation of sentence (6) shows that the application of the proper plural operator to intransitive verb phrases also gives rise to some problems with negation. Since ${}^{\bullet}die(j \oplus m)$ is equivalent[66] to

$$die(j) \wedge die(m) \wedge j \neq m$$

[66] Still assuming that $\| j \|, \| m \| \in A$.

the translation of sentence (6) is equivalent to

$$\neg die(j) \vee \neg die(m) \vee j = m$$

So it is quite possible that "John and Mary don't die" and "John dies and Mary dies" are true at the same time. That is hardly satisfactory; but this problem would disappear—like the problems with sentence (3)—if "John and Mary die" were treated as equivalent to "John dies and Mary dies". For then the translation of "John and Mary do not die" would simply be equivalent to

$$\neg die(j) \vee \neg die(m)$$

which is as it should be, I believe.

Of course, the discussion of the above representations of natural language sentences has already involved some discussion of the inferences which they warrant. But it will be useful also to discuss three further inference patterns stated by Link himself; these patterns are without doubt intended to be particularly characteristic of his system. And indeed they are, but in one case an important extra proviso has to be added.

First, there is the inference-pattern which Link calls *expansion* (but more commonly called *cumulativity*). In [Lin83], it is stated as follows:

(41)*a.* John and Paul are pop stars and George is a pop star,
so John, Paul, and George are pop stars.
b. $^{\bullet}P(a \oplus b) \wedge Pc \Rightarrow {}^{\bullet}P(a \oplus b \oplus c)$
[Lin83, page 311]

This captures the often discussed *cumulative reference property* of plurals, and the pattern is valid within the present framework.

Next, there is the inference pattern which Link calls *contraction*:

(42)*a.* John, Paul, George, and Ringo are pop stars,
so Paul and Ringo are pop stars.
b. $^{\bullet}P(a \oplus b \oplus c \oplus d) \Rightarrow {}^{\bullet}P(b \oplus d)$
[Lin83, page 311]

This pattern captures the *dissectiveness* property of plurals, albeit in a form restricted to proper plural predicates. Unfortunately, the inference is actually ***not*** valid within the present framework, and it is harder to mend this than one might initially believe. For the inference to be valid we must add the proviso that $b \neq d$. But much as was the case with sentence (3) above, this proviso is more in the way of a pragmatic presupposition than it is a logico-linguistic necessity. When we join terms such as in "John, Paul, George, and Ringo" it is without doubt a bit of default reasoning that they are mutually different, unless there is explicit information to the opposite effect. But as I have already shown with the example "John and Mary are identical", this is nothing more than a pragmatic expectation.

At any rate, the inference is invalid, and my axiomatisation shall not warrant any inferences of this sort, except of course when the needed proviso is explicitly added.

I now turn to the last inference pattern given by Link. First I must state his higher-order predicate *Distr*, which gives the definition of distributivity of a predicate:

(27) $Distr(P) \leftrightarrow \forall x(Px \rightarrow At\, x)$
[Lin83, page 309]

where $At\, x$ is true if and only if $\| x \| \in A$, i.e. x is an atom. The inference itself is this:

(28) *a.* John, Paul, George, and Ringo are pop stars.
Paul is a pop star.
[Lin83, page 309]

Symbolically,

(29) *a.* $^*P(a \oplus b \oplus c \oplus d)$
b. $Distr(P)$
c. $b \Pi a \oplus b \oplus c \oplus d$
d. $^*Pb$
e. Pb
[Lin83, page 309]

This inference is valid and captures the general property of ***dissectiveness***. It is a straight-forward algebraic consequence of the way in which the semi-lattice closure of a set—respectively, a predicate—is constructed.

It is worth noting, though, that the premise *Distr(P)* (cf. (29) b.) is *not* present in the natural language inference of (28), and hence the entire inference (28)-(29) presupposes that this kind of "extrinsic information" is somehow accessible for the inference system[67].

9.4 Link's Translation Relation

In order to provide a translation relation from natural language into **LP**, Link takes over the general framework of [PTQ] and implements his own notions into this by stating the additions and modifications which become necessary. A rather important modification regards the translation rule T4, which in its original version says

> T4. If $\delta \in P_{t/\mathrm{IV}}$, $\beta \in P_{\mathrm{IV}}$, and δ, β translate into δ', β' respectively, then $F_4(\delta, \beta)$ translates into $\delta'(\hat{}\beta')$
> [PTQ, page 261]

Link modifies this rule with the remark

> "In accordance with what I did above the translation rule T4 involving distributive expressions has to be such that their translations always enter this rule under the star operator."
> [Lin83, page 318]

Actually it is not quite clear to me what Link is referring to with the phrase "in accordance with what I did above"—there is, however, one example of an intransitive verb phrase under the proper plural operator[68] in the previous text, namely the verb phrase "are pop stars", which we have seen in the inference pattern of (42) above.

It is also worth noting that Link anticipates a more general applicability of the closure- and the plural-operator by explicitly letting[69]

$$^{*}\varsigma,\ ^{\bullet}\varsigma \in \mathrm{ME}_{\langle\tau\rangle},\ \text{for } \tau = \langle e,t\rangle \text{ or } \tau = \langle s,\langle e,t\rangle\rangle,\ \varsigma \in \mathrm{ME}_{\langle\tau\rangle}$$

Let us assume, then, that basic intransitive verbs are subcategorised as distributive and non-distributive. Then T4 will provide us with the translations foreseen in the examples (1)-(5) above.

---

[67] Such information could for instance be gleaned from the lexicon by inspection in this, assuming that nouns and verbs are equipped with syntactic features of the form $\pm$*distr*. Ambiguous cases could be handled by dual lexical entries; inferences for those cases would naturally have to be relativised to the reading being considered for the lexical item in question (distributive or non-distributive).

[68] The symbol which Link uses for this operator is a star within a circle.

[69] Cf. [Lin83, page 317]

As for common nouns, the [PTQ]-rules T2 and T3 must be modified in a largely analogous manner. In [PTQ], T2 specifies the translation of noun phrases formed from a determiner and a common noun (e.g. "the woman"); T3 specifies the translation of common nouns with subordinate relative clauses (e.g. "man such that he dies")[70]. There is one major difference from the [PTQ] w.r.t. T2 and T3, namely that determiners are not introduced syncategorematically by Link, but rather in the more traditional compositional way.

Link gives the following translations for the determiners[71]:

1. a, some $\triangleright$ $\lambda Q \lambda P \exists x[Q(x) \wedge P(x)]$
2. the $\triangleright$ $\lambda Q \lambda P \exists y[Q(y) \wedge \forall x[Q(x) \rightarrow x \,\Pi\, y] \wedge P(y)]$
3. every, all $\triangleright$ $\lambda Q \lambda P \forall x[Q(x) \rightarrow P(x)]$

Finally, before setting out on the implementation, let me mention two minor oddities in Link's system.

First, a bare plural noun phrase such as "children" in "children died" is always seen to be equivalent to "some children". Maybe this is linguistically acceptable, provided that generic readings are kept out of the picture. At any rate, in my implementation I am going to follow Link in this decision. But it is certainly not quite self-evident, so it should be mentioned before the reader encounters it in concrete examples[72].

Second, Link actually gives no syntactic rule for joining terms as in "John and Mary" or "the students and some masters". There are some vague hints towards the end of Link's paper, but nothing concrete.

---

[70]Link does not discuss those rules, however; but clearly they have to be reformulated in a more exhaustive account of his theory.

[71]Link deals with bare plural noun phrases by introducing a "zero determiner" $\emptyset_{pl}$, which is given the same translation as "some". However, I have chosen to deal with such cases by separate syntactic rules (Cf. **Appendix B**), and so I have ignored this determiner here.

Furthermore, Link stipulates a rather complex translation of the putative determiner "all the", of which he claims that it has certain particular pragmatic properties. I have chosen to ignore this determiner and his discussion of its properties, since this issue hardly bears on the general nature of his theory of plurals.

[72]An obvious alternative to this treatment would be to regard bare plural noun phrases as *ambiguous*; e.g. the noun phrase "children" can be regarded as ambiguous between "all children", "most children", "some children" etc. This would be in line with Parsons' views on bare mass noun phrases, (cf. p. 33), which are in this respect structurally analogous to bare plural noun phrases. See also section 10.10 for a bit further discussion of this subject.

9.5 A Revised Translation Relation

The details of my translation relation (including the lexicon) must be found in **Appendix B**, but some general comments will probably be helpful.

My implementation of Link's translation relation is intended to constitute together with the axiomatisation of his semantics a fully defined system. That is, it should be possible to follow mechanically each and every step from a natural language input sentence to its (natural language) consequences. The intention is that an inference engine for this system shouldn't need to have recourse to the lexicon to see whether a predicate is distributive or not. This policy has some consequences for the output logical forms. For instance, in my translation relation "a college" translates into[73]

$$\text{(a) } \lambda Q \exists x [{}^{*}\mathit{college}(x) \wedge Q(x)]$$

rather than the form you would expect, viz

$$\text{(b) } \lambda Q \exists x [\mathit{college}(x) \wedge Q(x)]$$

An inference engine working on translation (b) would need recourse to a lexicon or a database with meta-information on the predicates of **LP** in order to be able to draw an inference like

$$\begin{array}{l} \models \ \mathit{college}(x) \\ \models \ {}^{*}\mathit{college}(x) \end{array}$$

which is in turn necessary for some of the natural language inferences studied here. But I wan't to avoid any such reliance on extrinsic information, and translation (a) indeed eliminates the need for it.

For the same reasons I have refrained from assuming the availability of a post-parsing phase, although in the long run it will very likely be preferable to incorporate such a phase[74].

Apart from this, the points which seem to me to deserve some elucidation are the following:

[73]Intuitively translation (a) seems to be to weak, since it quantifies over plural as well as over atomic entities; but this problem will be addressed under the heading **Atomicity** below, and at any rate it does not affect the point which I am trying to make here.

[74]See also my comments on the post-parsing phase of Pelletier & Schubert's theory, footnote 28 and section 7.1.3.

Atomicity: In Link's model-theoretic system, it is taken for granted that we always know of an entity whether it is atomic or not. However, for an axiomatic system designed to operate directly on the syntactic forms of the translations no such assumption can be made. If it were necessary to know whether an entity is atomic in order to get the proper inferences, the translation relation should provide this sort of information, for instance by translating expressions like "a college" into

$$\lambda Q \exists x[{}^{*}college(x) \wedge atomic(x) \wedge Q(x)].$$

However, the notion of *atomicity* is redundant w.r.t. the inferences studied in this section, and so it is simply omitted here. While this may render the translations of singular expressions like "a college" intuitively unsatisfactory, it must be remembered that *inference* is at the heart of any axiomatisation. As long as the notion isn't seen to be strictly necessary for getting the proper inferences, the principle of *Occam's Razor* calls for its elimination.

To omit the notion until it becomes strictly necessary has the advantage of simplifying proofs and also of making it clear exactly for what the notion is ultimately needed. I may mention already now that *atomicity* is indeed reintroduced, when transitive verbs are incorporated.

Joining terms: The lexical item "and" must have a separate entry in the lexicon for its term-joining function. Its translation must ensure that composite terms like "John and Mary" or "some students and the masters" are joined as intended. The rather complicated lambda-expression which does this can be found in the lexicon of **Appendix B**; to be fully understood it must be compared with the rule for composite noun phrases (see immediately below).

Composite noun phrases: A new syntactic category CNP (composite noun phrase) is introduced to be able to construct composite noun phrases. A non-composite noun phrase (NP) is an immediate subnode of an CNP.

Determiners: As for determiners I deviate significantly from Link in giving dual lexical entries for all those capable of combining with both singular and plural nouns. This gives more flexibility to the translation relation, since we need not decide whether to "star" a predicate or not before we have seen with which determiner it combines. For instance, a plural noun should not be under the proper plural operator when combining with determiners like "all" or "no". At any rate, I think that as the fragment is extended to cover ever more complicated cases—and in particular, when mass terms are to be involved on a significant scale—it will be impossible to uphold uniform translations of the determiners.

I draw attention to the translations of the definite article: the plural determiner the_{pl} translates into

$$\lambda P \lambda Q\, \exists x[{}^{\bullet}P(x) \wedge \forall y[{}^{*}P(y) \leftrightarrow y \,\Pi\, x] \wedge Q(x)]$$

This translation—together with the translation rules—gives logical forms which are more convenient for an inference system than Link's translation, since it incorporates directly facts that you otherwise have to infer[75].

The singular determiner the_{sing} translates into

$$\lambda P \lambda Q\, \exists x[\forall y[{}^{*}P(y) \leftrightarrow y = x] \wedge Q(x)]$$

Again, this translation directly gives us some information in the singular case which we would otherwise have to infer.

Finally, let me mention that in the current section I abide by the principle of translating distributive intransitive verb phrases into *proper* plural predicates (in spite of the problems with this principle, which were discussed in connection with sentences (3) and (6) on page 64). By abiding by this principle its implications will become more clear. I conclude this subsection by listing some translations according to my translation relation:

[75]Consider an inference like "the students die, John is a student, so John dies".

In Links translation relation, the first premise translates into

$$\exists x[{}^{\bullet}\mathit{student}(x) \wedge \forall y[{}^{\bullet}\mathit{student}(y) \rightarrow y \,\Pi\, x] \wedge {}^{\bullet}\mathit{die}(x)]$$

To prove the inference we shall have to prove that this premise entails

$$\exists x[{}^{\bullet}\mathit{student}(x) \wedge \forall y[{}^{*}\mathit{student}(y) \leftrightarrow y \,\Pi\, x] \wedge {}^{\bullet}\mathit{die}(x)]$$

(To see this in greater detail compare with natural language inference 5 and its proof on page 82).

This implication can indeed be proved in Links's system, but in my translation relation this information is already available in the translation itself and so those steps can be omitted.

9.5.1 Sentences with a Distributive Verb Phrase

1. some student dies

   $\triangleright\ \exists x[{}^{*}student(x) \wedge {}^{*}die(x)]$

2. the student dies

   $\triangleright\ \exists x[\forall y[{}^{*}student(y) \leftrightarrow y = x] \wedge {}^{*}die(x)]$

3. some students die

   $\triangleright\ \exists x[{}^{\bullet}student(x) \wedge {}^{\bullet}die(x)]$

4. students die

   $\triangleright\ \exists x[{}^{\bullet}student(x) \wedge {}^{\bullet}die(x)]$

5. the students die

   $\triangleright\ \exists x[{}^{\bullet}student(x) \wedge \forall y[{}^{*}student(y) \leftrightarrow y \,\Pi\, x] \wedge {}^{\bullet}die(x)]$

6. John and Mary die

   $\triangleright\ {}^{\bullet}die(john \oplus mary)$

7. some students and Mary die

   $\triangleright\ \exists x[{}^{\bullet}student(x) \wedge {}^{\bullet}die(x \oplus mary)]$

8. some students and the masters die

   $$\triangleright\quad \exists x[{}^{\bullet}student(x) \wedge \exists z[{}^{\bullet}master(z) \wedge \forall y[{}^{*}master(y) \leftrightarrow y \,\Pi\, z] \wedge {}^{\bullet}die(x \oplus z)]]$$

9.5.2 Sentences with a Collective Verb Phrase

9. some students convene

   $\triangleright\ \exists x[{}^{\bullet}student(x) \wedge convene(x)]$

10. students convene

    $\triangleright\ \exists x[{}^{\bullet}student(x) \wedge convene(x)]$

11. the students convene

    $\triangleright\ \exists x[{}^{\bullet}student(x) \wedge \forall y[{}^{*}student(y) \leftrightarrow y \,\Pi\, x] \wedge convene(x)]$

12. John and Mary convene

$\triangleright$ $convene(john \oplus mary)$

13. some students and Mary convene

$\triangleright$ $\exists x[{}^{\bullet}student(x) \wedge convene(x \oplus mary)]$

14. some students and the masters convene

$\triangleright$ $\exists x[{}^{\bullet}student(x) \wedge$
$\exists z[{}^{\bullet}master(z) \wedge \forall y[{}^{*}master(y) \leftrightarrow y \, \Pi \, z] \wedge convene(x \oplus z)]]$

9.6 Axiomatisation of Link

The axiomatisations given here will be restricted to deal with purely extensional fragments only (as mentioned in the introduction (part I)). As a consequence my axiomatisation of Link's theory covers only its extensional sub-fragment. Moreover, I state only those axioms and theorems which are relevant to the natural language inferences studied in the next section.

The importance of the present axiomatisation lies in the fact that it realises my previously stated goal: together with the translation relation the axiomatisation makes it possible to follow mechanically each and every step from a set of natural language input sentences to their consequences. This result would also seem to corroborate my suggestion in part I that theories from formal semantics can form part of the basis for the development of logic programming languages.

As I have already suggested in section 9.5, we shall not need recourse to the lexicon to be able to tell whether a predicate is distributive or not. This will be implicit in the translations, and hence an inference engine will be able to use the following axioms and theorems directly.

9.6.1 Basic Semi-Lattice Properties

Axiom 1 (Idempotency of '⊕') $\forall x : (x \oplus x) = x$

Axiom 2 (Commutativity of '⊕') $\forall x, y : (x \oplus y) = (y \oplus x)$

Axiom 3 (Associativity of '⊕')
$\forall x, y, z : (x \oplus y) \oplus z = x \oplus (y \oplus z)$

Axiom 4 (Reflexivity of 'Π') $\forall x : x \,\Pi\, x$

Axiom 5 (Transitivity of 'Π')
$\forall x, y, z : x \,\Pi\, y \wedge y \,\Pi\, z \rightarrow x \,\Pi\, z$

Axiom 6 (Antisymmetry of 'Π')
$\forall x, y : x \,\Pi\, y \wedge y \,\Pi\, x \leftrightarrow x = y$

Axiom 7 (Interdefinability of '⊕' and 'Π')
$\forall x, y : x \,\Pi\, y \leftrightarrow x \oplus y = y$

Theorem 1 $\forall x, y : x \,\Pi\, (x \oplus y)$

Proof 1

$\vdash_1$	$(x \oplus x) \oplus y = x \oplus y$	Axiom 1
$\vdash_2$	$x \oplus (x \oplus y) = x \oplus y$	Axiom 3
$\vdash_3$	$x \oplus (x \oplus y) = x \oplus y \leftrightarrow x \,\Pi\, (x \oplus y)$	Axiom 7
$\vdash_4$	$x \,\Pi\, (x \oplus y)$	replacement

9.6.2 Axioms for the Closure-Operator (*)

Axiom 8 (Cumulativity)
$\forall x, y : {}^*P(x) \wedge {}^*P(y) \rightarrow {}^*P(x \oplus y)$

Axiom 9 (Dissectiveness)
$\forall x, y : {}^*P(x) \wedge y \,\Pi\, x \rightarrow {}^*P(y)$

Theorem 2 (Relation between '⊕' and '∧')
$\forall x, y : {}^*P(x \oplus y) \leftrightarrow {}^*P(x) \wedge {}^*P(y)$

Proof 2

$\vdash_1$	$^*P(x \oplus y)$	premise
$\vdash_2$	$x \Pi (x \oplus y)$	theorem 1
$\vdash_3$	$^*P(x \oplus y) \wedge x \Pi (x \oplus y)$	conjunction
$\vdash_4$	$^*P(x \oplus y) \wedge x \Pi (x \oplus y) \rightarrow {^*P(x)}$	axiom 9
$\vdash_5$	$^*P(x)$	detachment
$\vdash_6$	$y \Pi (x \oplus y)$	theorem 1
	$\vdots$	(as for x)
$\vdash_7$	$^*P(y)$	
$\vdash_8$	$^*P(x) \wedge {^*P(y)}$	conjunction
$\vdash_9$	$^*P(x \oplus y) \rightarrow {^*P(x)} \wedge {^*P(y)}$	deduction theorem
$\vdash_{10}$	$^*P(x) \wedge {^*P(y)} \rightarrow {^*P(x \oplus y)}$	axiom 8
$\vdash_{11}$	$^*P(x) \wedge {^*P(y)} \leftrightarrow {^*P(x \oplus y)}$	conjunction and material equivalence

9.6.3 Axioms for the Plural-Operator ($^\bullet$)

Axiom 10 $\forall x : {^\bullet P(x)} \rightarrow {^*P(x)}$

Axiom 11 (Expansion)
$\forall x, y$, where $x \neq y$: $^*P(x) \wedge {^*P(y)} \rightarrow {^\bullet P(x \oplus y)}$

Theorem 3 $\forall x, y : {^\bullet P(x)} \wedge {^*P(y)} \rightarrow {^\bullet P(x \oplus y)}$

Proof 3
(1) $x \neq y$:

$\vdash_1$	$^\bullet P(x) \wedge {^*P(y)}$	premise
$\vdash_2$	$^\bullet P(x)$	simplification
$\vdash_3$	$^\bullet P(x) \rightarrow {^*P(x)}$	axiom 10
$\vdash_4$	$^*P(x)$	detachment
$\vdash_5$	$^*P(y)$	simplification of 1
$\vdash_6$	$^*P(x) \wedge {^*P(y)}$	conjunction
$\vdash_7$	$^*P(x) \wedge {^*P(y)} \rightarrow {^\bullet P(x \oplus y)}$	axiom 11
$\vdash_8$	$^\bullet P(x \oplus y)$	detachment
$\vdash_9$	$^\bullet P(x) \wedge {^*P(y)} \rightarrow {^\bullet P(x \oplus y)}$	deduction theorem

(2) $x = y$:

$\vdash_{10}$	$^{\bullet}P(x) \wedge {}^{*}P(y)$	premise
$\vdash_{11}$	$^{\bullet}P(x)$	simplification
$\vdash_{12}$	$^{\bullet}P(x \oplus x)$	axiom 1
$\vdash_{13}$	$^{\bullet}P(x \oplus y)$	since $x = y$
$\vdash_{14}$	$^{\bullet}P(x) \wedge {}^{*}P(y) \rightarrow {}^{\bullet}P(x \oplus y)$	deduction theorem

Theorem 4
$\forall x, y$, *where* $x \neq y$: $^{*}P(x) \wedge {}^{*}P(y) \leftrightarrow {}^{\bullet}P(x \oplus y)$

Proof 4 $x \neq y$:

$\vdash_{1}$	$^{\bullet}P(x \oplus y)$	premise
$\vdash_{2}$	$^{\bullet}P(x \oplus y) \rightarrow {}^{*}P(x \oplus y)$	axiom 10
$\vdash_{3}$	$^{*}P(x \oplus y)$	detachment
$\vdash_{4}$	$^{*}P(x) \wedge {}^{*}P(y)$	theorem 2 + replacement
$\vdash_{5}$	$^{\bullet}P(x \oplus y) \rightarrow {}^{*}P(x) \wedge {}^{*}P(y)$	deduction theorem
$\vdash_{6}$	$^{*}P(x) \wedge {}^{*}P(y) \rightarrow {}^{\bullet}P(x \oplus y)$	axiom 11
$\vdash_{7}$	$^{*}P(x) \wedge {}^{*}P(y) \leftrightarrow {}^{\bullet}P(x \oplus y)$	conjunction + material equivalence

Axiom 12 [76] $\forall x : {}^{\bullet}P(x) \wedge {}^{*}Q(x) \rightarrow {}^{\bullet}Q(x)$

[76] After the axiomatisation has been completed, I have realised that this axiom can in fact be proved as a theorem, using axiom 13. The main outline of the proof is as follows:

$\vdash$ $^{\bullet}P(x) \wedge {}^{*}Q(x)$
$\vdash$ $y \,\Pi\, x \wedge z \,\Pi\, x \wedge y \neq z$
$\vdash$ $^{\bullet}Q(y \oplus z)$
$\vdash$ $^{\bullet}Q(y \oplus z) \wedge {}^{*}Q(x)$
$\vdash$ $^{\bullet}Q(y \oplus z \oplus x)$
$\vdash$ $^{\bullet}Q(x)$

Theorem 5
$\forall x$: $[{}^\bullet P(x) \wedge \forall z[{}^*P(z) \rightarrow {}^*Q(z)]] \rightarrow {}^\bullet Q(x)$

Proof 5

$\vdash_1$	${}^\bullet P(x) \wedge \forall z[{}^*P(z) \rightarrow {}^*Q(z)]$	premise
$\vdash_2$	${}^\bullet P(x)$	simplification
$\vdash_3$	${}^*P(x)$	axiom 10 + detachment
$\vdash_4$	$\forall z[{}^*P(z) \rightarrow {}^*Q(z)]]$	simplification of 1
$\vdash_5$	${}^*P(x) \rightarrow {}^*Q(x)$	U.I.
$\vdash_6$	${}^*Q(x)$	detachment
$\vdash_7$	${}^\bullet P(x) \wedge {}^*Q(x)$	conjunction of 2,6
$\vdash_8$	${}^\bullet P(x) \wedge {}^*Q(x) \rightarrow {}^\bullet Q(x)$	axiom 12
$\vdash_9$	${}^\bullet Q(x)$	detachment
$\vdash_{10}$	$[{}^\bullet P(x) \wedge \forall z[{}^*P(z) \rightarrow {}^*Q(z)]] \rightarrow {}^\bullet Q(x)$	deduction theorem

Theorem 6 (Contraction)
$\forall x, y, z$, *where* $x \neq y$: ${}^\bullet P(x \oplus y \oplus z) \rightarrow {}^\bullet P(x \oplus y)$

Proof 6 $x \neq y$:

$\vdash_1$	${}^\bullet P(x \oplus y \oplus z)$	premise
$\vdash_2$	${}^\bullet P(x \oplus y \oplus z) \rightarrow {}^*P(x \oplus y \oplus z)$	axiom 10
$\vdash_3$	${}^*P(x \oplus y \oplus z)$	detachment
$\vdash_4$	${}^*P(x \oplus y \oplus z) \leftrightarrow {}^*P(x \oplus y) \wedge {}^*P(z)$	theorem 2
$\vdash_5$	${}^*P(x \oplus y) \wedge {}^*P(z)$	replacement
$\vdash_6$	${}^*P(x \oplus y)$	simplification
$\vdash_7$	${}^*P(x) \wedge {}^*P(y)$	theorem 2 + replacement
$\vdash_8$	${}^*P(x) \wedge {}^*P(y) \rightarrow {}^\bullet P(x \oplus y)$	axiom 11
$\vdash_9$	${}^\bullet P(x \oplus y)$	detachment
$\vdash_{10}$	${}^\bullet P(x \oplus y \oplus z) \rightarrow {}^\bullet P(x \oplus y)$	deduction theorem

Theorem 7 $\forall x, y : {}^\bullet P(x) \rightarrow [y \,\Pi\, x \rightarrow {}^*P(y)]$

Proof 7

$\vdash_1$	${}^\bullet P(x) \wedge y \,\Pi\, x$	premise
$\vdash_2$	${}^\bullet P(x)$	simplification
$\vdash_3$	${}^*P(x)$	axiom 10 + detachment
$\vdash_4$	$y \,\Pi\, x$	simplification of 1
$\vdash_5$	${}^*P(x) \wedge y \,\Pi\, x$	conjunction
$\vdash_6$	${}^*P(x) \wedge y \,\Pi\, x \rightarrow {}^*P(y)$	axiom 9
$\vdash_7$	${}^*P(y)$	detachment
$\vdash_8$	${}^\bullet P(x) \wedge y \,\Pi\, x \rightarrow {}^*P(y)$	deduction theorem
$\vdash_9$	${}^\bullet P(x) \rightarrow [y \,\Pi\, x \rightarrow {}^*P(y)]$	exportation

Axiom 13
$\forall x : {}^{\bullet}P(x) \rightarrow \exists y[y \,\Pi\, x \wedge \exists z[z \,\Pi\, x \wedge y \neq z]]$

Theorem 8 $\forall x : [{}^{\bullet}P(x) \rightarrow \exists y[{}^{*}P(y) \wedge \exists z[{}^{*}P(z) \wedge y \neq z]]$

Proof 8

$\vdash_1$	${}^{\bullet}P(x)$	premise
$\vdash_2$	${}^{\bullet}P(x) \rightarrow \exists y[y \,\Pi\, x \wedge \exists z[z \,\Pi\, x \wedge y \neq z]]$	axiom 13
$\vdash_3$	$\exists y[y \,\Pi\, x \wedge \exists z[z \,\Pi\, x \wedge y \neq z]]$	detachment
$\vdash_4$	$\alpha \,\Pi\, x \wedge \beta \,\Pi\, x \wedge \alpha \neq \beta$	E.I.
$\vdash_5$	${}^{\bullet}P(x) \wedge \alpha \,\Pi\, x$	conjunction of 1, 4 + simplification
$\vdash_6$	${}^{\bullet}P(x) \wedge \alpha \,\Pi\, x \rightarrow {}^{*}P(\alpha)$	theorem 7
$\vdash_7$	${}^{*}P(\alpha)$	detachment
	$\vdots$	(likewise for β)
$\vdash_8$	${}^{*}P(\beta)$	
$\vdash_9$	$\alpha \,\Pi\, x \wedge {}^{*}P(\alpha) \wedge \beta \,\Pi\, x \wedge {}^{*}P(\beta) \wedge \alpha \neq \beta$	conjunction of 4,7,8
$\vdash_{10}$	${}^{*}P(\alpha) \wedge {}^{*}P(\beta) \wedge \alpha \neq \beta$	simplification
$\vdash_{11}$	$\exists y[{}^{*}P(y) \wedge \exists z[{}^{*}P(z) \wedge y \neq z]]$	E.G.
$\vdash_{12}$	${}^{\bullet}P(x) \rightarrow \exists y[{}^{*}P(y) \wedge \exists z[{}^{*}P(z) \wedge y \neq z]]$	deduction theorem

Theorem 9 $\exists x \; {}^{\bullet}P(x) \leftrightarrow \exists y[{}^{*}P(y) \wedge \exists z[{}^{*}P(z) \wedge y \neq z]]$

Proof 9

$\vdash_1$	$\exists x \; {}^{\bullet}P(x)$	premise
$\vdash_2$	${}^{\bullet}P(\alpha)$	E.I.
$\vdash_3$	${}^{\bullet}P(\alpha) \rightarrow \exists y[{}^{*}P(y) \wedge \exists z[{}^{*}P(z) \wedge y \neq z]]$	theorem 8
$\vdash_4$	$\exists y[{}^{*}P(y) \wedge \exists z[{}^{*}P(z) \wedge y \neq z]]$	detachment
$\vdash_5$	$\exists x \; {}^{\bullet}P(x) \rightarrow \exists y[{}^{*}P(y) \wedge \exists z[{}^{*}P(z) \wedge y \neq z]]$	deduction theorem
$\vdash_6$	$\exists y[{}^{*}P(y) \wedge \exists z[{}^{*}P(z) \wedge y \neq z]]$	premise
$\vdash_7$	${}^{*}P(\alpha) \wedge {}^{*}P(\beta) \wedge \alpha \neq \beta$	E.I.
$\vdash_8$	$[{}^{*}P(\alpha) \wedge {}^{*}P(\beta) \wedge \alpha \neq \beta] \rightarrow {}^{\bullet}P(\alpha \oplus \beta)$	axiom 11
$\vdash_9$	${}^{\bullet}P(\alpha \oplus \beta)$	detachment
$\vdash_{10}$	$\exists x \; {}^{\bullet}P(x)$	E.G.
$\vdash_{11}$	$\exists y[{}^{*}P(y) \wedge \exists z[{}^{*}P(z) \wedge y \neq z]] \rightarrow \exists x \; {}^{\bullet}P(x)$	deduction theorem
$\vdash_{12}$	$\exists x \; {}^{\bullet}P(x) \leftrightarrow \exists y[{}^{*}P(y) \wedge \exists z[{}^{*}P(z) \wedge y \neq z]]$	conjunction of 5,12 + material equivalence

9.7 Natural Language Inferences

The following eleven natural languages inferences are (presumably) intuitively valid, but as the proofs show a few of them are formally valid only if an additional proviso is added (see for instance inference 7). I shall comment on the individual proofs to the extent which seems to me to be necessary, and the need for extra provisos will be explained in these comments.

The translations of the natural language premises will occur as premises in the proofs, and where the inferences are formally valid the translations of the consequences will occur as the last step in the proof. Otherwise the natural language consequences will occur together with a suitably added proviso as the last step of the proof. (Again, cf. inference 7.)

Many more inferences could have been studied in detail, but of course I have tried to select such inference patterns which cover the essential logical properties of *plurals*. Or rather, they demonstrate the properties of *distributive* plural constructions; for, as regards the *collective* plural constructions, the whole point is that *no* special inferences can be made for those. However, that does not imply that the distributive cases are the most interesting ones; on the contrary these cases always reduce in various ways to a sum of atomic cases. The proofs of the natural language inferences give examples of how that reduction can be carried out. The collective cases, on the other hand, indicate an expressive power of **LP** over and above "standard" predicate calculus.

Of course, all the "normal" logical inferences can be carried out on expressions involving collective predicates and plural entities, but there is no reason why we should study already well-known inference patterns here.

9.7.1 Potentially Valid Inferences

1. The students die, so every student dies.

Proof 1

$\vdash_1$	$\exists x[{}^{\bullet}student(x) \wedge \forall y[{}^{*}student(y) \leftrightarrow y \,\Pi\, x] \wedge {}^{\bullet}die(x)]$	premise
$\vdash_2$	${}^{\bullet}student(\alpha) \wedge \forall y[{}^{*}student(y) \leftrightarrow y \,\Pi\, \alpha] \wedge {}^{\bullet}die(\alpha)$	E.I.
$\vdash_3$	$\forall x, y[{}^{\bullet}die(x) \rightarrow [y \,\Pi\, x \rightarrow {}^{*}die(y)]]$	theorem 8
$\vdash_4$	${}^{\bullet}die(\alpha) \rightarrow \forall y[y \,\Pi\, \alpha \rightarrow {}^{*}die(y)]$	U.I
$\vdash_5$	${}^{\bullet}die(\alpha)$	simplification of 2
$\vdash_6$	$\forall y[y \,\Pi\, \alpha \rightarrow {}^{*}die(y)]$	detachment
$\vdash_7$	$\forall y[{}^{*}student(y) \leftrightarrow y \,\Pi\, \alpha]$	simplification of 2
$\vdash_8$	$\forall y[{}^{*}student(y) \rightarrow y \,\Pi\, \alpha]$	material equivalence + simplification
$\vdash_9$	$\forall y[{}^{*}student(y) \rightarrow y \,\Pi\, \alpha] \wedge \forall y[y \,\Pi\, \alpha \rightarrow {}^{*}die(y)]$	conjunction of 6,8
$\vdash_{10}$	$\forall y[{}^{*}student(y) \rightarrow y \,\Pi\, \alpha \wedge y \,\Pi\, \alpha \rightarrow {}^{*}die(y)]$	distributive laws for '$\wedge$' and '$\forall$'
$\vdash_{11}$	$\forall y[{}^{*}student(y) \rightarrow {}^{*}die(y)]$	hypothetical syllogism

Comment:
In all its simplicity this is a very important inference pattern, which highlights the nature of the "plural uniqueness condition" introduced by the translation of the_{pl}. The premise is stronger than the consequence, since it introduces not only a universal quantification but also an explicit statement of existence—existence of a proper plural entity, of course. For these reasons "the students die" entails "every student dies", but *not* conversely.

2. The students die, so some students die.

Proof 2

$\vdash_1$	$\exists x[{}^{\bullet}student(x) \wedge \forall y[{}^{*}student(y) \leftrightarrow y \,\Pi\, x] \wedge {}^{\bullet}die(x)]$	premise
$\vdash_2$	$\exists x[{}^{\bullet}student(x) \wedge {}^{\bullet}die(x)]$	simplification

Comment:
Obviously the system makes at least this inference very easy to come by. Again the characteristic feature of the inference is the explicit assertion of existence introduced by the translation of the_{pl}. Of course, the consequence

could not be obtained from "every student dies".

3. Some students die, so some student dies.

Proof 3

$\vdash_1$	$\exists x[^{\bullet}student(x) \wedge {}^{\bullet}die(x)]$	premise
$\vdash_2$	$^{\bullet}student(\alpha) \wedge {}^{\bullet}die(\alpha)$	E.I.
$\vdash_3$	$^{\bullet}student(\alpha)$	simplification
$\vdash_4$	$\forall x[^{\bullet}student(x) \rightarrow {}^{\star}student(x)]$	axiom 10
$\vdash_5$	$^{\bullet}student(\alpha) \rightarrow {}^{\star}student(\alpha)$	U.I.
$\vdash_6$	$^{\star}student(\alpha)$	detachment
	$\vdots$	(likewise for $^{\bullet}die(\alpha)$)
$\vdash_7$	$^{\star}die(\alpha)$	
$\vdash_8$	$^{\star}student(\alpha) \wedge {}^{\star}die(\alpha)$	conjunction
$\vdash_9$	$\exists x[^{\star}student(x) \wedge {}^{\star}die(x)]$	E.G.

Comment:
This proof is perhaps slightly unnatural in the sense that it does not explicitly go "down" to atomic entities, as the natural language consequence intuitively demands. As I have already mentioned the notion of atomicity isn't required from a technical point of view to get the right inferences for this part of the fragment, and so it has simply been left out. If you feel uncomfortable with the proof nonetheless, I can mention here that when transitive verbs have been incorporated you will see similar proofs, where the atomicity-aspect of the consequence is explicitly asserted (see for instance inference 9, p. 133).

4. The students die, so some student dies.

Proof 4

$\vdash_1$	$\exists x[^{\bullet}student(x) \wedge \forall y[^{\star}student(y) \leftrightarrow y \,\Pi\, x] \wedge {}^{\bullet}die(x)]$	premise
$\vdash_2$	$\exists x[^{\bullet}student(x) \wedge {}^{\bullet}die(x)]$	proof 2
$\vdash_3$	$\exists x[^{\star}student(x) \wedge {}^{\star}die(x)]$	proof 3

Comment:
The inference as well as the proof is a straightforward combination of inferences/proofs 2 and 3.

5. The students die, John is a student, so John dies.

Proof 5

$\vdash_1$	$\exists x[^{\bullet}student(x) \wedge \forall y[^{*}student(y) \leftrightarrow y \,\Pi\, x] \wedge {}^{\bullet}die(x)]$	premise 1
$\vdash_2$	$\forall z[^{*}student(z) \rightarrow {}^{*}die(z)]$	proof 1
$\vdash_3$	$^{*}student(john) \rightarrow {}^{*}die(john)$	U.I
$\vdash_4$	$^{*}student(john)$	premise 2
$\vdash_5$	$^{*}die(john)$	detachment

Comment:
In this inference it is the universal quantification introduced by the translation of the_{pl} which comes into use.

6. The students die, John and Mary are students, so John and Mary die.

Proof 6

$\vdash_1$	$\exists x[^{\bullet}student(x) \wedge \forall y[^{*}student(y) \leftrightarrow y \,\Pi\, x] \wedge {}^{\bullet}die(x)]$	premise 1
$\vdash_2$	$\forall z[^{*}student(z) \rightarrow {}^{*}die(z)]$	proof 1
$\vdash_3$	$^{\bullet}student(john \oplus mary)$	premise 2
$\vdash_4$	$^{\bullet}student(john \oplus mary) \wedge \forall z[^{*}student(z) \rightarrow {}^{*}die(z)]$	conjunction
$\vdash_5$	$^{\bullet}student(john \oplus mary) \wedge \forall z[^{*}student(z) \rightarrow {}^{*}die(z)] \rightarrow {}^{\bullet}die(john \oplus mary)$	theorem 6
$\vdash_6$	$^{\bullet}die(john \oplus mary)$	detachment

Comment:
At first glance this inference may seem to be a quite trivial extension of inference 5. However, it is more than that, for the consequence says not only that "John dies and Mary dies", but also implies that "John" and "Mary" are not identical. Theorem 6 allows us to draw the desired inference nonetheless and without too much complication. But it is worth noting that the proof would have been even more trivial without the demand that "John" be different from "Mary".

7. John is a student, Mary is a student, so John and Mary are students.

Proof 7

$\vdash_1$	$^{*}student(john)$	premise 1
$\vdash_2$	$^{*}student(mary)$	premise 2
$\vdash_3$	$^{*}student(john) \wedge {}^{*}student(mary)$	conjunction
$\vdash_4$	$^{*}student(john) \wedge {}^{*}student(mary) \wedge john \neq mary$ $\rightarrow {}^{\bullet}student(john \oplus mary)$	axiom 11
$\vdash_5$	$[{}^{*}student(john) \wedge {}^{*}student(mary)]$ $\rightarrow [john \neq mary \rightarrow {}^{\bullet}student(john \oplus mary)]$	exportation
$\vdash_6$	$john \neq mary \rightarrow {}^{\bullet}student(john \oplus mary)$	detachment

Comment:

This is the first example of an intuitively valid inference, which is not "quite valid" formally. Strictly speaking it is simply invalid, of course. But various logical operations allow us to get a formal consequence (step 6), whose "internal antecedent" is the lacking proviso that "John" is different from "Mary", and whose "internal consequence" is the translation of the natural language consequence. Provided that the translation relation—as envisaged by Link and implemented by me—is as it ought to be, it follows that language users are being slightly imprecise in making inferences like 7. A correct natural language inference could be obtained by replacing the consequence of 7 as in the following inference:

(7') "John is a student, Mary is a student,
so if John is different from Mary,
then John and Mary are students"

which is valid, intuitively as well as formally.

On the other hand we might take the stance that inference 7 in its original version is indeed valid, and that it is the translation relation which has to be revised. In my opinion this is a more reasonable course to choose, and the translation relation will then have to be revised in a way which does not automatically incorporate the "plural presupposition" into all plural verb phrases. This can be very easily achieved by simply letting translations of verb phrases enter the overall translation under the closure-operator instead of under the proper plural operator. Examples of this will be provided when transitive verbs are incorporated.

8. John and Mary are students, so Mary and John are students.

Proof 8

$\vdash_1$	$^{\bullet}student(john \oplus mary)$	premise
$\vdash_2$	$^{\bullet}student(mary \oplus john)$	axiom 2

Comment:
This very trivial inference has been included because Link actually singles out this pattern under the heading *symmetry*. (See example (40) in [Lin83, page 311]). What comes into use here is simply the *commutativity* property of the semi-lattice join-operator.

9. John and Mary are students, Paul is a student, so John and Paul and Mary are students.

Proof 9

$\vdash_1$	$^{\bullet}student(john \oplus mary)$	premise 1
$\vdash_2$	$^{*}student(paul)$	premise 2
$\vdash_3$	$^{\bullet}student(john \oplus mary) \wedge {}^{*}student(paul)$	conjunction
$\vdash_4$	$^{\bullet}student(john \oplus mary) \wedge {}^{*}student(paul) \rightarrow {}^{\bullet}student(john \oplus mary \oplus paul)$	theorem 3 + U.I.
$\vdash_5$	$^{\bullet}student(john \oplus mary \oplus paul)$	detachment

Comment:
This inference illustrates how the previously discussed inference-pattern of *expansion* comes out in my implementation.

10. John and Paul and Mary are students, so Paul and Mary are students.

Proof 10

$\vdash_1$	${}^{\bullet}student(john \oplus mary \oplus paul)$	premise
$\vdash_2$	${}^{*}student(john \oplus mary \oplus paul)$	axiom 10 + detachment
$\vdash_3$	${}^{*}student(john) \wedge {}^{*}student(paul \oplus mary)$	theorem 2 + replacement
$\vdash_4$	${}^{*}student(paul \oplus mary)$	simplification
$\vdash_5$	${}^{*}student(paul) \wedge {}^{*}student(mary)$	theorem 2 + replacement
$\vdash_6$	${}^{*}student(paul) \wedge {}^{*}student(mary) \wedge paul \neq mary \rightarrow {}^{\bullet}student(paul \oplus mary)$	axiom 11 + U.I.
$\vdash_7$	$paul \neq mary \rightarrow {}^{\bullet}student(paul \oplus mary)$	exportation + detachment

Comment:
This inference illustrates how the previously discussed inference-pattern of *contraction* comes out in my implementation. As mentioned in that discussion (cf. p. 65), the pattern stated by Link is actually not valid without an added proviso, namely that the involved entities be different. This proviso has been incorporated as an antecedent in the final step of the proof, and with this addition the inference is valid. See also the remarks on inference 7, which apply here as well.

11. John and Mary are students, so Mary is a student.

Proof 11

$\vdash_1$	${}^{\bullet}student(john \oplus mary)$	premise
$\vdash_2$	${}^{*}student(john \oplus mary)$	axiom 10 + detachment
$\vdash_3$	${}^{*}student(john) \wedge {}^{*}student(mary)$	theorem 2 + replacement
$\vdash_4$	${}^{*}student(mary)$	simplification

Comment:
Like inference 3, although perhaps in a more natural manner, this inference illustrates once more how atomic entities can be "extracted" from plural entities of which they are a part.

I now go on to list a few more interesting inferences, but without proofs:

12. John and Mary are the students, the students die, so

1. John and Mary die

2. John dies

13. John and Mary are students, the students die, so

1. John and Mary die

2. John dies

14. The masters and some students die, so

1. the masters die

2. some masters die

3. some master dies

4. some students die

5. some student dies

6. some master and some student die

9.7.2 Invalid Inferences

As a final illustration of the present system, I list some invalid inferences. I shall not here construct the models to show why each inference is invalid, but nonetheless this list can throw some further light on the system—and on the problem of *plurals* in general.

15. The students die, so the student dies.
16. Some students die, so the students die.
17. Students die, so every student dies.
18. Some student dies, John is a student and John dies, so some students die.
19. John and Mary are students, some students die, so

1. John and Mary die

2. John dies

20. John is a student, Mary is a student, John is Mary, so John and Mary are students.

10 Extending Link's Framework to Transitive Verbs

In [Lin83] Link has little to say about transitive verbs, which may imply that they did not seem to him to involve any particular problems. When a transitive verb phrase such as "see" is combined with its object, e.g. "a college", an intransitive verb phrase "see a college" is formed; so it would not seem unnatural to suppose that such non-basic intransitive verb phrases could be dealt with simply along the lines of basic intransitive verbs. It is suggestive, though, that Link's paper does not contain a single example of such a phrase under the closure- or plural-operator.

One thing is clear, however: there *are* transitive verbs which give rise to unambiguously distributive phrases—take again "see a college". A brief but unfortunately somewhat imprecise remark by Link indicates that he would acknowledge this:

> "Most of the basic count nouns like *child* are taken as distributive, similarly IV phrases like *die* or *see*."
>
> [Lin83, page 318]

I think we may assume that Link means IV phrases *formed* from "see".

In my attempt to extend [Lin83] to transitive verbs[77] I shall deviate a lot more from his framework than I did in the previous section. This is not meant to imply that it is impossible to deal with transitive verbs by subsuming them under the treatment of basic intransitive verb phrases. But even if such an approach is feasible, there are certain questions which it would leave unaddressed[78]. Furthermore I believe that my implementation has the merit of providing logical forms which are not quite as hard to understand for the human reader and more computationally tractable than the logical forms which emerge from a straightforward extension of Link's framework. This will soon be illustrated by various examples.

It should perhaps also be noted that model-theoretic considerations play a far more prominent role in the following account than in previous sections. The reason for this shift of emphasis is that things are getting far more complicated now, so I feel that the concepts involved had better be

[77]More precisely, only bitransitive verbs are dealt with in detail; when this has been done I shall offer some deliberations on the subject of how to generalise the account to n-transitive verbs. Cf. section 10.5.

[78]For instance, what role meaning postulate 1 in the following would have to play.

sorted out in detail at a model-theoretic level before setting out to give an axiomatic system.

My point of departure, then, will be a tacit assumption of Link's model-theoretic account. For the purposes of the following discussion I shall also go back to using Link's translation relation, albeit with one exception. The exception is distributive intransitive verb phrases, which will enter translations under the closure-operator rather than the proper plural operator (in accordance with earlier discussion).

Finally, let me mention that I introduce the following bit of extra notation: if a is an object in a model, then $\vec{a}$ is to be a constant denoting a, i.e. $\| \vec{a} \| = a$. As for extra terminology I shall use the expression *D-ambiguous* to describe verbs that are systematically ambiguous between distributive and collective readings.

10.1 A Discussion of Link 83 on Transitive Verbs

As a starting point let us examine somewhat closer one of the central meaning postulates which emerges from [PTQ]. It will be recalled that Montague in [PTQ] stated some meaning postulates on purely extensional transitive verbs, which applied to a verb phrase like "see a college" would amount to[79]

Meaning Postulate 1 *(Applied to "see a college"):*
$\forall x\, [see(\lambda Q \exists z[college(z) \wedge Q(z)])(x) \leftrightarrow \exists z[college(z) \wedge see_*(x,z)]]$

This meaning postulate makes use of the fact that from a two-place predicate δ of type $\langle\langle\langle e,t\rangle,t\rangle,\langle e,t\rangle\rangle$ we can "derive" a two-place predicate δ_* of type $\langle e,\langle e,t\rangle\rangle$ [80]. If the predicate δ is purely extensional, then δ and δ_* stand in a one-to-one correspondance to each other.

So the meaning postulate tells us that for purely extensional verbs we don't need the cumbersome translations of the form given in the lefthand side of its biconditional. We can do with the simpler forms indicated in its

[79] I have taken a few liberties here. Montague doesn't actually state any such meaning postulate directly, but rather a number of meaning postulates and definitions, which taken together validate this meaning postulate. Furthermore, since coordinates like *possible worlds* and *moments of time* are not under consideration here, I have dropped the modal operator $\Box$ (in whose scope the entire meaning postulate should be, otherwise); and for the same reason I have dropped the intensional operators which occur in the corresponding [PTQ] meaning postulates.

[80] For further details I refer to [DWP79, 219–227], where a very thorough account of such "derived" predicates δ_* is given, and their relation to so-called sublimation-concepts of the form $\lambda Q \exists z[college(z) \wedge Q(z)]$ is explained.

righthand side—forms which are easier to understand for the human reader as well as computationally more tractable.

Now there is as far as I can see no reason to believe that intensionality can affect our perception of distributive verbs[81], and at any rate I am currently working within a purely extensional fragment. So it seems all the more desirable to be able to apply meaning postulate 1 in some suitably modified form to transitive verbs under the closure-operator, too. That is, if we consider for instance the following translation (assuming that "see a college" is unambiguously distributive):

John and Mary see a college
$$\triangleright {}^{*}see(\lambda Q \exists z[college(z) \wedge Q(z)])(john \oplus mary)$$

it ought to be possible to "dissolve" the translation along the lines indicated by meaning postulate 1.

One prerequisite for being able to do this is that it must be possible to apply the closure-operator directly to two-place predicates of the form δ_*. That is, we must ensure that for instance ${}^{*}see_{*}$ is a meaningful expression. This raises a few technical questions (apart from the more serious semantic concerns to be investigated in due course). Link has already made the closure-operator "heterogeneous" in the sense that it can be applied to expressions of differing types (Cf. p. 67). There is no reason why we should not follow him in this. With respect to our particular goals we need to modify his rule into the following:

$${}^{*}\varsigma,\ {}^{\bullet}\varsigma \in \mathrm{ME}_{\langle\tau\rangle},\ \text{for}\ \tau \in \{\langle e,t\rangle, \langle e,\langle e,t\rangle\rangle\},\ \varsigma \in \mathrm{ME}_{\langle\tau\rangle}$$

This modification leads on to another condition for generalising meaning postulate 1: when the closure-operator is applied to one- or two-place predicates, the semantic construction of denotations for both kinds of predicate must follow one and the same general rule. This requirement will indeed be met by the definitions established later on. For the time being think of the closure of a two-place predicate, e.g. ${}^{*}see_{*}$, as being merely the intuitive counterpart of the closure of a one-place predicate, e.g. ${}^{*}die$.

Of course, it is still the case that there are really two different closure-operators involved here. I follow Link, however, in using one and the same

[81]Nor have I found any evidence to that effect in the literature on the subject. If, however, it should turn out in the long run that intensionality *does* affect our notion of distributivity, no harm will come from having started the investigation within a purely extensional framework. For then it will be all the clearer exactly what the involvement of intensional constructions has to contribute to the notion.

symbol for both, and I think this is justified by their very close affinity. No problems can arise from this if one always remembers that in the expression

$$^{*}see(\lambda Q\exists z[college(z) \wedge Q(z)])(\alpha)$$

the closure operator ranges over the entire translation of the intransitive verb phrase, i.e. over

$$(see(\lambda Q\exists z[college(z) \wedge Q(z)]))$$

whereas in an expression of the form $^{*}see_{*}(\alpha, \beta)$, or equivalently,

$$^{*}see_{*}(\beta)(\alpha)$$

the closure-operator ranges over see_{*} only.

With these tedious but necessary precautions in mind we may now at least tentatively consider $^{*}see_{*}$ to be a meaningful expression. On that basis we can now reconsider meaning postulate 1. It is tempting to try a straightforward reformulation like the following:

$$\forall x[^{*}see(\lambda Q\exists z[college(z) \wedge Q(z)])(x) \leftrightarrow \exists z[college(z) \wedge {}^{*}see_{*}(x, z)]]$$

But here caution is called for. For if we uncritically apply this to all (distributive) transitive verbs, and in particular to those with a plural noun phrase as subject term, then for instance the translation of the sentence

"some students see a college"

i.e.

$$\exists x[^{\bullet}student(x) \wedge {}^{*}see(\lambda Q\exists y[college(y) \wedge Q(y)])(x)]$$

would be equivalent to

$$\exists x[^{\bullet}student(x) \wedge \exists y[college(y) \wedge {}^{*}see_{*}(x, y)]]$$

which means that it was one and the same college, which each of the students saw—but the sentence may be true even if each student saw a different college.

Obviously, we are faced with an implicit scope problem here. In effect meaning postulate 1 gives the object of a transitive verb wide scope, and this yields a logical form for the sentence in question, which is certainly too restricted to represent its meaning in general. Given a sensible meaning of ${}^*\delta_*$, where ${}^*\delta$ is the translation of some transitive verb phrase, it still remains to somehow revise meaning postulate 1 into a more suitable form. But before trying to face up to this task there is another notion which should be examined more closely, namely the notion of *mixed extension*.

10.1.1 Mixed Extension vs. Ambiguity

Given that some intransitive verb phrases are unambiguously distributive and others unambiguosly collective, there still remains a large group which are neither the one nor the other, such as "build a college". I suppose that the immediate and natural reaction to this fact would be to consider the remaining verbs as ambiguous in a way indicated by the two following logical forms:

1. $build(\lambda Q \exists y[{}^*college(y) \wedge Q(y)])$ (the *collective* reading).
2. ${}^*build(\lambda Q \exists y[{}^*college(y) \wedge Q(y)])$ (the *distributive* reading).

This is not the approach taken by Link, however. Instead he attributes to this group the property of having *mixed extension*. To explain what this means I shall stipulate a tiny model for the two-place predicate $carry_*$: suppose we have a model[82] such that

$$\begin{aligned} \|\,(\text{John})'\,\| &= a, \\ \|\,(\text{Mary})'\,\| &= b, \\ \|\,(\textit{the piano})'\,\| &= \beta \end{aligned}$$

and that the semantic value for "carry" is given by

$$\begin{aligned} \|\,carry_*\,\| = \{&(a,\alpha),(a,\beta),(b,\alpha \uplus \delta),(a \uplus b,\beta), \\ &(a \uplus b,\epsilon),(a \uplus f,\alpha \uplus \delta),(e \uplus f,\gamma)\} \end{aligned}$$

[82]Intuitively the model may be unsatisfactory, since it makes it true that John and Mary carried the piano together, and at the same time John carried the piano all by himself. It is part of our real-world knowledge that these events cannot take place independently at the same time. But currently I am merely examining the semantics at its "pre-pragmatic" level, i.e. at a purely model-theoretic level, and I choose examples of this type in order to demonstrate certain logical properties of the theory. See also the discussion of the determiner "the" on p. 120 f., and in particular footnote 116.

In fact the problem encountered may have some more general implications, and you will meet variants of it later in the text; the nature of this type of problem will be more closely addressed in section 10.10.

Now in a sense "carry" is not seen to be ambiguous—there is just one such verb with just one meaning, namely the one given above. On the other hand, the sentence "John and Mary carry the piano" *is* in a way ambiguous w.r.t. the above model; more precisely, we have the situation that the sentence may be true (respectively, false) for the following reasons:

$$\| \phi \| = \begin{cases} 1, & \text{for } (a \uplus b, \beta) \in \| \, carry_* \, \| \\ 1, & \text{for } (a, \beta) \in \| \, carry_* \, \| \\ & \text{and } (b, \beta) \in \| \, carry_* \, \| \\ 0, & \text{otherwise} \end{cases}$$

where

$$\phi = (\text{John and Mary carry the piano})'$$ [83].

In the case of the model above, the sentence happens to be true because $(a \uplus b, \beta) \in \| \, carry_* \, \|$ (corresponding to the collective reading "John and Mary carry the piano together").

Why is it, then, that Link doesn't deal with such verbs as being plainly ambiguous? Link is not very explicit on this point, but the argument he offers can be reconstrued as follows[84]:

In the case of unambiguously distributive verbs, we have a number of valid inferences; for instance,

> "John, Paul, George, and Ringo are pop stars,
> so Paul is a pop star" is a valid inference.

Symbolically,

$$\vdash {}^{\bullet}P(a \oplus b \oplus c \oplus d) \Rightarrow {}^{\star}P(b)$$

where P=(pop star)', a=(John)', etc.

However, for collective predicates like "convene" we have of course no analogous inference. And, Link claims, that observation holds for the *mixed*

[83] The use of the $(\cdot)'$ notation presupposes that the sentence in question has only one translation—and this holds w.r.t. Link's translation relation.

[84] Cf. [Lin83, pages 309–310].

extension verbs, too—for instance,

> "John and Mary carried the piano,
> so John carried the piano and Mary carried the piano"

is *not* a valid inference, for as we have seen in the discussion of the model above the antecedent may be true, while the consequence is false; that is, the antecedent does not lead infallibly to the consequence. Therefore, Link concludes, the class of verbs which permit of both distributive and collective readings should be treated as having *mixed extension* rather than as being ambiguous in the manner indicated by my tentative translations (1) and (2) of "build a college" above.

Although intuitively tempting I think Link's line of argumentation is inside out when compared with the project of formal semantics for natural languages[85]. Within the framework of that project, most natural language sentences are translated into more than one logical form. Each logical form can by itself enter unambiguosly into certain valid inferences. It would seem natural, then, to speak of natural language inferences as being valid *relative to particular readings* of the sentences involved. Otherwise, there wouldn't be many valid natural language inferences left. But for the sake of argument, let us suppose that we accept the rather restricted notion of validity for natural language inferences as implied by Link. Even so the conclusion that the verbs in question are not ambiguous does not follow. For—and this is really a very trivial point—we may still choose to treat them as ambiguous, without this making the disputed inference valid in Link's sense. Furthermore, on Link's line of reasoning classical cases of ambiguity such as "every man loves some woman" would no longer be treated as ambiguous, because that sentence does not lead infallibly to "there is one woman such that every man loves her"—i.e. the wide scope reading of the sentence would have to go.

[85] As I see it, Link's argumentation is here of a *philosophical* rather than *semantical* nature. While there is in principle nothing wrong with that, it does in the present context clash with Link's own statements on methodology in [Lin83] (Cf. section 11.2). There, it is repeatedly stressed that "language itself" must be our guide in doing formal semantics, rather than philosophical or ontological considerations.

On these grounds I reject the reasons given by Link for treating the verbs in question by *mixed extension*[86].

Even so, it perhaps still remains to make a more constructive case for the approach which I have so far been hinting at—i.e. their treatment as systematically ambiguous along the lines indicated above. Now I am inclined to believe that the real reason why Link resists the "ambiguity approach" is different from the one he actually offers. First, in a strictly compositional semantics we would naturally feel uncomfortable with sentence-ambiguity arising form the properties of a single word rather than from structural properties of the sentence as a whole. Second, from a strictly linguistic point of view there seems to be no good reason for considering sentences like "John and Mary carry the piano" to be ambiguous, as it is implied by the "ambiguity approach"[87].

I think such worries are not entirely without reason, and that they should not be brushed aside too lightly. But in my opinion there are other considerations which override them. As it should be clear from my statement of the truth-conditions for "John and Mary carry the piano" above, the semantic ambiguity of that sentence is of a quite systematic and general nature. The

---

[86]Maybe it should be pointed out that certain remarks by Link on "type ambiguity" of *mixed extension verbs* do *not* bear on the present discussion in the way one might at first expect. Link states that

> "Sums and collections are similar ... in that they both are just individuals, as concrete as the individuals which serve to define them, and of the same logical type as these. The latter feature is important because **there is no systematic type ambiguity inherent** in predicates like *carry*, *build*, *demolish*, *defend*, etc. [Lin83, page 305]

However, what is at stake here is not the sort of ambiguity which I have been discussing so far; these remarks by Link are intended to show that it is legitimate to conflate objects in A and objects in $E \setminus A$ into one and the same type. But w.r.t. the feasibility of subcategorising verbs as *+distributive* or *-distributive* (which amount to assigning different categories to these classes of verbs), Link later goes on to say that

> "...distributive predicates working on plural terms have to be starred, all the other predicates must not be. For distributivity seems to be a lexical feature. If we have a formal translation procedure the predicates have to be subcategorized accordingly." [Lin83, page 310]

[87]In a [PTQ]-style framework the "ambiguity approach" would force us to make the sentence in question *syntactically* ambiguous—in order to get the two different translations with and without the closure-operator applied to the verb.

sentence may be true either in terms of its *collective reading* or in terms of its *distributive reading*[88].

It is a basic tenet of Montague Grammar that semantics is in a sense prior to syntax—that is, the syntax we give is ultimately semantically motivated. So within that paradigm, at least, it is perfectly consistent to let semantic concerns have the final say.

From a methodological point of view, these observations could by themselves justify that we provide (at least) two logical forms for sentences with D-ambiguous verbs. However, since the worries about taking that approach were of a specifically linguistic nature, it should also be mentioned that certain broader linguistic concerns seem to recommend the "ambiguity approach". I conclude this section by stating those points:

- Even if we insist on maintaining the notion of *mixed extension*, we still have an obligation to account for the non-ambiguity of

  "John and Mary *each* built a college"

  "John and Mary built a college *together*".

  Proceeding from *mixed extension*, that will presumably lead to the introduction of extra gadgetry—and this will be of an *ad hoc* nature. But on the "ambiguity approach", the solution to this problem is straightforward and of a systematic nature (Cf. p. 122).

- In section 10.6 it will become clear that the distributive-collective distinction can be extended into a still more fine-grained subclassification of transitive verbs.

  As far as I can see, that subclassification cannot be accounted for if we adopt the notion of *mixed extension*, whereas the "ambiguity approach" lends itself to that task in a natural manner.

- In general I am not here concerned with "psychological reality", but it may still be in order to draw attention to one observation having to do with that issue: when somebody says "John and Mary carried a piano", he must be assumed to mean *either* that they did it together

[88] Technically, it may be true on both readings at the same time. They are not mutually exclusive, and there is no reason why they need be so in order to qualify for being called "ambiguous"; for instance, the wide-scope and the narrow-scope readings of "every man loves some woman" are not mutually exclusive (See also section 10.2.4).

or that each of them carried a piano, but not both. That is, although not mutually exclusive from the logical point of view, a speaker will presumably have only one of the two possible readings in mind when uttering that sentence[89].

10.2 Defining the Closure-Operator for Two-Place Predicates

As my point of departure for defining $^*\delta_*$, where δ_* denotes a two-place relation, I shall choose the transitive verbs "love" and "build" (respectively, the two-place predicates $love_*$ and $build_*$). This choice is quite deliberately not neutral. For they share a rather important property, namely the property that *they are always distributive over their object*[90]. That is, from

"John and Mary love Paul and Sue"

it follows not only that

"John loves Paul and Sue,
and Mary loves Paul and Sue"

but also that

"John and Mary love Paul,
and John and Mary love Sue".

Similar results obtain for a sentence like "John and Mary build a college and a chapel"[91]. Not all D-ambiguous verbs have this property, and so the following account is not completely general. I shall address the implicated problems in due course; here I simply mention them lest the reader should think that my examples are unduly biased.

---

[89]This case should not be confused with the rather different problems with the sentence "John and Mary carried *the* piano". Cf. footnote 82.

[90]Steve Pulman has pointed out to me that "love" may have collective readings over its object phrase as in "Mary loves bread and butter". However, as long as "love" is restricted to descibe inter-personal relationships only the assumption seems to hold, and at any rate the *logical* points being made in the following hold regardless of the exact *linguistic* status of "love" in these respects. But if it seems preferable, "love" could be substituted by "kill" (I have not been able to find any counter-examples to "kill" being always distributive over its object phrases).

[91]This even holds for the collective reading of the sentence; "John and Mary build a college and a chapel [together]" entails "John and Mary build a college [together] and John and Mary build a chapel [together]".

10.2.1 Distributive Transitive Verbs

Consider a model $\mathcal{M}$ with particular respect to the basic two-place predicate $love_*$:

$$\| love_* \| = \{(c,\epsilon),(d,\beta),(d,\gamma),(d,\epsilon),(e,\beta),(e,\epsilon)\}$$

Given this extension of $love_*$ in $\mathcal{M}$, what should the semantic value of $^*love_*$ be—that is, what will the natural definition of $^*love_*$ be? To gain a merely intuitive answer to this question, we could reason along the following lines: "since c loves ϵ and d loves ϵ, c and d love ϵ"; "since d loves β and d loves γ, d loves β and γ; and so forth through all the obvious combinations. Proceeding systematically we would then arrive at the following extension for $^*love_*$:

$$\begin{aligned} \| {}^*love_* \| \;=\;& \| love_* \| \cup \\ & \{(d,\beta \uplus \gamma),(d,\gamma \uplus \epsilon),(d,\beta \uplus \epsilon), \\ & (d,\beta \uplus \gamma \uplus \epsilon),(e,\beta \uplus \epsilon), \\ & (d \uplus e,\beta),(c \uplus d,\epsilon),(c \uplus e,\epsilon), \\ & (d \uplus e,\epsilon),(c \uplus d \uplus e,\epsilon), \\ & (d \uplus e,\beta \uplus \epsilon)\} \end{aligned}$$

This construction has an immediate linguistic rationality (apart from satisfying basic intuition w.r.t. the verb "love"), namely the fact that it *comprises* the set of true (atomic) sentences with this verb. That is, every such sentence, whether in the singular or in the plural, will be true in virtue of its subject and its object (in that order[92]) denoting one of the ordered pairs in $\| {}^*love_* \|$. Thus the meaning given for $^*love_*$ is perfectly natural and linguistically straightforward. A general formulation of the construction of $^*\delta_*$, where δ translates a transitive verb, is given by the following recursive semantic rules:

Definition 1 (Semantic Definition of $^*\delta_*$)

1. *If* $(x,y) \in \| \delta_* \|$, *then* $(x,y) \in \| {}^*\delta_* \|$
2. *If* $(x,y) \in \| {}^*\delta_* \|$, *and* $(x,z) \in \| {}^*\delta_* \|$, *then* $(x,y \uplus z) \in \| {}^*\delta_* \|$

[92] I am ignoring the passive forms of the verb.

3. If $(x,y) \in \| {}^*\delta_* \|$, *and* $(z,y) \in \| {}^*\delta_* \|$, *then* $(x \uplus z, y) \in \| {}^*\delta_* \|$

4. Nothing else is in $\| {}^*\delta_* \|$.

Two propositions follow immediately from this definition[93]:

Proposition 1 $\models \forall x,y,z\,[{}^*\delta_*(x, y \oplus z) \leftrightarrow {}^*\delta_*(x,y) \wedge {}^*\delta_*(x,z)]$

Proposition 2 $\models \forall x,y,z\,[{}^*\delta_*(x \oplus y, z) \leftrightarrow {}^*\delta_*(x,z) \wedge {}^*\delta_*(y,z)]$

These propositions mirror the first intuitive construction of ${}^*love_*$ above, and show that that construction has given us exactly those properties characteristic of distributive predicates, namely *dissectiveness* and *cumulativity*.

10.2.2 An Attempt to Define Closure of Two-Place Predicates in Terms of Closure of One-Place Predicates

So far, the results of the investigation into transitive verbs have in my opinion a certain naturalness about them which is rather convincing. On the other hand, in [Lin83] it was apparently the intention to define closure of two-place predicates in terms of the closure of one-place predicates—a principle, which undeniably has something attractive about it. So let me scrutinise whether this principle carries over into the present framework.

It is well known that a two-place predicate, which translates an extensional transitive verb in the singular, can be expressed in terms of one-place predicates by using lambda-abstraction:

$$\forall x,y\,[\delta_*(x,y) \leftrightarrow (\lambda z\ \delta_*(z,y))(x) \wedge (\lambda z\ \delta_*(x,z))(y)]$$

For instance, given $\mathcal{M}$ above, clearly

$$\models_{\mathcal{M}} \forall x,y\,[love_*(x,y) \leftrightarrow (\lambda z\ love_*(z,y))(x) \wedge (\lambda z\ love_*(x,z))(y)]$$

In our present context, then, it might be tempting to try to define closure for two-place predicates of the form δ_* in terms of the closure of their corresponding lambda-abstracted one-place predicates.

To gather a precise picture of the meaning of expressions like $\lambda z\ love_*(z,\vec{\beta})(e)$, and in particular ${}^*(\lambda z\ love_*(z,\vec{\beta}))(e)$, consider some examples w.r.t. $\| love_* \|$ in $\mathcal{M}$; clearly,

[93] Of course, they are to be theorems (or axioms) within the final axiomatisation of the entire system; but at present I do not wish to state them as such (and I have made no assumptions w.r.t. the completeness of my axiomatisations, so I am not automatically entitled to do so, either).

- $\| \lambda x\ love_*(\vec{d}, x) \| = \{\beta, \gamma, \epsilon\}$
- $\| \lambda x\ love_*(\vec{e}, x) \| = \{\beta, \epsilon\}$
- $\| \lambda x\ love_*(x, \vec{\beta}) \| = \{d, e\}$
- $\| \lambda x\ love_*(x, \vec{\gamma}) \| = \{d\}$
- $\| \lambda x\ love_*(x, \vec{\epsilon}) \| = \{c, d, e\}$

It is a straightforward matter to construct the values of the closures of the one-place predicates involved:

- $\| {}^*(\lambda x\ love_*(\vec{d}, x)) \| = \{\beta, \gamma, \epsilon, \beta \uplus \gamma, \beta \uplus \epsilon, \gamma \uplus \epsilon, \beta \uplus \gamma \uplus \epsilon\}$
- $\| {}^*(\lambda x\ love_*(\vec{e}, x)) \| = \{\beta, \epsilon, \beta \uplus \epsilon\}$
- $\| {}^*(\lambda x\ love_*(x, \vec{\beta})) \| = \{d, e, d \uplus e\}$
- $\| {}^*(\lambda x\ love_*(x, \vec{\gamma})) \| = \{d\}$
- $\| {}^*(\lambda x\ love_*(x, \vec{\epsilon})) \| = \{c, d, e, c \uplus d, c \uplus e, d \uplus e, c \uplus d \uplus e\}$

Define the set X to be the set of pairs such that

$$\begin{aligned} X \ = \ & \{(x,y) | \, x \in \| {}^*(\lambda z\ love_*(z, y)) \|\} \cup \\ & \{(x,y) | \, y \in \| {}^*(\lambda z\ love_*(x, z)) \|\} \end{aligned}$$

Then by instantiating $\vec{d}$, $\vec{e}$ for x and $\vec{\beta}$, $\vec{\epsilon}$, $\vec{\gamma}$ for y in this schema, it should be easy to see from the above values that we get a subset in $\| {}^*love_* \|$, namely

$$\begin{aligned} &\{(d,\beta), (d,\gamma), (d,\epsilon), (e,\beta), (e,\epsilon), (c,\epsilon), (d, \beta \uplus \gamma), \\ &(d, \gamma \uplus \epsilon), (d, \beta \uplus \epsilon), (d, \beta \uplus \gamma \uplus \epsilon), (e, \beta \uplus \epsilon), (d \uplus e, \beta), \\ &(c \uplus d, \epsilon), (c \uplus e, \epsilon), (d \uplus e, \epsilon), (c \uplus d \uplus e, \epsilon)\} \end{aligned}$$

but the pair $(d \uplus e, \beta \uplus \epsilon)$ (cf. the set $\| {}^*love_* \|$ on p. 97) is conspicuously lacking among these. The simple reason for this is that expressions of the form ${}^*(\lambda z\ love_*(z, \alpha))$ and ${}^*(\lambda z\ love_*(\alpha, z))$ only "work" when α is atomic. In fact, it is always the case that[94]

$$\forall y \in E \setminus A : \| {}^*(\lambda z\ love_*(z, y)) \| = \| {}^*(\lambda z\ love_*(y, z)) \| = \emptyset$$

[94]Recall that A is the set of atomic entities and E is the semi-lattice closure of A. Cf. section 9.2.

What I am getting at is that $\| {}^*love_* \|$ cannot be defined in terms of the closure of the corresponding lambda-abstracted one-place predicates alone[95].

10.2.3 Lambda-Abstraction and the Closure-Operator

The discussion of the previous section should make it reasonably clear that[96]

- $\not\models_{\mathcal{M}} \lambda x\, ({}^*love_*(x, \mu)) = {}^*(\lambda x\, love_*(x, \mu))$;
  thus $d \uplus e \in \| \lambda x\, ({}^*love_*(x, \beta \uplus \epsilon)) \|$,
  but $d \uplus e \notin \| {}^*(\lambda x\, love_*(x, \beta \uplus \epsilon)) \|$
  (which is an empty set);
- $\not\models_{\mathcal{M}} \lambda x\, ({}^*love_*(\nu, x)) = {}^*(\lambda x\, love_*(\nu, x))$;
  thus $\beta \uplus \epsilon \in \| \lambda x\, ({}^*love_*(d \uplus e, x)) \|$,
  but $\beta \uplus \epsilon \notin \| {}^*(\lambda x\, love_*(d \uplus e, x)) \|$
  (which is also an empty set).

On the other hand, the following rules w.r.t. lambda-conversion still obtain:

- $\models_{\mathcal{M}}\ love_*(\nu, \mu) \leftrightarrow \lambda x\, (love_*(x, \mu))(\nu)$
- $\models_{\mathcal{M}}\ love_*(\nu, \mu) \leftrightarrow \lambda x\, (love_*(\nu, x))(\mu)$
- $\models_{\mathcal{M}}\ {}^*love_*(\nu, \mu) \leftrightarrow \lambda x\, ({}^*love_*(x, \mu))(\nu)$
- $\models_{\mathcal{M}}\ {}^*love_*(\nu, \mu) \leftrightarrow \lambda x\, ({}^*love_*(\nu, x))(\mu)$

These results are of course not restricted to the model $\mathcal{M}$ or the verb "love". They emerge from the very definition of ${}^*\delta_*(x, y)$[97] and hence carry

[95] For theoretical completeness it should perhaps be noted that the notion of ${}^*love_*$, albeit very natural, is not *indispensable*, in the strictest meaning of this word. We could translate a sentence of the form "d and e love β and ϵ" into

$$ {}^*(\lambda x\ love_*(x, \vec{\beta}))(\vec{d} \oplus \vec{e}) \wedge {}^*(\lambda x\ love_*(x, \vec{\epsilon}))(\vec{d} \oplus \vec{e}) \wedge $$
$$ {}^*(\lambda x\ love_*(\vec{d}, x))(\vec{\beta} \oplus \vec{\epsilon}) \wedge {}^*(\lambda x\ love_*(\vec{e}, x))(\vec{\beta} \oplus \vec{\epsilon}) $$

Obviously, this "strategy" would lead to a combinatorial explosion in the complexity of the output logical forms as a function of the number of terms in the input sentence. Apart from the disastrous practical consequences of this, it is certainly a strong warning signal of some failure of grasping a general rule of language—indeed the very sort of rule that the closure-operator is intended to account for. So this sort of approach would annihilate any linguistic rationale for using the closure-operator all.

[96] In the following, I use μ, ν for any constants of type $\langle e \rangle$.

[97] Of course, the closure-oparator (*) applies to δ_* only and not to the entire expression $\delta_*(x, y)$. I merely include the argument (x, y) to make it clear that we have to do with a two-place predicate.

over to all such expressions w.r.t all models. There are two important consequences of this:

First,

> **lambda-conversion is restricted from taking place, when the lambda-abstracted expression is *within* the scope of the closure-operator, and the argument of the expression is *outside* its scope.**

Although this places a restriction on the use of the lambda-operator in our logical language, it actually enhances the expressive power of this, as will be shown in section 10.6.

Second, we cannot directly apply the axioms and theorems for the closure of one-place predicates to the closure of two-place predicates. This would have been the case if $^*\delta_*(x, y)$ had been directly definable in terms of $^*(\lambda z\, \delta_*(z, y))(x)$ and $^*(\lambda z\, \delta_*(x, z))(y)$—but that is not possible, as we have seen.

However, the implications of this are not as serious as they may seem at first glance. From the propositions (1) and (2) it can easily be proved that all the axioms (and *a fortiori* all the theorems) for one-place predicates under the closure-operator have an obvious "dual counterpart" for two-place predicates. For instance, axiom 8 for one-place predicates

$$\forall x, y\, [^*P(x) \wedge {}^*P(y) \rightarrow {}^*P(x \oplus y)]$$

has the obvious dual counterpart

(i) $\forall x, y, z\, [^*\delta_*(x, z) \wedge {}^*\delta_*(y, z) \rightarrow {}^*\delta_*(x \oplus y, z)]$
(ii) $\forall x, y, z\, [^*\delta_*(x, z) \wedge {}^*\delta_*(x, y) \rightarrow {}^*\delta_*(x, z \oplus y)]$

Actually, little work has to be done to get all the desirable theorems for two-place predicates under the closure-operator, as it will become clear in the final axiomatisation[98].

[98] In fact the general definition of closure on p. 105 makes it possible to do with just one set of generalised axioms. See also my comments on the axiomatisation on p. 125.

Let me summarise the results of this section:

Proposition 3
For all two-place predicates δ_*:

1. $\models \forall x, y\,[\delta_*(x,y) \leftrightarrow \lambda z\,(\delta_*(z,y))(x)]$
2. $\models \forall x, y\,[\delta_*(x,y) \leftrightarrow \lambda z\,(\delta_*(x,z))(y)]$
3. $\models \forall x, y\,[{}^*\delta_*(x,y) \leftrightarrow \lambda z\,({}^*\delta_*(z,y))(x)]$
4. $\models \forall x, y\,[{}^*\delta_*(x,y) \leftrightarrow \lambda z\,({}^*\delta_*(x,z))(y)]$
5. $\not\models \forall x, y\,[{}^*\delta_*(x,y) \leftrightarrow {}^*(\lambda z\,(\delta_*(z,y)))(x)]$
6. $\not\models \forall x, y\,[{}^*\delta_*(x,y) \leftrightarrow {}^*(\lambda z\,(\delta_*(x,z)))(y)]$

10.2.4 D-Ambiguous Transitive Verbs

The treatment of verb phrases which can be read both "collectively" and "distributively" follow in a rather straightforward manner from the decision to treat them as strictly ***ambiguous*** between those readings, and from the definition of ${}^*\delta_*(x,y)$. But there still are a few separate points to be cleared up for these verbs.

I shall assume for the moment that the collective reading is in a sense the basic one, and that this one can include pairs in $A \times A$ as trivial cases. Consider again our model $\mathcal{M}$, and let

$$\|\ build_*\ \| = \{(a,\beta),(a,\gamma),(b,\alpha),(b,\beta),(a \uplus b,\beta),(a \uplus c,\gamma),(b \uplus c,\delta_*)\}$$

where all the pairs are understood to be "collective". In particular I draw attention to the presence of the pair $(a \uplus b, \beta)$ in $\|\ build_*\ \|$. Of course I have included it for illustrative purposes, but logically speaking it is a mere coincidence that this pair and the pairs (a,β) and (b,β) are all in $\|\ build_*\ \|$ simultaneously (cf. point 4 on p. 105, and also footnote 82).

To arrive at the desirable semantic value for ${}^*build_*$ (corresponding to the distributive readings), we must start from the "atomic pairs" in $\|\ build_*\ \|$.

Define

$$\| {}^{a}\mathit{build}_* \| =_{\text{def}} \| \mathit{build}_* \| \cap (A \times A).$$

That is, in $\mathcal{M}$ we have

$$\| {}^{a}\mathit{build}_* \| = \{(a,\beta),(a,\gamma),(b,\alpha),(b,\beta)\}$$

Then ${}^{*}\mathit{build}_*$ is defined as given by Definition 1, except that the first rule

If $(x,y) \in \| \delta_* \|$, *then* $(x,y) \in \| {}^{*}\delta_* \|$

is now (in the case of build_*) to be instantiated as

If $(x,y) \in \| {}^{a}\mathit{build}_* \|$, *then* $(x,y) \in \| {}^{*}\mathit{build}_* \|$

The extension of a basic two-place predicate which is unambiguously distributive, such as love_*, is always in $A \times A$. For that reason it is no problem to modify Definition 1 so as to comprise both types of verb. I state the revised definition of ${}^{*}\delta_*(x,y)$:

Definition 2 (Revised Semantic Definition of ${}^{*}\delta_*$)

1. If $(x,y) \in \| \delta_* \| \cap (A \times A)$, *then* $(x,y) \in \| {}^{*}\delta_* \|$
2.-4.: As Definition 1.

This definition establishes the following general relation between collective and distributive "variants" of a D-ambiguous verb:

Proposition 4 $\models \forall x, y \in A\,[\delta_*(x,y) \leftrightarrow {}^{*}\delta_*(x,y)]$
where δ is an extensional transitive verb subcategorised as both +distributive and -distributive[99].

[99] In this fragment, the restriction of proposition 4 to verbs which are actually D-ambiguous is not strictly necessary. For the purely collective verbs do not permit of any atoms in their extension within this fragment, and so there can be no danger that the proposition could inadvertently be used for those. When group terms are included, this may have to change, and if so the restriction here will indeed be required. That in turn may have consequences for my implementation in the following sections; for the counterpart of proposition 4, which will be axiom 18, cannot be restricted like proposition 4 without presupposing that an inference engine has access to the lexicon. In the current fragment axiom 18 works all right; but there may be a problem to be solved here, when the fragment is to be extended to include group terms (depending on which treatment those will ultimately receive).

From Definition 2 we get the following semantic value for ${}^{*}build_{*}$:

$$\| {}^{*}build_{*} \| = \| {}^{a}build_{*} \| \bigcup \{(a \oplus b, \beta), (a, \beta \oplus \gamma), (b, \alpha \oplus \beta)\}$$

Let me sum up the crucial points about this treatment of D-ambiguous verbs (in a somewhat loose order):

1. I have assumed that the collective reading should be considered to be the basic one. In my opinion this is a natural and well-founded choice: first, it is the reading which carries the least information in the sense that it warrants less inferences than the distributive reading; second, the distributive reading arises from the application of the two-place predicate modifier (*) and is therefore naturally regarded as derived from something else.

2. Having once decided to consider verbs like "build" to be ambiguous between a collective and a distributive reading, it should be made clear that there is a linguistic connection between those readings. The "ambiguity approach" leads to dual lexical entries for "build" (and similar verbs); but the various entries for any such verb do not represent semantically disjoint verbs, which simply happen to be homographs. Rather, the entries in question are semantically related. The construction given above for the variants of "build" imply such a semantic relation (a fact which is also expressed by proposition 4).

3. The collective and distributive readings always coincide w.r.t. their "atomic pairs". For instance, the coinciding readings in our present case of $build_{*}$ are

$$\| build_{*} \| \bigcap \| {}^{*}build_{*} \| = \| {}^{a}build_{*} \| \bigcup \{(a \uplus b, \beta)\}$$

   Generally speaking, it is always the case that

$$\| {}^{a}\delta_{*} \| \subseteq (\| \delta_{*} \| \bigcap \| {}^{*}\delta_{*} \|)$$

   These cases are at the same time the trivial cases of the collective and the distributive readings. In semantics as in mathematics there is no reason to be suspicious of trivial cases; and obviously the collective-distributive distinction becomes trivial exactly when we "come down" to speaking of atoms only.

4. The presence of the pair $\{(a \uplus b, \beta)\}$ in ($\| \mathit{build}_* \| \cap \| {}^*\mathit{build}_* \|$) illustrates a point which I have made earlier, namely the fact that even in the non-trivial cases the two readings are not mutually exclusive. Again, this is as it should be, at least from a purely model-theoretic point of view: the truth of "John built the college and Mary built the college" is quite independent from the truth of "John and Mary built the college together"[100].

10.3 The General Definition of "Closure"

During the previous sections I have a few times suggested that the semantic construction of the closures of one- and two-place predicates are really just instances of the same generalised definition of closure. By now, sufficient background has been created for actually stating this definition:

Definition 3 (General Closure)

For all n-place predicates δ_*, for all $z, y, x_1, \ldots, x_n \in E$:

1. *If* $(x_1, x_2, \ldots, x_n) \in \| \delta_* \| \cap A^n$,
   then $(x_1, x_2, \ldots, x_n) \in \| {}^*\delta_* \|$
   $(n \geq 1)$
2. *If* $(x_1, \ldots, x_m, y, x_{m+1}, \ldots, x_n) \in \| {}^*\delta_* \|$,
   and $(x_1, \ldots, x_m, z, x_{m+1}, \ldots, x_n) \in \| {}^*\delta_* \|$,
   then $(x_1, \ldots, x_m, y \oplus z, x_{m+1}, \ldots, x_n) \in \| {}^*\delta_* \|$
   $(n \geq m \geq 0)$
3. *Nothing else is in* $\| {}^*\delta_* \|$

With this definition we can now clearly see that closure of one-place predicates is really just a special case of closure of n-place predicates[101].

The discussion has of necessity been very technical so far. However, having once grasped the necessary technical notions it seems to me that the above definition is the obvious intuitive extension of Link's idea of closure.

[100] The reader may feel that a college—at least as a physical entity—can only be built once and that these examples are therefore somewhat odd. In my opinion such intuitions have to do with real-world knowledge rather than semantics, but if the example is felt to be unclear, substitute "built the college" with "lifted the stone". (Cf. further discussion of this issue in section 10.7).

[101] At least as long as we are considering predicates of the "sub-asterisk type" $\delta_*(x_1, x_2, \ldots, x_n)$.

At least as long as we are considering purely extensional verbs the definition can be used to construct in a quite simple manner those closures which represent all the possible distributive readings of such verbs. As I have argued earlier in the special case of definition 1, these closures reflect linguistic intuition on such verbs and comprise the set of true sentences with them (Cf. p. 97).

Strictly speaking, though, there is not just one closure-operator, but in fact one for each different *n*. But the semantic construction of each closure is an instance of one and the same "construction algorithm", i.e. the one given by definition 3; and they all comply with the general rule

$$^{*}\varsigma, \in \mathrm{ME}_{\langle\tau\rangle}, \quad \text{for}\, \tau \in B, \;\; \varsigma \in \mathrm{ME}_{\langle\tau\rangle}$$

where B is the set of types of n-place predicates of the form δ_*[102].

Unfortunately, the general definition of closure does *not* solve all our problems w.r.t. the notion of "distributivity", or, for that matter, the problems of ambiguity between collective and distributive readings. It only forms part of a solution, and in some cases we shall have to combine the use of closure with other means in order to achieve an adequate account. This will become evident in the next two sections.

10.4 A Meaning Postulate Regained

Since $^{*}\delta_*(x, y)$ is now a well-understood expression, we are in a better position to discuss meaning postulate 1. It will be recalled that its "application" gave rise to a scope-problem, when plurals were involved. Thus the sentence

(1) "some students see a college",

which translates into

(2) $\exists x[^{\bullet}\mathit{student}(x) \wedge {}^{*}\mathit{see}_*(x, \lambda Q \exists y[\mathit{college}(y) \wedge Q(y)])]$

would by straightforward application become equivalent to

(3) $\exists x[^{\bullet}\mathit{student}(x) \wedge \exists y[\mathit{college}(y) \wedge {}^{*}\mathit{see}_*(x, y)]]$

While (3) does represent a possible reading of (1) (loosely: "there is a college such that each of the students in question sees it"), it is certainly too

[102] I.e. $B = \{\langle e,t\rangle, \langle e,\langle e,t\rangle\rangle, \langle e,\langle e,\langle e,t\rangle\rangle\rangle, \ldots, \ldots\}$

restricted to provide a "general meaning" for that sentence. So let me try to state in ordinary English which meaning is required:

(4) "there are some students,
such that each of those students sees a college"

As I see it, (4) is a paraphrase of (1), and this paraphrase apparently leads to the logical form

$$(5)\ \exists x[{}^{\bullet}student(x) \wedge \forall z[z \,\Pi\, x \rightarrow \exists y[college(y) \wedge {}^{*}see_{*}(z,y)]]]$$

The logical form (5) seems to rid us of the scope-problem in (3): while (3) in effect gives the object phrase "a college" wide scope[103], it is given narrow scope in (5). And yet, the problem with (3) repeats itself, although for other reasons: (5) still implies that there is at least one college such that all the students in question see it. To see this, let

$$\| student \|_{\mathcal{M}} = \{a, b\}$$
$$\| college \|_{\mathcal{M}} = \{\alpha, \beta, \gamma, \delta\}$$
$$\| see_{*} \|_{\mathcal{M}} = \{(a,\alpha), (a,\beta), (b,\gamma), (b,\delta)\}$$

In this model there are two students each of whom sees two colleges, but no college is seen by both students. Then $\mathcal{M}$ intuitively satisfies (4) and (1). But if we instantiate $(a \oplus b)$ for x in (5) we get

$${}^{\bullet}student(a \oplus b) \wedge \forall z[z \,\Pi\, (a \oplus b) \rightarrow \exists y[college(y) \wedge {}^{*}see_{*}(z,y)]]$$

which implies (since $(a \oplus b) \,\Pi\, (a \oplus b)$):

$$\exists y[college(y) \wedge {}^{*}see_{*}(a \oplus b, y)]$$

or equivalently

$$\exists y[college(y) \wedge {}^{*}see_{*}(a,y) \wedge {}^{*}see_{*}(b,y)]$$

But in $\mathcal{M}$ that last formula is false.

Generally speaking, (5) implies that for every subgroup of "some students", there is a college such that the entire subgroup see it. The reason for this is that *the dissectiveness property of closures is too strong for our*

[103]Perhaps this can be seen more clearly in the logical form

$$\exists x[{}^{\bullet}student(x) \wedge \exists y[college(y) \wedge \forall z[z \,\Pi\, x \rightarrow {}^{*}see_{*}(z,y)]]]$$

which is equivalent to (3).

present case. As indicated by (4), we only want to state that for each *individual member* there is a college which he sees; but not, of course, that the same should necessarily be the case for each subgroup. So which notions and formal means are required here? As far as I can see, this is the exact point in the present framework where the notion of "atomicity" becomes indispensable[104] in order to account for "distributivity".

So let there be a predicate *atomic* such that

$$\| \, atomic(x) \, \| = \begin{cases} 1, & \text{for } \| \, x \, \| \in A \\ 0, & \text{otherwise} \end{cases}$$

Then we can finally state the desired reading by

(6) $\exists x[{}^{\bullet}student(x) \wedge \forall z[atomic(z) \wedge z \, \Pi \, x \rightarrow \exists y[college(y) \wedge {}^{*}see_{*}(z,y)]]]$

We can now state a meaning postulate akin to meaning postulate 1, but in a form which makes sense for translations of distributive plural verb phrases, too:

Meaning Postulate 2
$\forall x[{}^{*}\delta_{*}(\mathcal{P})(x) \leftrightarrow \forall z[atomic(z) \wedge z \, \Pi \, x \rightarrow \mathcal{P}(\lambda y \, {}^{*}\delta_{*}(z,y))]]$
where ${}^{*}\delta_{*}$ *translates any distributive and extensional transitive verb, and* $\mathcal{P}$ *ranges over terms.*

Perhaps attention should be drawn to the fact that the closure-operator in the lefthand side of the meaning postulate ranges over the entire expression $\delta_{*}(\mathcal{P})$, whereas the closure-operator of the righthand side ranges over δ_{*} only.

In fact, I am not going to use the results embodied in meaning postulate 2 in exactly that form; rather, I am going to provide the desired logical forms directly from the translation relation[105]. But since meaning postulate 1 was

[104]While it is true that I have relied on this notion in order to give the *semantic* definition of closure, "atomicity" has not been seen to be necessary for any previously encountered types of inference. In the absence of proofs of "soundness" and "completeness", I cannot make a definitive statement on the subject, but it does seem as if the notion could be dispensed with in axiomatising the inference system up to the present point.

[105]Strictly speaking, the previous proofs of natural language inferences in section 9.7 ought to be modified in accordance with the introduction of "atomicity" into the translation relation. However, with the exception of inferences (3) and (11), the structure of

a starting point for this part of my investigation, it seems to me appropriate that the result should be explicitly stated as a "revised" meaning postulate. The differences and similarities between the two meaning postulates should be obvious by now.

Finally, I draw attention to the following fact:

If we apply meaning postulate 2 to representation (2) above, the result will be representation (6). But then, if we wish our example "some students see a college" to be dealt with as ambiguous between the wide scope reading (3) and the narrow scope reading (6), the grammar and translation relation must be made to provide (3) as a separate logical form. So to cover both cases separately we should have

"some students see a college"

$\trianglerighteq_1$ $\exists x[{}^{\bullet}\boldsymbol{student}(x) \wedge \forall z[atomic(z) \wedge z \,\Pi\, x \rightarrow \exists y[college(y) \wedge {}^{*}see_{*}(z,y)]]]$

$\trianglerighteq_2$ $\exists x[{}^{\bullet}\boldsymbol{student}(x) \wedge \exists y[college(y) \wedge {}^{*}see_{*}(x,y)]]$

10.5 A Limitation of the Present Framework

Unfortunately, the introduction of the explicit universal quantifier together with the notion of "atomicity" limits the usefulness of closures. That is, although we have a general definition of closure, that does not by itself yield a general treatment of "distributivity". Furthermore, one might object that the introduction of the universal quantifier in meaning postulate 2 renders the closure of $\delta_*(y)(z)$ completely trivial w.r.t. its second-place argument z (which must in this context always be atomic).

However, there are other logical forms for which the full scale of closure properties are still required for the second-place argument, for instance in the wide-scope reading (3) of (1). Likewise the following sentences can be represented[106] by logical forms where the closure properties are non-trivial:

---

the proofs will essentially be unaltered, and as regards those two inferences you will see the kind of structure which should now be given to them in the inferences (4) and (9) in section 10.9. For these reasons I shall not use space and time here on recasting those previous proofs.

[106]For the sake of generality of the translation relation, I am going to subsume those logical forms under the pattern given by (6); but that is a matter of convenience and does not affect the point being made here: for some readings the full potential of closures *can* come into use in a natural manner.

(7) "John and Mary see Paul and George and Sue"

(8) $^{*}see_{*}(john \oplus mary, paul \oplus george \oplus sue)$

(9) "the students see the masters"

(10) $\exists x[^{\bullet}master(x) \wedge \forall y[^{*}master(y) \leftrightarrow y \,\Pi\, x] \wedge$
$\exists z[^{\bullet}student(z) \wedge \forall y[^{*}student(y) \leftrightarrow y \,\Pi\, z] \wedge$
$^{*}see_{*}(x, z)]]$

While "distributivity" as inherent in natural language is sometimes accounted for in terms of closure only, we may just have to acknowledge that there are other cases where it depends upon an interaction between quantifiers and closures[107]. There is nothing devious in this, although a completely uniform treatment would of course be preferable, if one were available.

The notion of closure does not seem to get us there within the context of the present framework, though; and that is one of the negative conclusions of this investigation. But then, that elusive goal of a completely general treatment may just be unattainable. At any rate, the present treatment of "distributivity" in terms of closure and sometimes quantifiers leads to an inferential system, which is more fully and rigorously defined than any other system known to me within this problem domain.

10.6 The Seven Readings

So far, I have deliberately restricted the discussion to a few paradigmatic cases involving the collective-distributive distinction. However, in principle a transitive verb may be D-ambiguous not only w.r.t. its subject phrase, but also w.r.t. its object phrase. When abstracting from idiosyncracies of individual verbs, a transitive verb is potentially four-ways ambiguous:

[107]This result must be seen as relative to an important decision at the outset of this investigation, viz the decision to concentrate on purely extensional predicates of the form δ_*. It may be that other directions of research would not lead to the (explicit) incorporation of the universal quantifier in the account of distributive verbs. But if that is so it remains to be shown for any significant fragment of natural language. It must also be emphasised that any such approach would apparently have to live with those superfluous sublimation-concepts which have here been eliminated by the use of the "sub-asterisk" type predicates of the form δ_*. Cf. p. 88 and also footnote 80.

- distributive subject — distributive object
- distributive subject — collective object
- collective subject — distributive object
- collective subject — collective object

This picture is further complicated by the fact that it sometimes also makes a difference whether the object phrase is taken to have wide scope over the subject phrase or not. Thus eight different readings could be expected in the worst cases, but in fact scope doesn't make any difference when both phrases are read "collectively", and so "only" seven different readings come into question[108].

One transitive verb which seems to be D-ambiguous w.r.t. both the subject and the object phrase is the verb "visit", and this gives rise to seven different readings for the sentence

"some students visit some masters".

I shall describe each of those readings by giving

- a brief verbal description
- a (minimal) model typical of the reading
- a logical form which represents the reading.

Before stating the individual readings it will be useful to explain expressions of the form

$$^{*}(\lambda u\ atomic(u) \wedge visit_{*}(u,\mu))(\nu)$$

Consider a tiny model $\mathcal{M}$ where

$$A = \{a, b, c, \alpha, \beta, \gamma\}$$
$$\parallel \mu \parallel = (\alpha \uplus \beta \uplus \gamma)$$

[108]The determiners in a natural language sentence may limit the number of possible readings. Thus "the students visit the masters" allows for no more than four different readings, since scope-ambiguity is irrelevant to the semantics of this sentence.

In the cases where all the seven readings seem possible it should be noted that some of them are rather stretched and might be difficult to recognise for the naïve language user. I do not wish to take any final stand on the question of whether some sentences are really seven-ways ambiguous in the manner discussed here; for the moment the only intention is to point out that this *may* be the case.

$$\| \nu \| = (a \uplus b)$$
$$\| \textit{visit}_* \| = \{(a, \alpha \uplus \beta \uplus \gamma), (b, \alpha \uplus \beta \uplus \gamma), (b \uplus c, \alpha \uplus \beta \uplus \gamma)\}$$

This corresponds to a "world" where a visited α, β and γ, who were together on the occasion (that is, they did not each receive a visit separately); and likewise b visited α, β and γ. In that situation a and b actually *each* visited the "collective entity" α, β and γ, and so we might wish to express this by means of a distributive predicate for a and b, and to be able to draw inferences accordingly[109].

Furthermore, in $\mathcal{M}$ b and c *together* visited α, β, and γ. Logically that event is quite independent from b's "separate visit" to α, β, and γ[110], and this fact must be respected by any attempt at a general formalisation (although meaning postulates may impose various constraints on the models in this respect). In other words, we must see to it that the representations of the various readings of "visit" are constructed so as to rule out any inference from $\textit{visit}_*(b \oplus c, \alpha \oplus \beta \oplus \gamma)$ to $\textit{visit}_*(b, \alpha \oplus \beta \oplus \gamma)$.

The question now is, can it be expressed with the aid of the closure-operator that "a and b [each] visited [the collective entity] α, β and γ"?

By means of lambda-abstraction we can construct new predicates which are distributive w.r.t. the argument bound by the lambda-operator, and collective w.r.t. its other argument. To get the right result we have to restrict the lambda-abstraction to range over atoms only. Then, the formula we need is

$$^*(\lambda u \; \textit{atomic}(u) \wedge \textit{visit}_*(u, \mu))(\nu)$$

To see this, consider the following values in $\mathcal{M}$:

$$\| (\lambda u \; \textit{atomic}(u) \wedge \textit{visit}_*(u, \mu)) \| = \{a, b\}$$
$$\| ^*(\lambda u \; \textit{atomic}(u) \wedge \textit{visit}_*(u, \mu)) \| = \{a, b, a \uplus b\}$$

---

[109] To say that a sentence of the form "a and b visited α, β and γ" may be read distributively for the subject phrase and collectively for the object phrase is the immediate linguistic counterpart of this model-theoretic line of reasoning.

[110] In a model with time-coordinates this would probably be more intuitively acceptable. However, as long as we insist that all events take place at exactly the same point of time (as we implicitly do here), one might ask how a and b can *independently* visit $(\alpha \oplus \beta \oplus \gamma)$ and at the same time visit them together. We have seen other closely related problems before, see for instance footnote 82. As pointed out earlier there is actually no problem here, at least not in purely semantic, i.e. model-theoretic terms. The wider intuitive problems will be addressed in section 10.10.

It is crucial to the expression $^*(\lambda u\ atomic(u) \wedge visit_*(u,\mu))$ that lambda-conversion cannot take place, in accordance with our previous results. The predicate is "irreducible", as it were—that is indeed the reason why it is possible to read $^*(\lambda u\ atomic(u) \wedge visit_*(u,\mu))(\nu)$ as collective w.r.t. μ and distributive w.r.t. ν. In this sense the restrictions on lambda-conversion within the present framework can be said to augment the expressive power of its logical language.

By now it should be easy to see that $\mathcal{M}$ satisfies the formula under consideration, i.e.

$$\models_{\mathcal{M}}\ ^*(\lambda u\ atomic(u) \wedge visit_*(u,\mu))(\nu)$$

The seven readings for the sentence "some students visit some masters" are:

1. distributive subject, distributive object, narrow scope:

description: For each student in question there is a group of masters such that he visits each master in that group individually.

For each pair of students the corresponding groups of masters may be disjoint, as in the model immediately below; but they may also coincide or even be identical in all cases as in the model under (2).

model:

$\|\ student\ \| = \{a, b\}$
$\|\ master\ \| = \{\alpha, \beta, \gamma, \delta\}$
$\|\ visit_*\ \| = \{(a,\alpha), (a,\beta), (b,\gamma), (b,\delta)\}$

logical form:

$\exists x[^\bullet student(x) \wedge \forall z[atomic(z) \wedge z\,\Pi\,x \rightarrow$
$\exists y[^\bullet master(y) \wedge {}^*visit_*(z,y)]]]$

2. distributive subject, distributive object, wide scope:

description: Like (1), except that there is only one group of masters involved; as noted earlier reading (2) entails reading (1), but not conversely.

model:

$\|\ student\ \| = \{a, b\}$
$\|\ master\ \| = \{\alpha, \beta\}$

$\| \textit{visit}_* \| = \{(a,\alpha),(a,\beta),(b,\alpha),(b,\beta)\}$

logical form:
$\exists x[{}^{\bullet}\textit{student}(x) \wedge \exists z[{}^{\bullet}\textit{master}(z) \wedge {}^{*}\textit{visit}_*(x,z)]]$

3. distributive subject, collective object, narrow scope:

description: For each student in question there is a group of masters such that he visits that group as a collective entity. As with (1), the groups of masters involved may be disjoint, but they may also coincide or even be identical in all cases as in the model under (4).

model:
$\| \textit{student} \| = \{a,b\}$
$\| \textit{master} \| = \{\alpha,\beta,\gamma,\delta\}$
$\| \textit{visit}_* \| = \{(a,\alpha \uplus \beta),(b,\gamma \uplus \delta)\}$

logical form:
$\exists x[{}^{\bullet}\textit{student}(x) \wedge \forall y[\textit{atomic}(y) \wedge y \,\Pi\, x \rightarrow$
$\exists z[{}^{\bullet}\textit{master}(z) \wedge {}^{*}(\lambda u\ \textit{atomic}(u) \wedge \textit{visit}_*(u,z))(y)]]]$

(An alternative to this representation would be the simpler logical form

$\exists x[{}^{\bullet}\textit{student}(x) \wedge \forall y[\textit{atomic}(y) \wedge y \,\Pi\, x \rightarrow \exists z[{}^{\bullet}\textit{master}(z) \wedge \textit{visit}_*(y,z)]]]$

The more complicated form has the advantage, however, that is is quite easy to show for instance

$$\vdash \ {}^{*}(\lambda u\ \textit{atomic}(u) \wedge \textit{visit}_*(u,\beta))(a)$$
$$\vdash \ {}^{*}(\lambda u\ \textit{atomic}(u) \wedge \textit{visit}_*(u,\beta))(b)$$
$$\vdash \ {}^{*}(\lambda u\ \textit{atomic}(u) \wedge \textit{visit}_*(u,\beta))(a \oplus b))$$

whereas no similar inference can be made from $\textit{visit}_*(\beta)(a)$ and $\textit{visit}_*(\beta)(b)$ (at least not without some additional information). As I have suggested on an earlier occasion I give priority to ease of inference over elegance of translations[111]. Similar remarks apply to the logical form of reading 5 below.)

[111] See p. 69; for some more general remarks on this policy, see p. 117.

4. distributive subject, collective object, wide scope:

description: There is exactly one group of masters such that each of the students in question visits that collective group.

model:

$\| \textit{student} \| = \{a, b\}$
$\| \textit{master} \| = \{\alpha, \beta\}$
$\| \textit{visit}_* \| = \{(a, \alpha \uplus \beta), (b, \alpha \uplus \beta)\}$

logical form:

$\exists x[{}^\bullet \textit{student}(x) \land \exists z[{}^\bullet \textit{master}(z) \land {}^*(\lambda u \ \textit{atomic}(u) \land \textit{visit}_*(u, z))(x)]]$

5. collective subject, distributive object, narrow scope:

description: For each of the masters in question there is a group of students, which (as a collective entity) visited that master. The groups of students involved may be completely disjoint, but they may also intersect or even all be identical.

model:

$\| \textit{student} \| = \{a, b, c, d\}$
$\| \textit{master} \| = \{\alpha, \beta\}$
$\| \textit{visit}_* \| = \{(a \uplus b, \alpha), (c \uplus d, \beta)\}$

logical form:

$\exists z[{}^\bullet \textit{master}(z) \land \forall y[\textit{atomic}(y) \land y \, \Pi \, z \rightarrow$
$\exists x[{}^\bullet \textit{student}(x) \land {}^*(\lambda u \ \textit{atomic}(u) \land \textit{visit}_*(x, u))(y)]]]$

6. collective subject, distributive object, wide scope:

description: This is the special case of (5), where there is exactly one group of students involved.

model:

$\| \textit{student} \| = \{a, b\}$
$\| \textit{master} \| = \{\alpha, \beta\}$
$\| \textit{visit}_* \| = \{(a \uplus b, \alpha), (a \uplus b, \beta)\}$

logical form:

$\exists z[{}^\bullet \textit{master}(z) \land \exists x[{}^\bullet \textit{student}(x) \land {}^*(\lambda u \ \textit{atomic}(u) \land \textit{visit}_*(x, u))(z)]]$

7. <u>collective subject, collective object:</u>

description: There is just one group of students and one group of masters involved, and all the students visited all the masters on one and the same occasion.

model:

$\| \textit{student} \| = \{a, b\}$
$\| \textit{master} \| = \{\alpha, \beta\}$
$\| \textit{visit}_* \| = \{(a \uplus b, \alpha \uplus \beta)\}$

logical form:

$\exists x[{}^{\bullet}\textit{student}(x) \wedge \exists x[{}^{\bullet}\textit{master}(z) \wedge \textit{visit}_*(x, z)]]$

There are two general linguistic comments to be made on the results here.

First, given that the verbs "see", "build", and "visit" systematically give rise to different patterns of ambiguity, there seems to be a real linguistic subclassification to be made for transitive verbs in this respect. That subclassification could hardly evolve from the notion of *mixed extension*, but it can be expressed within the present framework, as I have shown.

Second, the reader may have noticed the use of expressions such as "occasion" and "event" above. Actually, some of the nuances in the meanings discussed here may belong to the specific domain of event-semantics rather than to the general logic of individuals.

Be that as it may, though: in the following implementation of an inference system I am going to ignore any other readings than the most general distributive and the most general collective reading (cf. readings (1) and (7)). This is first and foremost to limit the task of implementation, but I may mention that the axiomatisation about to be presented with minor modifications carry over to all the readings discussed above. From the pattern given in (1) it is possible to demonstrate all the paradigmatic inferences w.r.t. "distributivity". It is the translation relation rather than the axiomatic system that would become more complicated by the incorporation of all the possible readings.

The strategy here, however, is to focus on the *concepts* of "distributivity" and "collectiveness" as inherent in natural language rather than trying to achieve linguistic breadth.

I also consider event-semantics to be beyond the scope of the present investigation and shall not discuss that subject further.

10.7 Consequences for the Translation Relation

The translation relation must now be revised so as to implement the results of the previous sections. At this point I had better state a principle, which I have actually been following all along, but which will become more evident with this translation relation than the previous ones. When devising a translation relation one is sometimes faced with a trade-off: either one can strive for optimal elegance and clarity of the translation relation at the cost of sometimes complicating proofs of natural language inferences, or one can give priority to the ease of proving inferences. The latter priority leads to a more complicated translation relation, for proofs are done on the basis of the logical forms provided by the relation, and so a maximal content of *explicit* logical information in these forms make proofs of inferences easier (whereas implicit information will itself have to be proved, of course). A clear example of giving priority to ease of inferences can be found in the rule T5 (*ii*) below, where the translation relation is complicated beyond what is strictly necessary in order to utilise the linguistic information in the input sentences to the full.

When faced with trade-offs of this kind, I think that in general we should prefer simplicity of proving inferences to simplicity of the translation relation. The task of proving inferences is more likely to suffer combinatorial explosion than is the task of parsing sentences (given that natural language sentences are in practice rather restricted in length).

For the full details of the translation relation I refer you to **Appendix C**, but there are some points which deserve a closer comment:

Atomicity:
The results w.r.t. *atomicity* have the consequence that we need to represent explicitly whether entities are atomic. In the present fragment all singular common nouns are count nouns and clearly denote atomic entities. So I shall let the translations of the determiners with which they combine reflect this; thus, for instance, the translation of the singular definite article the_{sing} is to be

$$\lambda P \lambda Q \exists x[\forall y[{}^{*}P(y) \leftrightarrow y = x] \wedge atomic(x)]$$

In the revised translation relation, a sentence like "the student dies" translates[112] into

$$\exists x[\forall y[{}^{*}\mathit{student}(y) \leftrightarrow y = x] \wedge {}^{*}\mathit{die}(x) \wedge \mathit{atomic}(x)]$$

The other determiners are dealt with in an analogous manner.

All those *names* which are considered to denote atomic individuals[113] are dealt with along the following lines: if n is a name, then

$$n \quad \triangleright \lambda P\,[P(n) \wedge \mathit{atomic}(n)]$$

So for instance "John dies" translates into

$${}^{*}\mathit{die}(\mathit{john}) \wedge \mathit{atomic}(\mathit{john})$$

Transitive Verbs:
Transitive verbs will be subcategorised as $\pm$*distributive*. The D-ambiguous ones will receive dual lexical entries, one entry having the "feature-value" *+distributive*, and the other having the value *-distributive*.

The rule for transitive verbs will implement the meaning postulates 1 and 2 directly into the translations—meaning postulate 1 for the collective readings, where it still obtains without problems, and meaning postulate 2 for the distributive readings of plural transitive verbs. Furthermore, all singular transitive verb phrases are dealt with by a simplified version of meaning postulate 2 (see below).

All D-ambiguous transitive verbs will have dual lexical entries; or to put it in the terms of a [PTQ]-style grammar, there will for each such verb be two homographs[114] belonging to two different categories.

As mentioned in section 10.6, I shall devise the translation relation to provide only the most general distributive and collective readings, i.e. those readings where *both* arguments of the verbs are read distributively or *both* are read collectively. While this is without doubt too linguistically restricted for a large number of verbs such as "visit", it still allows me to study all the relevant inference patterns at stake. It will not be difficult to refine the translation relation to provide all the possible readings, though, but it

[112]In section 9.5, I have discussed why a translation of this form is preferred to the simpler logical form $\exists x[\forall y[\mathit{student}(y) \leftrightarrow y = x] \wedge \mathit{die}(x)]$.

[113]And that happens to be all names in the present fragment.

[114]With a certain semantic relation between them as stated in proposition 4.

would require a further subcategorisation of the verbs than the one suggested above[115].

In this place I shall state the rules as a modification of the [PTQ]-rule T5 for transitive verbs. Let

A = IV/T be the category of **plural distributive transitive verbs**;

B = IV//T be the category of **singular transitive verbs**;

C = IV///T be the category of **collective transitive verbs**.

Then, the revised [PTQ]-rule is as follows:

T5 (*i*) If $\delta \in P_A, \beta \in P_T$, and β translates into $\mathcal{P}$, then $F_{5-\text{i}}(\delta, \beta)$ translates into $\lambda z[\forall x[atomic(x) \wedge x \,\Pi\, z \rightarrow \mathcal{P}(\lambda y\, {}^*\delta_*(x, y))]]$

T5 (*ii*) If $\delta \in P_B, \beta \in P_T$, and β translates into $\mathcal{P}$, then $F_{5-\text{ii}}(\delta, \beta)$ translates into $\lambda x\, \mathcal{P}(\lambda y\, {}^*\delta_*(x, y))$

T5 (*iii*) If $\delta \in P_C, \beta \in P_T$, and β translates into $\mathcal{P}$, then $F_{5-\text{iii}}(\delta, \beta)$ translates into $\lambda x\, \mathcal{P}(\lambda y\, \delta_*(x, y))$

We have already indirectly seen this translation rule at work in the discussion of meaning postulate 2; more examples of translations under this rule will follow below.

As for the rule T5 (*ii*), this could be subsumed under T5 (*i*). However, singular verb phrases will in the present fragment always have a subject anphrase denoting an atomic entity, and for this reason the "internal antecedent" in meaning postulate 2 will always be redundant. To see this, consider a translation of "John loves a student" along the lines of T5 (*i*); then

John loves a student

$$\triangleright\ \forall x[atomic(x) \wedge x \,\Pi\, john \rightarrow \exists y[{}^*student(y) \wedge atomic(y) \wedge {}^*love_*(x, y)]] \wedge atomic(john)$$

But since $atomic(john)$ is part of the premise and $john \,\Pi\, john$ by axiom 4, we have almost immediately that

$$john \,\Pi\, john \wedge atomic(john)$$

[115]For instance, it would be necessary to sub-categorise the verb "visit" into no less than four different "versions", according to the intended "distributivity-values" of its subject and object, respectively.

and hence we always have in this case (and similarly in similar cases) that

$$\exists y[{}^{*}\mathit{student}(y) \wedge \mathit{atomic}(y) \wedge {}^{*}\mathit{love}_{*}(\mathit{john}, y)] \wedge \mathit{atomic}(\mathit{john})$$

We can prevent an inference engine (and ourselves) from always having to take those few steps by simply implementing this result directly into the translation relation; this is what has been done with the rule T5 (*ii*).

The verb "be":
A sentence of the form "X are students" is ambiguous between a purely predicative reading and an "identity reading", the latter meaning approximately "X are identical with certain students". (Recall that "students" as a bare plural noun phrase is interpreted as "some students".) It is therefore necessary to distinguish between a *be-of-identity* and a *be-of-predication*. (In **Appendix C**, the former is taken care of by the rule VP $\Longrightarrow$ Vaux+CNP, the latter by the rule VP $\Longrightarrow$ Vaux+Nbar).

A sentence like "John and Mary are students" on the *be-of-identity*-reading translates into

$$\exists x[{}^{\bullet}\mathit{student}(x) \wedge x = (j \oplus m) \wedge \mathit{atomic}(j) \wedge \mathit{atomic}(m)]$$

and on the *be-of-predication*-reading into

$${}^{*}\mathit{student}(j \oplus m) \wedge \mathit{atomic}(j) \wedge \mathit{atomic}(m)$$

The only semantic difference is that the *be-of-identity*-reading carries the "plural presupposition" and the *be-of-predication*-reading does not. Obviously, the former immediately entails the latter.

Some Intuitive Problems with the Determiner "the":
For a starting-point, consider the following sentence and its distributive translation

some students build the college

$$\trianglerighteq\ \exists x[{}^{\bullet}\mathit{student}(x) \wedge \forall v[\mathit{atomic}(v) \wedge v\,\Pi\, x \rightarrow \exists z[\forall y[{}^{*}\mathit{college}(y) \leftrightarrow y = z] \wedge \mathit{atomic}(z) \wedge {}^{*}\mathit{build}_{*}(v, z)]]]$$

This translation gives the distributive reading of the sentence and implies that each and every student all by himself built that unique college in question. Then the sentence is inherently false[116] and perhaps we should rule

[116] At least as long as we read "the college" as denoting a physical entity. If we think

out translation no. 1. That is, it could be argued that this sentence is really unambiguous and should receive only one (non-distributive) translation. For comparison one could also point to the related sentence

each of the students built the chapel

which may seem to be semantically anomalous.

However, I believe that the anomaly of translation no. 1 has to do with real-world knowledge (and perhaps lexicography) rather than semantics in its strictest sense. For consider a sentence like

"some pupils handed in the exercise"

where it is the *distributive* reading which seems to be preferable. Since "handed in the exercise" is itself obviously a D-ambiguous verb phrase[117], it seems that the discussed judgments in these two cases have to do with our general knowledge with respect to "building colleges" and "handing in exercises".

The tentative conclusion is that the translation relation should indeed not try to decide which reading is preferable in these cases, but this should rather be left to any actual theory building on the framework here—that is, it should be up to such theories to specify meaning postulates on transitive verbs so as to rule out one reading or the other.

For the translation of a sentence like "the students see the masters" I

---

of "the college" as an institution, and assume that this institution has been housed in several different buildings during its existence, then we might argue that translation no. 1 *can* make sense—i.e. that it is at least *possibly* true. However, even if we accept this it does not solve our problem with the "physical reading". At any rate I shall disregard the suggested "institutional reading" here.

[117]It might be objected that "the" in this sentence has a different semantic function than it has in "some students build the college"—"the exercise" obviously does not refer to one single physical entity, but rather to a certain "type", of which the pupils hand in "tokens". I suppose that is true, but the point being made here is simply that the use of the determiner "the" does not by itself exclude a distributive reading.

draw attention to the following fact:

the students see the masters

$$\triangleright\ \exists x[{}^{\bullet}student(x) \wedge \forall y[{}^{*}student(y) \leftrightarrow y \,\Pi\, x] \wedge \forall z[atomic(z) \wedge z \,\Pi\, x \rightarrow \exists v[{}^{\bullet}master(v) \wedge \forall u[{}^{*}master(u) \leftrightarrow u \,\Pi\, v] \wedge {}^{*}see_{*}(z, v)]]]$$

which is equivalent to

$$\exists x[{}^{\bullet}student(x) \wedge \forall y[{}^{*}student(y) \leftrightarrow y \,\Pi\, x] \wedge \exists z[{}^{\bullet}master(z) \wedge \forall u[{}^{*}master(u) \leftrightarrow u \,\Pi\, z] \wedge {}^{*}see(x, z)]]$$

"each" and "together" as "disambiguators":
In the present framework, it is very easy to account for the non-ambiguity of the sentences "John and Mary each built a college" and "John and Mary built a college together". If these "floated quantifiers"—"each" and "together"—are equipped with suitable syntactic features, e.g. "each" with the feature *+distributive* and "together" with *-distributive*, then a parse for the first sentence will be forced to select the distributive variant of "build", and a parse for the second sentence will be forced to select the collective variant. No other parses will be possible.

I have mentioned previously that this circumstance in my opinion speaks for the treatment of D-ambiguous verbs as being indeed ambiguous rather than as having *mixed extension*. (Cf. section 10.2.4.)

Dropping sub-asterisks:
Finally it should be mentioned that in the translation relation of **Appendix C**, I have ventured to drop the sub-asterisks (as in $^{*}\delta$ as opposed to $^{*}\delta_{*}$) in the translation of transitive verbs—since *all* translations of transitive verbs are of the form $^{*}\delta_{*}$ or simply δ_{*} no ambiguity can arise. The dropping of the sub-asterisks is simply a matter of convenience in the context of the translation relation of **Appendix C**.

10.7.1 Some Translations of Sentences with a D-Ambiguous Verb Phrase

1. some student builds a college

   $\triangleright$ $\exists x[{}^{*}student(x) \wedge atomic(x) \wedge$
   $\exists y[{}^{*}college(y) \wedge atomic(y) \wedge {}^{*}build_{*}(x, y)]]$

2. the student builds a college

   $\triangleright$ $\exists x[\forall y[{}^{*}student(y) \leftrightarrow y = x] \wedge atomic(x) \wedge$
   $\exists z[{}^{*}college(z) \wedge atomic(z) \wedge {}^{*}build_{*}(x, z)]]$

3. some students build a college

   $\trianglerighteq_1$ $\exists x[{}^{\bullet}student(x) \wedge \forall y[atomic(y) \wedge y \, \Pi \, x \rightarrow$
   $\exists z[{}^{*}college(z) \wedge atomic(z) \wedge {}^{*}build_{*}(y, z)]]]$

   $\trianglerighteq_2$ $\exists x[{}^{\bullet}student(x) \wedge \exists y[{}^{*}college(y) \wedge atomic(y) \wedge build_{*}(x, y)]]$

4. students build a college

   $\trianglerighteq_1$ $\exists x[{}^{\bullet}student(x) \wedge \forall y[atomic(y) \wedge y \, \Pi \, x \rightarrow$
   $\exists z[{}^{*}college(z) \wedge atomic(z) \wedge {}^{*}build_{*}(y, z)]]]$

   $\trianglerighteq_2$ $\exists x[{}^{\bullet}student(x) \wedge \exists y[{}^{*}college(y) \wedge atomic(y) \wedge build_{*}(x, y)]]$

5. some students build the college

   $\trianglerighteq_1$ $\exists x[{}^{\bullet}student(x) \wedge \forall v[atomic(v) \wedge v \, \Pi \, x \rightarrow$
   $\exists z[\forall y[{}^{*}college(y) \leftrightarrow y = z] \wedge atomic(z) \wedge {}^{*}build_{*}(v, z)]]]$

   $\trianglerighteq_2$ $\exists x[{}^{\bullet}student(x) \wedge$
   $\exists z[\forall y[{}^{*}college(y) \leftrightarrow y = z] \wedge atomic(z) \wedge build_{*}(x, z)]]$

6. John and Mary build a college

   $\trianglerighteq_1$ $\forall y[atomic(y) \wedge y \, \Pi \, (j \oplus m) \rightarrow$
   $\exists z[{}^{*}college(z) \wedge atomic(z) \wedge {}^{*}build_{*}(y, z)]] \wedge$
   $atomic(j) \wedge atomic(m)$

   $\trianglerighteq_2$ $\exists x[{}^{*}college(x) \wedge atomic(x) \wedge build_{*}(j \oplus m, x)] \wedge$
   $atomic(j) \wedge atomic(m)$

7. some students and Mary build a college

$\trianglerighteq_1$ $\exists x[^{\bullet}student(x) \wedge \forall y[atomic(y) \wedge y \,\Pi\, (x \oplus m) \rightarrow$
$\exists z[^{*}college(z) \wedge atomic(z) \wedge {}^{*}build_{*}(y, z)]]] \wedge atomic(m)$

$\trianglerighteq_2$ $\exists x[^{\bullet}student(x) \wedge \exists y[^{*}college(y) \wedge atomic(y) \wedge build_{*}(x \oplus m, y)]] \wedge$
$atomic(m)$

8. some students and the masters build a college and the chapel

$\trianglerighteq_1$ $\exists x[^{\bullet}student(x) \wedge \exists z[^{\bullet}master(z) \wedge \forall y[^{*}master(y) \leftrightarrow y \,\Pi\, z] \wedge$
$\forall y_0[atomic(y_0) \wedge y_0 \,\Pi\, (x \oplus z) \rightarrow \exists u[^{*}college(u) \wedge atomic(u) \wedge$
$\exists v[\forall y_1[chapel(y_1) \leftrightarrow y_1 = v] \wedge atomic(v) \wedge {}^{*}build_{*}(y_0, u \oplus v)]]]]]$

$\trianglerighteq_2$ $\exists x[^{\bullet}student(x) \wedge \exists z[^{\bullet}master(z) \wedge \forall y[^{*}master(y) \leftrightarrow y \,\Pi\, z] \wedge$
$\exists u[^{*}college(u) \wedge atomic(u) \wedge \exists v[\forall y_0[^{*}chapel(y_0) \leftrightarrow y_0 = v] \wedge$
$atomic(v) \wedge build_{*}(x \oplus z, u \oplus v)]]]]]$

9. some student builds the colleges

$\triangleright$ $\exists x[^{*}student(x) \wedge atomic(x) \wedge$
$\exists z[^{\bullet}college(z) \wedge \forall y[^{*}college(y) \leftrightarrow y \,\Pi\, z] \wedge {}^{*}build_{*}(x, z)]]]$

10. John and Mary build some colleges

$\trianglerighteq_1$ $\forall y[atomic(y) \wedge y \,\Pi\, (j \oplus m) \rightarrow \exists z[^{\bullet}college(z) \wedge {}^{*}build_{*}(y, z)]] \wedge$
$atomic(j) \wedge atomic(m)$

$\trianglerighteq_2$ $\exists z[^{\bullet}college(z) \wedge build_{*}(m \oplus j, z)] \wedge$
$atomic(j) \wedge atomic(m)$

10.8 Axiomatisation for Bi-Transitive Verbs

Now I can finally state those minor amendments to the previous axiomatisation, which will yield a significant portion of those inferences for transitive verbs that are intuitively valid.

Please observe that although there seem to be separate axioms for the one-place predicates and the two-place predicates, respectively, these can be conflated into just one set of "generalised" axioms, along the lines of definition 3. It is simply a matter of convenience that I choose here to add a few axioms and theorems rather than rephrase all the previous ones.

The axioms 14 and 15 correspond to the previously stated propositions 1 and 2, which are in turn the obvious "dual counterpart" of theorem 2. The theorems 10 and 11 are the dual counterpart of axiom 9, so they express the property of dissectiveness for both arguments of a two-place predicate under the closure-operator.

The axioms 16, 17, and 18 all add something genuinely new to the entire framework, since they state those rules regarding *atomicity*, which are of importance for the inferences studied here. In particular I draw attention to axiom 18: this axiom corresponds to proposition 4, which stated the relationship between the "trivial cases" of collective and distributive readings (see also footnote 99).

Finally, theorem 12 is quite trivial, but helpful for some natural language inferences.

Axiom 14 $\forall x, y, z: {}^*\delta_*(x, y \oplus z) \leftrightarrow {}^*\delta_*(x, y) \wedge {}^*\delta_*(x, z)$

Axiom 15 $\forall x, y, z: {}^*\delta_*(x \oplus y, z) \leftrightarrow {}^*\delta_*(x, z) \wedge {}^*\delta_*(y, z)$

Theorem 10 $\forall x, y, z: {}^*\delta_*(x, z) \wedge y \,\Pi\, x \rightarrow {}^*\delta_*(y, z)$

Proof 10

$\vdash_1$	$y \,\Pi\, x$	premise 1
$\vdash_2$	$y \oplus x = x$	axiom 7+replacement
$\vdash_3$	${}^*\delta_*(x, z)$	premise 2
$\vdash_4$	${}^*\delta_*(y \oplus x, z)$	since $y \oplus x = x$
$\vdash_5$	${}^*\delta_*(x, z) \wedge {}^*\delta_*(y, z)$	axiom 15+replacement
$\vdash_6$	${}^*\delta_*(y, z)$	simplification
$\vdash_7$	${}^*\delta_*(x, z) \wedge y \,\Pi\, x$	conjunction of 1, 3
$\vdash_8$	${}^*\delta_*(x, z) \wedge y \,\Pi\, x \rightarrow {}^*\delta_*(y, z)$	deduction theorem

Theorem 11 $\forall x,y,z\colon {}^*\delta_*(x,z) \wedge y\,\Pi\,z \rightarrow {}^*\delta_*(x,y)$

Proof 11

$\vdash_1$	$y\,\Pi\,z$	premise 1
$\vdash_2$	$y \oplus z = z$	axiom 7+replacement
$\vdash_3$	${}^*\delta_*(x,z)$	premise 2
$\vdash_4$	${}^*\delta_*(x, y \oplus z)$	since $y \oplus z = z$
$\vdash_5$	${}^*\delta_*(x,z) \wedge {}^*\delta_*(x,y)$	axiom 14+replacement
$\vdash_6$	${}^*\delta_*(x,y)$	simplification
$\vdash_7$	${}^*\delta_*(x,z) \wedge y\,\Pi\,z$	conjunction of 1, 3
$\vdash_8$	${}^*\delta_*(x,z) \wedge y\,\Pi\,z \rightarrow {}^*\delta_*(x,y)$	deduction theorem

Axiom 16
$\forall x,y,z\colon atomic(x) \wedge atomic(y) \rightarrow$
$[atomic(z) \wedge z\,\Pi\,x \oplus y \leftrightarrow z = x \vee z = y]$

Axiom 17
$\forall z\colon {}^\bullet P(z) \rightarrow \exists x[atomic(x) \wedge x\,\Pi\,z \wedge \exists y[atomic(y) \wedge y\,\Pi\,z \wedge x \neq y]]$

Axiom 18 [118] $\forall x,y\colon atomic(x) \wedge atomic(y) \rightarrow [\delta_*(x,y) \leftrightarrow {}^*\delta_*(x,y)]$

Theorem 12 $\forall x,y,z\colon x\,\Pi\,y \rightarrow x\,\Pi\,y \oplus z$

Proof 12

$\vdash_1$	$x\,\Pi\,y$	premise 1
$\vdash_2$	$x \oplus y = y$	axiom 7+replacement
$\vdash_3$	$x \oplus y \oplus z = y \oplus z$	since $x \oplus y = y$
$\vdash_4$	$x\,\Pi\,y \oplus z$	axiom 7+replacement
$\vdash_5$	$x\,\Pi\,y \rightarrow x\,\Pi\,y \oplus z$	deduction theorem

[118] With the general definition of closure and proposition 4 it has been made clear that the relationship between trivial collective and trivial distributive readings is of a quite general nature—so perhaps the general version of this axiom should be stated. (This can also serve as an example of how to generalise the axioms and theorems of the entire system).
Axiom 18 Generalised
$\forall x_1, x_2, \ldots, x_n\ (n \geq 1)\colon$
$atomic(x_1) \wedge atomic(x_2) \wedge \ldots \wedge atomic(x_n) \rightarrow$
$[\delta_*(x_1, x_2, \ldots, x_n) \leftrightarrow {}^*\delta_*(x_1, x_2, \ldots, x_n)]$

10.9 Natural Language Inferences

The underlying inference system for the natural language inferences in this section has been explained rather thoroughly, and so it should not be necessary with individual comments for each inference and its proof. But it may be helpful to point out explicitly the following pattern among the inferences:

The inferences 1, 2, 3, and 8 are all examples of inferences w.r.t. the object phrase of the transitive verb in question. They therefore exhibit some proof structures which are genuinely different from anything previously encountered.

The inferences 4, 5, 6, 7, and 10 are examples of inferences w.r.t the subject phrase of the verb. In this they closely resemble previously encountered natural language inferences with (basic) intransitive verb phrases, but their proofs differ somewhat in structure from previous proofs because of the way the translation relation works for transitive verbs.

Inferences 9 and 11 are examples of an inference-pattern that involves both the subject and the object phrase at the same time. Inference 12 also involves both the subject and the object phrase, but apart from this it poses some particular problems, which I will discuss in an individual comment to the inference.

In inferences 2 and 12 the difference between *be-of-identity-* and *be-of-predication*-readings (cf. p. 120) are of potential importance. I shall address this issue in the individual comments to those inferences.

All the inferences studied here with the exception of 12 are valid formally as well as intuitively.

1. John sees some colleges, so John sees a college.

Proof 1

$\vdash_1$	$\exists x[{}^{\bullet}college(x) \wedge {}^{*}see_{*}(j,x)] \wedge atomic(j)$	premise 1
$\vdash_2$	${}^{\bullet}college(\mu) \wedge {}^{*}see_{*}(j,\mu)$	E.I., simplification
$\vdash_3$	${}^{\bullet}college(\mu) \rightarrow$ $\exists y[atomic(y) \wedge y \,\Pi\, \mu \wedge \exists z[atomic(z) \wedge z \,\Pi\, \mu \wedge y \neq z]]$	axiom 17
$\vdash_4$	$\exists y[atomic(y) \wedge y \,\Pi\, \mu]$	detachment + simplification
$\vdash_5$	$atomic(\nu) \wedge \nu \,\Pi\, \mu$	E.I.
$\vdash_6$	${}^{\bullet}college(\mu) \wedge \nu \,\Pi\, \mu$	conjunction of 2', 5'
$\vdash_7$	${}^{*}college(\nu)$	theorem 7, detachment
$\vdash_8$	${}^{*}see_{*}(j,\mu) \wedge \nu \,\Pi\, \mu$	conjunction of 2',5'
$\vdash_9$	${}^{*}see_{*}(j,\nu)$	theorem 11, detachment
$\vdash_{10}$	${}^{*}college(\nu) \wedge atomic(\nu) \wedge {}^{*}see_{*}(j,\nu) \wedge atomic(j)$	conjunction of 1',5',7,9
$\vdash_{11}$	$\exists y[{}^{*}college(y) \wedge atomic(y) \wedge {}^{*}see_{*}(j,y)] \wedge atomic(j)$	E.G.

2. John sees the colleges, Clare and Kings are colleges, so John sees Clare and Kings.

Proof 2

$\vdash_1$	$\exists x[{}^{\bullet}college(x) \wedge \forall y[{}^{*}college(y) \leftrightarrow y \,\Pi\, x] \wedge$ ${}^{*}see_{*}(j,x)] \wedge atomic(j)$	premise 1
$\vdash_2$	${}^{\bullet}college(\mu) \wedge \forall y[{}^{*}college(y) \leftrightarrow y \,\Pi\, \mu] \wedge {}^{*}see_{*}(j,\mu)$	E.I., simplification
$\vdash_3$	${}^{*}college(c \oplus k) \wedge atomic(c) \wedge atomic(k)$	premise 2
$\vdash_4$	${}^{*}college(c \oplus k)$	simplification
$\vdash_5$	${}^{*}college(c \oplus k) \leftrightarrow (c \oplus k) \,\Pi\, \mu$	U.I. in 2'
$\vdash_6$	$(c \oplus k) \,\Pi\, \mu$	replacement in 4
$\vdash_7$	$(c \oplus k) \,\Pi\, \mu \wedge {}^{*}see_{*}(j,\mu)$	conjunction of 2', 6
$\vdash_8$	${}^{*}see_{*}(j, c \oplus k)$	theorem 11 + detachment
$\vdash_9$	${}^{*}see_{*}(j, c \oplus k) \wedge atomic(j) \wedge atomic(c) \wedge atomic(k)$	conjunction of 1',3',8

Comment:
Strictly speaking two different proofs for this inference should be considered, since "Clare and Kings are colleges" is ambiguous between the *be-of-predication*-reading

$$^{*}college(c \oplus k) \wedge atomic(c) \wedge atomic(k)$$

and the *be-of-identity*-reading

$$\exists x[^{\bullet}college(x) \wedge x = (c \oplus k)] \wedge atomic(c) \wedge atomic(k)$$

But the latter immediately entails the former (cf. axiom 10), and given that this is so a proof of the inference on the *be-of-identity*-reading really boils down to the same structure as we have in the proof here.

3. John sees the colleges, Clare is a college, so John sees Clare.

Proof 3

$\vdash_1$	$\exists x[^{\bullet}college(x) \wedge \forall y[^{*}college(y) \leftrightarrow y \,\Pi\, x] \wedge$ $^{*}see_{*}(j,x)] \wedge atomic(j)$	premise 1
$\vdash_2$	$^{\bullet}college(\mu) \wedge \forall y[^{*}college(y) \leftrightarrow y \,\Pi\, \mu] \wedge {}^{*}see_{*}(j,\mu)$	E.I. in 1'
$\vdash_3$	$^{*}college(c) \wedge atomic(c)$	premise 2
$\vdash_4$	$^{*}college(c) \leftrightarrow c \,\Pi\, \mu$	U.I. in 2'
$\vdash_5$	$c \,\Pi\, \mu$	replacement in 3'
$\vdash_6$	$^{*}see_{*}(j,\mu) \wedge c \,\Pi\, \mu$	conjunction of 2',5
$\vdash_7$	$^{*}see_{*}(j,c)$	theorem 11, detachment
$\vdash_8$	$^{*}see_{*}(j,c) \wedge atomic(c) \wedge atomic(j)$	conjunction of 1',3',7

4. John and Mary see a college, so John sees a college and Mary sees a college.

Proof 4

$\vdash_1$	$\forall z[atomic(z) \wedge z\,\Pi\,(m \oplus j) \rightarrow$ $\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(z,y)]] \wedge$ $atomic(j) \wedge atomic(m)$	premise
$\vdash_2$	$j\,\Pi\,(j \oplus m)$	theorem 1
$\vdash_3$	$j\,\Pi\,(j \oplus m) \wedge atomic(j)$	conjunction of 1',2
$\vdash_4$	$j\,\Pi\,(j \oplus m) \wedge atomic(j) \rightarrow$ $\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(j,y)]$	U.I. in 1'
$\vdash_5$	$\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(j,y)]$	detachment
	$\vdots$	likewise for m
$\vdash_6$	$\exists x[{}^*college(x) \wedge atomic(x) \wedge {}^*see_*(m,x)]$	
$\vdash_7$	$\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(j,y)] \wedge$ $\exists x[{}^*college(x) \wedge atomic(x) \wedge {}^*see_*(m,x)] \wedge$ $atomic(j) \wedge atomic(m)$	conjunction of 1',5,6

5. John sees a college and Mary sees a college, so John and Mary see a college.

Proof 5

$\vdash_1$	$\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(j,y)] \wedge$ $\exists x[{}^*college(x) \wedge atomic(x) \wedge {}^*see_*(m,x)] \wedge$ $atomic(j) \wedge atomic(m)$	premise 1
$\vdash_2$	$\mu = j$	premise 2
$\vdash_3$	$\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(\mu,y)]$	replacing μ for j in 1'
$\vdash_4$	$\mu = j \rightarrow$ $\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(\mu,y)]$	deduction theorem
$\vdash_5$	$\mu = m$	premise 3
	$\vdots$	likewise for m
$\vdash_6$	$\mu = m \rightarrow$ $\exists x[{}^*college(x) \wedge atomic(x) \wedge {}^*see_*(\mu,x)]$	
$\vdash_7$	$\mu = j \vee \mu = m \rightarrow$ $\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(\mu,y)] \vee$ $\exists x[{}^*college(x) \wedge atomic(x) \wedge {}^*see_*(\mu,x)]$	constructive dilemma
$\vdash_8$	$\mu = j \vee \mu = m \rightarrow$ $\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(\mu,y)]$	using laws for '$\exists$', '$\vee$'
$\vdash_9$	$atomic(j) \wedge atomic(m)$	simplification of 1
$\vdash_{10}$	$\mu = j \vee \mu = m \leftrightarrow atomic(\mu) \wedge \mu\, \Pi\, (m \oplus j)$	axiom 16, detachment
$\vdash_{11}$	$atomic(\mu) \wedge \mu\, \Pi\, (m \oplus j) \rightarrow$ $\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(\mu,y)]$	replacement in 8
$\vdash_{12}$	$\forall z[atomic(z) \wedge z\, \Pi\, (m \oplus j) \rightarrow$ $\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(z,y)]]$	U.G.
$\vdash_{13}$	$\forall z[atomic(z) \wedge z\, \Pi\, (m \oplus j) \rightarrow$ $\exists y[{}^*college(y) \wedge atomic(y) \wedge {}^*see_*(z,y)]] \wedge$ $atomic(j) \wedge atomic(m)$	conjunction of 1', 12

6. John and Mary see some colleges, so John sees some colleges and Mary sees some colleges.

Sketch of Proof: Analogously to proof (4) above.

7. John sees some colleges and Mary sees some colleges, so John and Mary see some colleges.

Sketch of Proof: Analogously to proof (5) above.

8. John sees a college and some chapels, so John sees some chapels.

Proof 8

$\vdash_1$	$\exists x[{}^*college(x) \wedge atomic(x) \wedge$ $\exists y[{}^\bullet chapel(y) \wedge {}^*see_*(j, x \oplus y)]] \wedge atomic(j)$	premise
$\vdash_2$	${}^*college(\mu) \wedge {}^\bullet chapel(\nu) \wedge {}^*see_*(j, \mu \oplus \nu)$	E.I. in 1'
$\vdash_3$	${}^*see_*(j, \mu \oplus \nu) \leftrightarrow {}^*see_*(j, \mu) \wedge {}^*see_*(j, \nu)$	axiom 14
$\vdash_4$	${}^\bullet chapel(\nu) \wedge {}^*see_*(j, \nu)$	replacement in 2 + simplification
$\vdash_5$	$\exists x[{}^\bullet chapel(x) \wedge {}^*see_*(j, x)]$	E.G.
$\vdash_6$	$\exists x[{}^\bullet chapel(x) \wedge {}^*see_*(j, x)] \wedge atomic(j)$	conjunction of 1',5

9. John and some students see some colleges, so a student sees a college.

Proof 9

$\vdash_1$	$\exists x[{}^{\bullet}student(x) \wedge \forall z[atomic(z) \wedge z\,\Pi\,(x \oplus j) \rightarrow$ $\exists y[{}^{\bullet}college(y) \wedge {}^{*}see_{*}(z,y)]]] \wedge atomic(j)$	premise
$\vdash_2$	${}^{\bullet}student(\mu) \wedge \forall z[atomic(z) \wedge z\,\Pi\,(\mu \oplus j) \rightarrow$ $\exists y[{}^{\bullet}college(y) \wedge {}^{*}see_{*}(z,y)]]$	E.I. in 1'
$\vdash_3$	${}^{\bullet}student(\mu)$	simplification
$\vdash_4$	${}^{\bullet}student(\mu) \rightarrow$ $\exists u[atomic(u) \wedge u\,\Pi\,\mu \wedge \exists v[atomic(v) \wedge v\,\Pi\,\mu \wedge u \neq v]]$	axiom 17
$\vdash_5$	$\exists u[atomic(u) \wedge u\,\Pi\,\mu \wedge \exists v[atomic(v) \wedge v\,\Pi\,\mu \wedge u \neq v]]$	detachment
$\vdash_6$	$atomic(\nu) \wedge \nu\,\Pi\,\mu$	E.I. in 5'
$\vdash_7$	$\nu\,\Pi\,\mu$	simplification
$\vdash_8$	$\nu\,\Pi\,(\mu \oplus j)$	theorem 12 + detachment
$\vdash_9$	$atomic(\nu) \wedge \nu\,\Pi\,(\mu \oplus j)$	conjunction of 6', 8
$\vdash_{10}$	$atomic(\nu) \wedge \nu\,\Pi\,(\mu \oplus j) \rightarrow$ $\exists y[{}^{\bullet}college(y) \wedge {}^{*}see_{*}(\nu,y)]$	U.I. in 2'
$\vdash_{11}$	$\exists y[{}^{\bullet}college(y) \wedge {}^{*}see_{*}(\nu,y)]$	detachment
	$\vdots$	cf. proof 1
$\vdash_{12}$	$\exists y[{}^{*}college(y) \wedge atomic(y) \wedge {}^{*}see_{*}(\nu,y)]$	
$\vdash_{13}$	${}^{\bullet}student(\mu) \wedge \nu\,\Pi\,\mu$	conjunction of 3, 7
$\vdash_{14}$	${}^{*}student(\nu)$	theorem 7 + detachment
$\vdash_{15}$	${}^{*}student(\nu) \wedge atomic(\nu) \wedge$ $\exists y[{}^{*}college(y) \wedge atomic(y) \wedge {}^{*}see_{*}(\nu,y)]$	conjunction of 6', 12, 14
$\vdash_{16}$	$\exists x[{}^{*}student(x) \wedge atomic(x) \wedge$ $\exists y[{}^{*}college(y) \wedge atomic(y) \wedge {}^{*}see_{*}(x,y)]]$	E.G.

10. The students and the masters see a chapel, so some students see a chapel.

Proof 10

$\vdash_1$	$\exists u[{}^{\bullet}student(u) \wedge \forall y_0[{}^{*}student(y_0) \leftrightarrow y_0 \,\Pi\, u] \wedge \exists v[{}^{\bullet}master(v) \wedge \forall y_1[{}^{*}master(y_1) \leftrightarrow y_1 \,\Pi\, v] \wedge \forall z[atomic(z) \wedge z \,\Pi\, (u \oplus v) \rightarrow \exists y[{}^{*}chapel(y) \wedge atomic(y) \wedge {}^{*}see_{*}(z, y)]]]]$	premise 1
$\vdash_2$	${}^{\bullet}student(\mu) \wedge {}^{\bullet}master(\nu) \wedge \forall z[atomic(z) \wedge z \,\Pi\, (\mu \oplus \nu) \rightarrow \exists y[{}^{*}chapel(y) \wedge atomic(y) \wedge {}^{*}see_{*}(z, y)]]$	E.I. in 1'
$\vdash_3$	$atomic(\alpha) \wedge \alpha \,\Pi\, \mu$	premise 2
$\vdash_4$	$\alpha \,\Pi\, (\mu \oplus \nu)$	theorem 12 + detachment
$\vdash_5$	$atomic(\alpha) \wedge \alpha \,\Pi\, (\mu \oplus \nu)$	conjunction of 3',4
$\vdash_6$	$atomic(\alpha) \wedge \alpha \,\Pi\, (\mu \oplus \nu) \rightarrow \exists y[{}^{*}chapel(y) \wedge atomic(y) \wedge {}^{*}see_{*}(\alpha, y)]$	U.I. in 2'
$\vdash_7$	$\exists y[{}^{*}chapel(y) \wedge atomic(y) \wedge {}^{*}see_{*}(\alpha, y)]$	detachment
$\vdash_8$	$atomic(\alpha) \wedge \alpha \,\Pi\, \mu \rightarrow \exists y[{}^{*}chapel(y) \wedge atomic(y) \wedge {}^{*}see_{*}(\alpha, y)]$	deduction theorem
$\vdash_9$	$\forall z[atomic(z) \wedge z \,\Pi\, \mu \rightarrow \exists y[{}^{*}chapel(y) \wedge atomic(y) \wedge {}^{*}see_{*}(z, y)]]$	U.G.
$\vdash_{10}$	${}^{\bullet}student(\mu) \wedge \forall z[atomic(z) \wedge z \,\Pi\, \mu \rightarrow \exists y[{}^{*}chapel(y) \wedge atomic(y) \wedge {}^{*}see_{*}(z, y)]]$	conjunction of 2',9
$\vdash_{11}$	$\exists x[{}^{\bullet}student(x) \wedge \forall z[atomic(z) \wedge z \,\Pi\, x \rightarrow \exists y[{}^{*}chapel(y) \wedge atomic(y) \wedge {}^{*}see_{*}(z, y)]]]$	E.G.

11. The students see some colleges and a chapel, John is a student, so John sees a college.

Proof 11

$\vdash_1$	$\exists x[{}^\bullet student(x) \wedge \forall y[{}^* student(y) \leftrightarrow y \,\Pi\, x] \wedge$ $\forall z[atomic(z) \wedge z \,\Pi\, x \rightarrow$ $\exists u[{}^\bullet college(u) \wedge$ $\exists v[{}^* chapel(v) \wedge atomic(v) \wedge {}^* see_*(z, u \oplus v)]]]]$	premise 1
$\vdash_2$	${}^* student(j) \wedge atomic(j)$	premise 2
$\vdash_3$	${}^\bullet student(\mu) \wedge \forall y[{}^* student(y) \leftrightarrow y \,\Pi\, \mu] \wedge$ $\forall z[atomic(z) \wedge z \,\Pi\, \mu \rightarrow$ ${}^\bullet college(\alpha) \wedge$ ${}^* chapel(\beta) \wedge atomic(\beta) \wedge {}^* see_*(z, \alpha \oplus \beta)]$	E.I. in 1
$\vdash_4$	${}^* student(j)$	simplification of 2
$\vdash_5$	${}^* student(j) \leftrightarrow j \,\Pi\, \mu$	U.I. in 3'
$\vdash_6$	$j \,\Pi\, \mu$	replacement in 4
$\vdash_7$	$atomic(j) \wedge j \,\Pi\, \mu$	conjunction of 2',6
$\vdash_8$	$atomic(j) \wedge j \,\Pi\, \mu \rightarrow$ ${}^\bullet college(\alpha) \wedge {}^* chapel(\beta) \wedge atomic(\beta) \wedge$ ${}^* see_*(j, \alpha \oplus \beta)$	U.I. in 3'
$\vdash_9$	${}^\bullet college(\alpha) \wedge {}^* chapel(\beta) \wedge atomic(\beta) \wedge$ ${}^* see_*(j, \alpha \oplus \beta)$	detachment
$\vdash_{10}$	${}^\bullet college(\alpha) \wedge {}^* chapel(\beta) \wedge atomic(\beta) \wedge$ ${}^* see_*(j, \alpha) \wedge {}^* see_*(j, \beta)$	axiom 14 + replacement
$\vdash_{11}$	${}^* chapel(\beta) \wedge atomic(\beta) \wedge {}^* see_*(j, \beta) \wedge atomic(j)$	conjunction of 2',10'
$\vdash_{12}$	$\exists y[{}^* chapel(y) \wedge atomic(y) \wedge {}^* see_*(j, y)] \wedge atomic(j)$	E.G.

12. John sees Pembroke and Kings,
Pembroke and Kings are colleges,
Mary sees Queens and Clare,
Queens and Clare are colleges,
so John and Mary see some colleges.

Proof 12

$\vdash_1$	$^{*}see_{*}(j, p \oplus k) \wedge atomic(p) \wedge atomic(k) \wedge atomic(j)$	premise 1
$\vdash_2$	$^{*}college(p \oplus k) \wedge atomic(p) \wedge atomic(k)$	premise 2
$\vdash_3$	$p \neq k$	premise 3
$\vdash_4$	$^{*}college(p) \wedge {}^{*}college(k)$	theorem 2 + replacement in 2'
$\vdash_5$	$p \neq k \wedge {}^{*}college(p) \wedge {}^{*}college(k)$	conjunction of 3, 4
$\vdash_6$	$^{\bullet}college(p \oplus k)$	axiom 11 + detachment
$\vdash_7$	$^{\bullet}college(p \oplus k) \wedge {}^{*}see_{*}(j, p \oplus k)$	conjunction of 1',6
$\vdash_8$	$\exists y[^{\bullet}college(y) \wedge {}^{*}see_{*}(j, y)]$	E.G.
$\vdash_9$	$p \neq k \rightarrow \exists y[^{\bullet}college(y) \wedge {}^{*}see_{*}(j, y)]$	deduction theorem
$\vdash_{10}$	$^{*}see_{*}(m, q \oplus c) \wedge atomic(q) \wedge atomic(c) \wedge atomic(m)$	premise 4
	⋮	as for j
$\vdash_{11}$	$q \neq c \rightarrow \exists x[^{\bullet}college(x) \wedge {}^{*}see_{*}(m, x)]$	
$\vdash_{12}$	$p \neq k \wedge q \neq c \rightarrow$ $\exists y[^{\bullet}college(y) \wedge {}^{*}see_{*}(j, y)] \wedge$ $\exists x[^{\bullet}college(x) \wedge {}^{*}see_{*}(m, x)] \wedge$ $atomic(m) \wedge atomic(j)$	conjunction of 1',9,10',11
	⋮	cf. inference 7
	⋮	and proof 5
$\vdash_{13}$	$p \neq k \wedge q \neq c \rightarrow$ $[\forall z[atomic(z) \wedge z\,\Pi\,(j \oplus m) \rightarrow$ $\exists y[^{\bullet}college(y) \wedge {}^{*}see_{*}(z, y)]] \wedge$ $atomic(j) \wedge atomic(m)]$	

Comment:
This inference is valid ***provided that*** Pembroke and Kings are not identical, and likewise for Queens and Clare. It was a consequence of the discussion of basic intransitive verb phrases that the translation relation was altered such as to let intransitive verb phrases enter translations under the closure- rather than the plural operator. The ***predicative readings*** of the verb "be" come under this general rule, and therefore the predicate "are colleges" warrants no inference w.r.t the "semantic plurality" of its subject term.

However, on the *be-of-identity*-reading the sentence "Pembroke and Kings are colleges" translates into

$$\exists x[{}^{\bullet}college(x) \land x = (p \oplus k)] \land atomic(p) \land atomic(k)$$

—and analogously for "Queens and Clare are colleges", of course. I have chosen here to concentrate on the *be-of-predication*-readings[119], but obviously the inference will turn out as unproblematically valid on the *be-of-identity*-readings.

On the readings studied here it is necessary to add a proviso; the proviso in question occurs in the final step of the proof (Cf. also natural language inference 7 of section 9.7).

[119]I am inclined to see those as the intuitively most natural.

10.10 Some Further Logico-Linguistic Problems

I conclude the investigation into Link's theory by briefly describing some further logico-linguistic problems for this paradigm.

Of course, a number of such problems have been discovered during the investigation. I shall not here recapitulate those problems which have already been discussed at some length and for which solutions have been proposed. Rather, I shall address some of those whose implications have been left open, and also mention a few problems which future research within a Linkian framework will encounter.

Simultaneous independent events:

One problem which has occurred in various versions is what might be called the ***problem of simultaneous independent events***; we saw it for instance in connection with the model for the predicate $carry_*$ on p. 91. In that model it was true that John carried the piano all by himself, and yet it was also true ***at the same time*** that John and Mary carried the piano together.

Intuitively that seems unacceptable, but the lattice-structure permits it[120]. Perhaps it should be stressed in this place that such counter-intuitive examples do not constitute conclusive evidence against a theory—after all, [PTQ] also contains semantic constructions without intuitive counterparts in language[121]. The crucial question is whether the theory provides all (and in particular ***only***) intuitively valid inferences. In the current case there are some apparently valid natural language inferences which do not come out as valid in the theory: for instance

(I) "John and Mary carried the piano together (at some fixed time t) so John didn't carry the piano all by himself (at t)".

But is that inference really a case of a general logical pattern? I think we are faced with a difficult problem here, which leaves much room for future research. So the following remarks are merely tentaive.

I argued earlier that in the case of the sentence "some pupils handed in the exercise" (cf. p. 121), a distributive reading is quite possible. It may be that

[120]The reason why I chose such counter-intuitive models for $carry_*$ was that I wanted to be able to demonstrate certain logical points about the theory.

[121]For instance, some of the semantical objects in the model-theory of intensional logic correspond to syntactic categories which are empty within natural language. (Cf. [DWP79, page 181 ff.])

at the same time John and Mary each hand in their individual solution to the exercise, and also collectively hand in a solution which they have made together. Then "John and Mary handed in the exercise (at some fixed time t)" does not warrant any inference like (I).

If (I) is nonetheless still accepted as valid, then its validity must rest on lexicographic properties of "carry"—which in turn seem to have to do with real-world knowledge. But then it will not be up to the basic semantic theory of plurals to validate this kind of inferences. Rather, that will be a task for further theories building on it. Such theories can specify the desired restrictions by means of meaning postulates. For instance, the case of "carry" could be handled by restricting the set of admissible models to those fulfilling

$$carry_*(x \oplus y, z) \rightarrow \neg carry_*(x, z)$$

which would immediately render the problematic inference (I) valid.

Or perhaps this kind of problem should be dealt with at the "event-level" in a more refined semantics, for what is at stake here may have to do with the logic of events rather than the logic of individuals. I suggested the same thing in section 10.6; but as I also remarked there I consider event-semantics to be beyond the scope of the present investigation.

Whichever way it is one thing seems certain: the semi-lattice structure represents a high level of abstraction when viewed in the light of such cases. To account for those we shall apparently need a good deal of extra machinery.

Bare plural noun phrases:

It may be worthwhile to briefly re-assess Link's treatment of this important feature of the problem domain. It will be recalled that a bare plural noun phrase, e.g. "students", is always treated as equivalent to "some students"[122].

As I remarked earlier (cf. p. 68 and footnote 72) that decision is not uncontroversial, for it can been argued that such phrases are really ambiguous

[122] Although in my implementation "children" differs from "some children" by having a separate predicative reading. For instance, "boys are children" has not only an *identity*-reading equivalent to "boys are some children", but also a predicative reading. Nonetheless the following remarks also hold for my implementation.

in a manner indicated by

$$\left\{\begin{array}{l} \text{all} \\ \text{most} \\ \text{some} \\ \vdots \end{array}\right\} \text{children}$$

However, it seems that a bare plural noun phrase like "children" also carries a plural presupposition that at least *some children* are involved; it is no coincidence that the determiner "no" does not occur in the list of possible readings.

So, "children die" may be read as e.g. "all children die" *only with the added proviso that there* are *some children who die.*

With this addition the entire picture changes, for now the reading "some children" become the common denominator of all the possible readings. To put it more technically, any sentence ϕ in which the bare noun phrase "children" occurs will entail a sentence ϕ', where ϕ' the result of substituting "children" with "some children" in ϕ. Thus "children die" entails "some children die" on whichever reading of the former is considered.

If we accept this, the decision to treat "children" and similar cases as equivalent to "some children" is a quite shrewd move, for then it gives the "minimal semantics" of bare plural noun phrases: sentences with such phrases will then entail all and only conclusions common to all the possible readings of them. In my opinion this is actually a desirable logical generalisation w.r.t. the inference patterns of natural language; but it is worth noting that it carries with it a good deal of regimentation, since it does not directly account for a number of linguistically quite possible readings of the sentences in question.

Determiners and features:

One important difference between Link's framework and my implementation is that I equip determiners with features whereas Link doesn't. This is not just a minor technicality, for Link several times stresses that the uniform translations of determiners is a principal merit of his treatment. For instance,

> "The quantifiers *some* and *the*, with one and the same translation, apply to both singular and plural cases ... it is only the incoming CN phrase which differentiates between the appropriate

singular and plural readings ... I think this is a nice instance of strict compositionality in times in which this principle has come under heavy fire in view of all sorts of recalcitrant data ..."
[Lin83, page 318]

So, Link's treatment relies heavily on the feature that the incoming plural CN phrases are under the proper plural-operator. While this works for some cases, as we have seen, it is hard to see how it could work for a determiner like "no" (and other determiners like "at most n", "no more than n" etc., where n is a numeral). Whatever the translation of "no" is to be it *must* combine with a predicate under the proper plural-operator in Link's translation relation, when the corresponding noun is in the plural; it is by this general treatment of plural noun phrases that he achieves the uniform translations of the determiners. But it is hard to see what translation of "no" could give the right semantics, then. And the treatment of nouns in the plural also renders his translation of "all" inadequate, "all children" coming out as

$$\lambda Q\, \forall x[^{\bullet} \mathit{child}(x) \rightarrow Q(x)]$$

which is not equivalent to "every child"[123].

So I think that even within Link's own semantic framework we shall have to abandon uniform translations of the quantifiers; and if I take the liberty of also considering the theory of Pelletier & Schubert in this place, that will provide further evidence of the need for the determiners themselves being differentiated by the noun with which they combine—mass or count, plural or singular[124].

In my opinion all this strongly implies that sooner or later we shall need several translations for most determiners, when mass terms and plurals are incorporated into larger fragments of natural language. The way I solved this problem was by giving various features of the lexical entries of the determiners, features whose values could then direct their translations.

[123]The difference can be illustrated by considering a model in which there is exactly one child, and that child doesn't die. Then "all children die" will be true in that model whereas "every child dies" will be false in it.

[124]Recall that in order to get the required inferences right it was necessary to give at least two translations for each determiner (according to whether their argument noun was count or mass).

Meta- and semi-distributivity:

It seems that there are kinds of distributivity which are slightly different from those discussed so far. Consider a verb like "gather"; this verb seems to be distributive down to but excluding the atomic level. For instance, "John, Mary, and George gather" implies "John and Mary gather", but not *"John gathers", of course. We might call this sort of distributivity "semi-distributivity".

It is interesting to contrast this example with "convene". That verb does not seem to be semi-distributive like "gather"—e.g. "John, Mary, and George convene" does not entail "John and Mary convene". The reason for this I take to be that "convene" has a sort of "institutional presupposition" about it: to convene implies to form a convention (however informal), and that convention is made up by *all* those convening, so one cannot "split it up" again. Observe how far these considerations seem to go in the direction of lexicography and pragmatics.

I suspect that most collective basic verbs are really semi-distributive. If that is so the theory of plurals has to be revised rather dramatically—the basic semi-lattice notions are still of use but it will be necessary to introduce a good bit of new features to get semi-distributivity into the picture.

On the other hand, consider an apparently strictly collective verb—take "convene" again. When this verb combines with group terms[125] in the plural as e.g. "the committees convene [in the college]", it suddenly ceases to be strictly collective. We may read the sentence in the usual collective way—*all the committees came together in the college to form a joint convention*—or we may read it distributively, meaning that *each committee convened by itself* (somewhere in the college, probably in different rooms). We may call this type of distributivity "meta-distributivity".

I admit that these observations are as yet brief and limited, but they do seem to indicate that the borderline between collective and distributive readings is not always so clear-cut as it appears to be in Link's theory. In a fuller account of collective and distributive readings we shall apparently get ever more involved in lexicography and real-world knowledge.

[125] Although group terms haven't been dealt with so far, I think it must be all right to point out one of the difficulties they will pose for future extensions of Link's theory. See also footnote 99.

11 Conclusions on the Theory of Plural

11.1 Link 86 on Distributive Transitive Verbs

In the introduction I mentioned that a later paper by G. Link contains a suggestion of how to deal with distributive transitive verbs in the plural. I should like to address this issue briefly before going on to the wider conclusions.

However, before I actually start quoting Link I have to explain a few notational items from his papers, which we haven't met before.

The symbol ":=" indicates a meta-linguistic abbreviation; thus $\epsilon_2 := \epsilon_1$ means that ϵ_2 is to be a meta-linguistic abbreviation for ϵ_1.

The symbol ${}^{\bullet}\Pi$ designates the relation *is-an-atomic-part-of*[126]; so $x\,{}^{\bullet}\Pi\, y$ in Link's notation is equivalent to $atomic(x) \wedge x\,\Pi\, y$ in my notation.

Finally, "LP+GQ" refers to the logic for plurals, which we already know from [Lin83], enhanced with generalised quantifiers.

I can now go on to give you the central quotation from Link's paper on this subject[127]. Link takes his clue from discussing *floated quantifiers* such as "each" and "together", and goes on to say:

> "Take the quantifier *all* hooked up on a verb phrase, VP. Then it acts like a distributivity operator on VP, which might be written ^{D}VP ...
>
> (48) ${}^{D}\mathrm{VP} := \lambda x \forall y[y\,{}^{\bullet}\Pi\, x \rightarrow \mathrm{VP}(y)]$
>
> The D-operator was designed to take care of the distributive reading of *lifting* sentences like (34a) [i.e. "Three men lifted the piano."] Note that it operates on possibly complex verb phrases; **this expressive power is needed in order to get scope relations right.** As an illustration of this, consider (49),
>
> (49) Three men lifted a piano.
>
> Let $\bigcup$ be the Montague-style tranlation relation between natural language and the second order logical representation for LP+GQ

[126] There is no connection between Link's use of (•) and my own use of the same symbol.

[127] Essentially that quotation gives you all he has to say on the subject; he does not give any more formal depth to it, although he does go on to discuss some more complicated examples with distributive transitive verbs and thus adds some linguistic breadth.

...For the distributive reading of (49) we then get the following translation ...

$$\exists x[(3men)'(x) \wedge \forall y[y \,{}^{\bullet}\Pi\, x \rightarrow \exists z[piano'(z) \wedge lifted'(y,z)]]]$$

[Lin86, pages 18–19]

First of all we may observe that the predicate *lifted* here really is to be $lifted_*$ in order for the translation to be well-formed. Furthermore, since the structure of the translation obviously carries over to cases like "some men [each] lifted a piano", we may ignore the numeral "3". And we may also ignore the tense of the verb here—it is sufficiently clear that the sentence

"some men [each] lift a piano"

is to be dealt with in the same manner.

Then the only difference between a Linkian treatment of "some men [each] lift a piano" and my own treatment of that sentence is that I apply the closure-operator directly to the predicate $lift_*$, my own translation coming out as

$$\exists x[{}^{\bullet}man(x) \wedge \forall y[atomic(y) \wedge y\,\Pi\, x \rightarrow$$
$$\exists z[piano(z) \wedge atomic(z) \wedge {}^{*}lift_*(x,z)]]]$$

Of course, the application of (*) to $lift_*$ presupposes a definition of the kind I worked out in the sections on distributive transitive verb phrases. As you see both Link and I introduce an explicit universal quantifier to account for this kind of cases. However, my use of the generalised closure-oparator allows the sentence here to enter into various inference patterns[128] which Link's translation does not by itself provide for.

So I conclude that my extension of Link's fragment to transitive verbs is actually quite in the spirit of his proposal here, but that it does make better use of the "semi-lattice properties" than his own suggestion[129].

[128] As given by for instance

$\models {}^{*}lift_*(a, \alpha \oplus \beta) \rightarrow {}^{*}lift_*(a,\alpha) \wedge {}^{*}lift_*(a,\beta)$

or $\models {}^{*}lift_*(a \oplus b, \alpha) \rightarrow {}^{*}lift_*(a,\alpha) \wedge {}^{*}lift_*(b,\alpha)$

[129] Actually Link's proposal may also make one wonder whether he intends to limit the use of the closure- and plural-operators to one-place predicates. This would apparently mean that they would be limited to translations of basic intransitive verbs, whereas more complicated verb phrases are to be dealt with along the lines of his suggestion here. If that is the intention it would pose a severe limitation on the usefulness of those notions in the description of natural language—and a quite unnecessary one, as the generalised definition 3 of closure has demonstrated.

11.2 Link on Methodology

Like Pelletier & Schubert, Link is commendably clear on the methodological views underlying his approach. Of course, his paper is not devoted to methodology as such, but rather to a specific logico-linguistic problem. For that reason we cannot expect to find any single sweeping statement of his methodological views; as was the case with Pelletier & Schubert we rather have to infer those from scattered remarks throughout the paper. The remarks that we do find are clear enough, however, and place Link firmly within that conception of formal semantics which I earlier called the *empiricist conception*.

In fact Link explicitly takes issue with nominalism and "reductionist ontological considerations", at least when they are employed in the semantic analysis of language. He first introduces the idea of using a semi-lattice structure in the analysis of plurals (and mass terms) with the following remark[130]:

> "In the case of group and mass objects ... [the general picture] naturally leads to the notion of a lattice structure ... However, its possible use in the present context has perhaps been obscured by reductionist ontological considerations which are, in my opinion, quite alien to the purpose of logically analyzing the inference structures of natural language. **Our guide in ontological matters has to be language itself, it seems to me.**"
>
> [Lin83, pages 303–304]

This is a most clear statement of the view that formal semantics should be concerned with the metaphysics (ontology) of language "as it is"—rather than be trying to impose a philosophically motivated semantic structure on language[131].

And Link further makes it clear that such a view has practical consequences for his approach to the linguistic problem at stake. I think it is

[130]Perhaps it should be pointed out that "group objects" in this quotation is intended to refer to *plural entities* of the form $x_1 \oplus x_2 \oplus \ldots \oplus x_n$. (Recall that Link does not deal with group terms such as the "committee").

[131]In fact the statement may even be taken to imply that "real" ontology should take its clue from the semantic analysis of language. If that is so the statement is in line with Pelletier & Schubert's suggestion that philosophical distinctions should be "generated" from linguistic analysis (or at least some such distinctions).

worthwhile quoting him at some length on this subject; but first I have to say that the notation $a + b$ is used in Link's description of mass nouns and therefore has not been discussed in the previous text. But the notation should be self-explanatory in the context of the quotation, in which it occurs here:

> "...let a and b denote two atoms in A. Then there are two more individuals to be called ... $a + b$ and $a \oplus b$. $a + b$ is still a singular object in A, the *material fusion of a and b*; $a \oplus b$ is the *individual sum* or *plural object* of a and b. The theory is such that $a + b$ constitutes, but is not identical with, $a \oplus b$. This looks like a wild platonistic caprice strongly calling for Occam's Razor. **Language, however, seems to function that way.** Take for a, b two rings recently made out of some old Egyptian gold. Then *the rings*, $a \oplus b$, are new, *the stuff*, $a + b$, is old."
>
> [Lin83, page 307]

Again the tendency is clear: the semantics that we give for language should be motivated by *the way language actually functions*, even if that leads to incorporating notions which seem otherwise highly suspicious. Of course, Link doesn't say that the suggested bit of semantics actually *is* a "wild platonistic caprice"—he only says that it looks like that. The crucial point here is that even if it were such a caprice we should accept it as part of the semantics of language, because language functions that way. But as I have already suggested in footnote 131 Link's views seem to go even further than that, as the following remark from his 86-paper perhaps also implies:

> "For those who like me think that language is the primary source for forming intuitions about our ontology there is a clear reason why the nominalists' project failed: According to the way language works the world is simply not a huge heap of particles of just one sort. To put it more formally: **there is more structure in the universe than just one part-whole relation of the kind the nominalists used.**"
>
> [Lin86, page 1]

The ultimate conclusion of such views might be that it is the analysis of language which leads to "the real ontology"—rather than the other way round. Link stops short of actually drawing that conclusion, though, and only suggests that it is the analysis of language which leads to a proper understanding of *our intuitions* about ontology.

What is perfectly clear, however, is the view that in the semantic analysis of language priority should be given to language as an empirical phenomenon rather than to philosophical considerations. I shall not try to evaluate in general the impact of such views upon the way formal semantics is conducted. In the conclusion below I shall, however, attempt a cautious assessment of the extent to which Link's theory actually lives up to his methodological views.

11.3 Conclusion

The theories of plurals within formal semantics are still in their infancy—even more so than the theories of mass terms. On that background Link's theory of plurals is a quite remarkable achievement, for it permits us to state the salient problems and concepts within this domain with more clarity and rigour than before. It is certainly true that it has been necessary to spell out many details and make non-trivial decisions in the implementation of the theory, but still the progress has essentially been based on Link's ideas.

However, the achievements of Link's theory do not necessarily imply that his approach is basically the right way to go. One question which arises almost immediately is whether it would not be more natural to rephrase the whole theory in terms of set-theory rather than lattice-theory; but I shall not consider that question further here, I merely draw attention to its existence. More fundamental questions are likely to arise w.r.t. the logico-linguistic soundness of the basic notions of the theory, as the list of outstanding issues in section 10.10 may indicate.

In my opinion the main shortcoming of the theory is its sheer complexity. While it works well and convincing for a set of paradigmatic core cases, it becomes ever more difficult to handle as we get involved with more complicated phenomena within the problem domain. Consider the complexity that would arise within this framework if we were to give a linguistically full account incorporating semi- and meta-distributivity, all the seven readings of four-ways D-ambiguous verbs, all the possible readings of bare plural noun phrases, etc.; the resulting machinery would obviously become very heavy-going.

In fairness it must be added, however, that I have not been able to find any single case for which such machinery couldn't conceiveably work. To that extent the theory must be said to be flexible and to possess considerable expressive power. But on the other hand it is not unfair to say that its complexity when it comes to handling more difficult cases gives cause for some worry w.r.t. the basic linguistic soundness of the entire appoach.

Another source of potential worry is the extent to which the borderline between semantics and lexicography/pragmatics becomes blurred in this framework. It is true that that borderline is always a hard one to draw; and in particular we must be aware of the fact that D-ambiguity of sentences arise out of the properties of individual words—namely their verbs—rather than from their structural properties. So in a way it is not surprising that the borderline may be especially hard to draw when dealing with the semantics of plurals; indeed, perhaps it *cannot* be done properly. If that is so it would add a piece of new evidence to the theoretical discussion of the relationship between semantics, lexicography, and pragmatics within linguistics. But it remains to be seen whether future research will be able to draw that borderline more clearly within the problem domain, and in the meantime we should simply be aware of this problem with Link's theory.

I think that the conclusion we must draw from the investigation into Link's theory is that it actually works better as an explication of theoretical concepts than as an account of linguistic data. Its real achievement is the considerable degree of clarity and rigour which it adds to our understanding of the distinction between collective and distributive readings.

But then it seems that we have here another case of discrepancy between the methodology professed and the method actually used. The semi-lattice structure has apparently been imposed upon the problem domain because of its power in explaining the basic notions rather than because of any demonstrated ability of accounting for a wide range of linguistic data. So it may do the theory more justice to view it as essentially a *theory of reasoning about collective and distributive readings*—based on linguistic observation, but not attempting to give any full account of the empirical data on such readings.

As with the theory of Pelletier & Schubert I find that this is actually a sensible approach to a substantial and little explored problem domain, but like them it fails to implement the stated *empiricist conception*. This may suggest that the *empiricist conception* of the project of formal semantics isn't completely adequate, or, for that matter, always helpful in putting that programme into practice; but here I shall not elaborate further on this point, important and interesting though it is. However, I think it is fair to say that where abstraction and regimentation are going to be used extensively it is better to realise that beforehand, for such realisation will enhance one's chances of getting one's priorities right.

In spite of all difficulties with Link's theory it is an important step on the way to a solution to the semantic description of plurals; for the clarity which it gives to the discussion of that problem domain ensures that it will be an important condition of future research, even if that will ultimately opt for a quite different approach to plurals.

Bibliography

[Bea79] G. Bealer. Predication and matter. In F. J. Pelletier, editor, *Mass Terms*, pages 279–294, D. Reidel, Dordrecht, 1979.

[Ben79] Michael Bennett. Mass terms and mass nouns in Montague Grammar. In Steven Davis and Marianne Mithun, editors, *Philosophy, Linguistics, and Montague Grammar*, pages 263–285, University of Texas Press, Austin, 1979.

[Bun85] Harry Bunt. *Mass Terms and Model-Theoretic Semantics.* Cambridge University Press, Cambridge, 1985.

[Cat86] Virginia Cathercole. Evaluating competing linguistic theories with child language data: the case of the mass-count distinction. *Linguistics and Philosophy*, 9(2):263–285, 1986.

[Cop78] Irving M. Copi. *Introduction to Logic.* MacMillan, New York, 1978.

[DWP79] David R. Dowty, Robert E. Wall, and Stanley Peters. *Introduction to Montague Semantics.* D. Reidel, Dordrecht, 1979.

[Gaz80] Gerald Gazdar. A cross-categorial semantics for coordination. *Linguistics and Philosophy*, 3(3):407–409, 1980.

[Gor85] Mike Gordon. *HOL: A Machine Oriented Formulation of Higher Order Logic.* Technical Report 68, University of Cambridge Computer Laboratory, 1985. Revised Version.

[Gor87] Mike Gordon. HOL: a proof generating system for higher order logic. 1987. Forthcoming in G. Birtwistle and P. A. Subrahmanyam, editors, *VLSI Specification, Verification, and Synthesis*, Kluwer.

[Gup80] Anil Gupta. *The Logic of Common Nouns.* Yale University Press, New Haven, 1980.

[Lan87] Fred Landman. Groups. 1987. Privately circulated paper.

[Lev78] Judith N. Levi. *The Syntax and Semantics of Complex Nominals.* Academic Press, London, 1978.

[Lin83] Godehard Link. The logical analysis of plurals and mass terms: a lattice-theoretical approach. In Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, editors, *Meaning, Use, and Interpretation of Language*, pages 302–323, de Gruyter, Berlin, 1983.

[Lin86] Godehard Link. *The Logic of Plurals, LP: Review of the Basic Ideas.* Technical Report, CSLI, Leland Stanford Junior University, 1986. Now published in P. Gärdenfors: *Generalized Quantifiers: Studies in Linguistics and Philosophy*, D. Reidel, Dordrecht.

[PTQ] Richard Montague. The proper treatment of quantification in ordinary English. In Richmond H. Thomason, editor, *Formal Philosophy*, pages 247–270, Yale University Press, 1976.

[Mon79] Richard Montague. The proper treatment of mass terms in English. In F. J. Pelletier, editor, *Mass Terms*, pages 173–178, D. Reidel, Dordrecht, 1979.

[Par79a] Terence Parsons. Afterthoughts on mass terms. In F. J. Pelletier, editor, *Mass Terms*, pages 167–171, D. Reidel, Dordrecht, 1979.

[Par79b] Terence Parsons. An analysis of mass terms and amount terms. In F. J. Pelletier, editor, *Mass Terms*, pages 137–166, D. Reidel, Dordrecht, 1979.

[PR83] Barbara H. Partee and Mats Rooth. Generalized conjunction and type ambiguity. In Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, editors, *Meaning, Use, and Interpretation of Language*, pages 361–383, de Gruyter, Berlin, 1983.

[Pel79] F. J. Pelletier. Non-singular reference: some preliminaries. In F. J. Pelletier, editor, *Mass Terms*, pages 1–14, D. Reidel, Dordrecht, 1979.

[PS82] F. J. Pelletier and L. K. Schubert. From English to logic: context-free computation of "conventional" logic translation. *American Journal of Computational Linguistics*, 8(1):26–44, 1982.

[PS84] F. J. Pelletier and L. K. Schubert. Two theories for computing the logical form of mass expressions. In *Proceedings of Coling84*, pages 108–111, Association for Computational Linguistics, 1984.

[PS86a] F. J. Pelletier and L. K. Schubert. Generically speaking, with remarks on the interpretation of pronouns and tenses. 1986. Forthcoming in G. Chierchia, Barbara H. Partee, and R. Turner, editors, *Property Theory, Type Theory, and Semantics*, D. Reidel, Dordrecht.

[PS86b] F. J. Pelletier and L. K. Schubert. Mass expressions. 1986. Forthcoming in Dov Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic, Vol. 4*, D. Reidel, Dordrecht.

[PS86c] F. J. Pelletier and L. K. Schubert. Problems in the representation of the logical form of generics, plurals, and mass nouns. 1986. Forthcoming in E. LePore, editor, *New Directions in Semantics*, Academic Press, London.

[Pul88] Stephen G. Pulman. *Introduction to Computational Linguistics*. 1988. Cambridge University Press, Cambridge. Forthcoming. (Title not yet final).

[Tho76] Richmond H. Thomason, editor. *Formal Philosophy. Selected Papers of Richard Montague*. Yale University Press, New Haven, 1976.

Appendices

A CFG + Translation Relation for Mass Term Fragment

Translation relation, lexicon, and test examples for Pelletier & Schubert's fragment of English with mass nouns.

B CFG + Translation Relation for Basic Link Fragment

Translation relation, lexicon, and test examples for the basic part of Link's fragment of English with plurals (restricted to contain *basic* intransitive verb phrases only).

C CFG + Translation Relation for Full Link Fragment

Translation relation, lexicon, and test examples for the full Linkian fragment of English with plurals.

D CFG + Translation Relation for Mass Terms and Plurals

An amalgamation of the translation relations from **Appendix A** and **Appendix C** above, together with test examples.

D.1 Grammar

```
; Appendix D Grammar

; SYMBOLS: In this grammar, the proper plural operator is designated by
"plural",
; the closure-operator is designated by "clos", the PI-relation by "ispartof",
; the beta-operator by "convkind", the zeta-operator by "kind-of",
; and the mu-operator by "maxkind".


((_ log-arg))

[FEATURES
 aux[+,-] DEFAULT -  ; auxiliary verb
 inv[+,-] DEFAULT - ; included for compatibility with other grammars
 vform[inf, ing, en, tnsd] DEFAULT X ; included for compatibility with
other grammars
 takes[-,cnp,inf,ing,en, cnp_pp, p_cnp, nbar] DEFAULT X ; only "takes
cnp", "takes -", and "takes nbar"
; are used in this fragment
 stype[imp, ynq, whq, state] DEFAULT X ; included for compatibility with
other grammars
 agr[sing, plur] DEFAULT X
 count[+,-] DEFAULT +
 distr[+,-] DEFAULT -


 CATEGORIES
   Sigma (stype) ; top symbol
   termjoin() ; noun phrase joining conjunction "and"
   conj() ; sentence joining conjunction "and"
   CNP (agr count) ; Composite Noun Phrase
   N (agr count)
   name (agr)
```

```
  Nbar (agr count)
  Det (agr count)
  V (aux agr takes vform distr)
  VP (vform  agr  distr)
  NP (agr)
  S (agr inv)

TOP Sigma


Rule% State
 Sigma[stype state] -->    S[inv -]

Semantics% S


Rule% S->S+conj+S
  S --> S[agr X2] conj S[agr X3]

Semantics% (conj (S2) (S3))


Rule% S->CNP+VP
S[agr X1] -->
        CNP[agr X1]
        VP[vform tnsd, agr X1]

Semantics% (CNP VP)


Rule% CNP->CNP+termjoin+NP
 ; Rule for forming proper composite noun phrases such as
 ; "John and Mary", "a master and John", "the students and some masters".
 ; Such phrases are always plural (within the current fragment).
CNP[agr plur] -->
   CNP termjoin NP

Semantics% ((termjoin (CNP2))(NP))


Rule% CNP->NP
 ; Trivial cases of composite noun phrases, e.g. "the man", "John".
CNP[agr X1] -->
   NP[agr X1]
```

```
Semantics% NP


Rule% NP->massNbar
 ; Raises bare mass nouns, singular or plural,
 ; to the category of noun phrase.
NP[agr X] -->
  Nbar[agr X, count -]

Semantics%
  (agr NP plur) => (L (_Q) (E (_x) (and ((plural (convkind Nbar)) _x)(_Q
_x))))
          => (L (_P) (and (_P (maxkind Nbar)) (atomic (maxkind Nbar))))


Rule% NP->countNbar
 ; Raises bare count nouns to the category of noun phrase.
 ; The value of "agr" must be plural.
 ; The semantic translation makes such bare noun phrases "N"
 ; equivalent to "some N"---e.g. "students die" is equivalent
 ; to "some students die".
NP[agr plur] -->
  Nbar[agr plur, count +]

Semantics% (L (_Q) (E (_x) (and ((plural Nbar) _x)(_Q _x))))


Rule% NP->Det+Nbar
NP[agr X] -->
     Det[agr X, count X1]
      Nbar[agr X, count X1]

Semantics% (Det Nbar)

Rule% NP->Name
 ; All words beginning with a capital letter are assumed to be names.
 ; This rule raises their category to that of noun phrase.
 ; (Within the current fragment such noun phrases
 ; are taken to be singular and to denote atomic entities.)
NP[agr sing] -->
  name
```

```
Semantics% (L (_P) (and (_P name)(atomic name)))


Rule% Nbar->N
Nbar[agr X, count X0] -->
  N[agr X, count X0]

Semantics% N


Rule% VP->V+CNP
 ; This rule combines transitive verbs
 ; (with the exception of "be") with their objects to form
 ; full intransitive verb phrases. The semantic translation
 ; depends on the "distributivity-value" of the verb.
VP[agr X, vform X1] -->
   V[agr X, aux -, vform X1, takes cnp, distr X4]
   CNP

Semantics% (and (agr VP sing) (distr V +)) =>
(L (_x) (CNP (L (_y) ((clos V) (_x _y)))))
   (and (agr VP plur) (distr V +)) =>
 (L (_z) (A (_x) (if
 (and (atomic (_x))(ispartof (_x _z)))
 (CNP (L (_y) ((clos V) (_x _y)))))))
                    =>    (L (_x) (CNP (L (_y) (V (_x _y)))))



Rule% VP->Vaux+CNP
 ; The rule gives the "be-of-identity"-reading of verb phrases
 ; such as "are students", "are some wines", "are the wines",
 ; "is a student", "is a wine", "is wine".
 ; (A PTQ-style translation of "be" is assumed.)
VP[agr X, vform X1] -->
   V[agr X, aux +, vform X1, takes cnp]
   CNP[agr X]

Semantics% (V CNP)

Rule% VP->Vaux+massNbar
 ; The rule gives the "be-of-predication"-reading of verb phrases
```

```
; with a mass noun predicate such as "are liquids", "is liquid".
VP[agr X, vform X1] -->
   V[agr X, aux +, vform X1, takes nbar]
   Nbar[agr X, count -]

Semantics%  (agr VP sing) => (clos (kind-of Nbar))
  => (clos (convkind Nbar))


Rule% VP->pluralVaux+singularmassNbar
 ; The rule gives the "be-of-predication"-reading of verb phrases
 ; with a mass noun predicate such as "are liquid", i.e. the special
 ; case where the verb is in the plural and the noun is in the singular.
VP[agr plur, vform X1] -->
   V[agr plur, aux +, vform X1, takes nbar]
   Nbar[agr sing, count -]

Semantics% (clos (kind-of Nbar))


Rule% VP->Vaux+countNbar
 ; The rule gives the "be-of-predication"-reading
 ; of verb phrases such as "are students".
 ; Both the verb and the noun must be in the plural.
VP[agr plur, vform X1] -->
   V[agr plur, aux +, vform X1, takes nbar]
   Nbar[agr plur, count +]

Semantics%  (clos Nbar)


Rule% VP->V
 VP[agr X, vform X1] -->
        V[agr X, vform X1, takes -, distr X0]

Semantics% (distr V +) => (clos V)
       => V


]
```

D.2 Lexicon

```
; LEXICON FOR APPENDIX D GRAMMAR
     ((_ log-var))
       (

; SYMBOLS: In this lexicon, the proper plural operator is designated by
"plural",
; the closure-operator is designated by "clos", the PI-relation by "ispartof",
; the zeta-operator by "kind-of", and the beta-operator by "convkind".

; REDUNDANT FEATURES: the feature "type" is included
; exclusively for compatibility with other grammars.




; COLLEGES, PERSONS, ETC.:
; any input word beginning with a Capital letter is assumed to be a name.
; (The current translation relation further "stipulates" that its
; agreement is singular and its denotation atomic.)


; DETERMINERS (the, every, all, no, some, a)

;the
        (the
         (Det agr
              sing
      type n
              count
              +
              (lambda (_P)
                      (lambda (_Q)
                              (some (_x)
(and ((and (all (_y)
 (equiv ((clos _P) _y)
 (equal (_y _x)))) (_Q _x)))
              (atomic _x))))
                      )
              )
         (Det agr
              plur
      type n
```

```
            count
            +
            (lambda (_P)
                    (lambda (_Q)
                            (some (_x)
      (and
                             ((plural _P) _x)
   ((and (all (_y)
  (equiv ((clos _P) _y)
  (ispartof (_y _x)))) (_Q _x)))))
                             )
            ))
         (Det agr
            sing
      type n
            count
            -
            (lambda (_P)
                    (lambda (_Q)
                            (some (_x)
 (and ((and (all (_y)
  (equiv ((clos (convkind _P)) _y)
  (equal (_y _x)))) (_Q _x)))
            (atomic _x))))
                    )
            )



         (Det agr
            plur
      type n
            count
            -
            (lambda (_P)
                    (lambda (_Q)
                            (some (_x)
      (and
                             ((plural (convkind _P)) _x)
   ((and (all (_y)
  (equiv ((clos (convkind _P)) _y)
  (ispartof (_y _x)))) (_Q _x)))))
                             )
            )))
```

```
;every
        (every
         (Det agr
              sing
      type n
              count
              +
              (lambda (_P)
                      (lambda (_Q)
                              (all (_x)
      (if (and ((clos _P) _x)(atomic _x)) (_Q _x))))
                      )
              )
          (Det agr
              sing
      type n
              count
              -
              (lambda (_P)
                      (lambda (_Q)
                      (all (_x)
      (if (and ((clos (convkind _P)) _x)(atomic _x)) (_Q _x))))
                      )
              ))
;all
        (all
         (Det agr
              plur
      type n
              count
              +
              (lambda (_P)
                      (lambda (_Q)
                              (all (_x) (if ((clos _P) _x) (_Q _x))))
                      )
              )
          (Det agr
              plur
      type n
              count
              -
              (lambda (_P)
                      (lambda (_Q)
```

```
                              (all (_x) (if ((clos (convkind _P)) _x) (_Q
_x))))
                          )
                  )
          (Det agr
                  sing
        type n
                  count
                  -
                  (lambda (_P)
                          (lambda (_Q)
                  (all (_x) (if ((and ((clos (kind-of _P)) _x) (atomic _x))
(_Q _x))))
                          )
                  )))

;no
          (no
           (Det agr
                  X1
        type n
                  count
                  +
                  (lambda (_P)
                          (lambda (_Q)
                                  (all (_x) (if ((clos _P) _x) (not (_Q _x)))))
                          )
                  ))
;some
          (some
           (Det agr
                  sing
        type n
                  count
                  +
                  (lambda (_P)
                          (lambda (_Q)
                             (some (_x) (and (and ((clos _P) _x) (_Q _x))
                  (atomic _x))))
                          )
                  )
          (Det agr
                  plur
        type n
```

```
count
+
(lambda (_P)
        (lambda (_Q)
                (some (_x) (and ((plural _P) _x) (_Q _x))))
        )
)
(Det agr
     sing
type n
     count
     -
     (lambda (_P)
             (lambda (_Q)
             (some (_x) (and (and ((clos (convkind _P)) _x) (_Q
_x))
     (atomic _x))))
             )
     )
     (Det agr
          plur
     type n
          count
          -
          (lambda (_P)
                  (lambda (_Q)
                  (some (_x) (and ((plural (convkind _P)) _x) (_Q
_x))))
                  )
          ))
;a (indef. art.)
        (a
         (Det agr
              sing
        type n
              count
              +
              (lambda (_P)
                      (lambda (_Q)
                         (some (_x) (and (and ((clos _P) _x) (_Q _x))
  (atomic _x))))
                      )
              )
         (Det agr
```

```
          sing
type n
          count
          -
          (lambda (_P)
          (lambda (_Q)
          (some (_x)
 (and (and ((clos (convkind _P)) _x) (_Q _x))(atomic _x)))))))


; CONJUNCTIONS (and)


;and
(and
 (conj and)
 (termjoin
   (lambda (_np1) (lambda (_np2) (lambda (_P)
   (_np1 (lambda (_z) (_np2 (lambda (_y) (_P (i-plus (_z _y))))))))))))
;          (nbarjoin
;    (lambda (_Nbar1) (lambda (_Nbar2)
;    (intersect (_Nbar1 _Nbar2))))))

; NOUNS

(claret (N agr sing count - claret))
(clarets (N agr plur count - claret))
(riesling (N agr sing count - riesling))
(rieslings (N agr plur count - riesling))
(substance (N agr sing count - substance))
(substances (N agr plur count - substance))
(wine (N agr sing count - wine))
(wines (N agr plur count - wine))
(liquid (N agr sing count - liquid))
(liquids (N agr plur count - liquid))

        (student (N agr sing count + student))
        (students (N agr plur count + student))
        (master (N agr sing count + master))
        (masters (N agr plur count + master))
        (college (N agr sing count + college))
        (colleges (N agr plur count + college))
        (chapel (N agr sing count + chapel))
```

```
(chapels (N agr plur count + chapel))
(member (N agr sing count + member))
(members (N agr plur count + member))
(person (N agr sing count + person))
(persons (N agr plur count + person))
(man (N agr sing count + man))
(men (N agr plur count + man))
(woman (N agr sing count + woman))
(women (N agr plur count + woman))
(child (N agr sing count + child))
(children (N agr plur count + child))


; VERBS (Forms for 1. and 2. person singular are disregarded)

(build (V aux - takes cnp vform tnsd agr plur distr + build)
       (V aux - takes cnp vform tnsd agr plur distr - build))
(builds (V aux - takes cnp vform tnsd agr sing distr + build))

(see (V aux - takes cnp vform tnsd agr plur distr + see))
(sees  (V aux - takes cnp vform tnsd agr sing distr + see))
(drink (V aux - takes cnp vform tnsd agr plur distr + drink))
(drinks  (V aux - takes cnp vform tnsd agr sing distr + drink))
(die  (V aux - takes - vform tnsd agr plur distr + die))
(dies  (V aux - takes - vform tnsd agr sing distr + die))
(convene  (V aux - takes - vform tnsd agr plur distr - convene))
(convenes  (V aux - takes - vform tnsd agr sing distr - convene))

(are (V aux + takes cnp vform tnsd agr plur distr X
          (lambda (_np)
              (lambda (_x)(_np (lambda (_y) (equal _x _y))))))
  (V aux + takes nbar vform tnsd agr plur distr X emptystring))

(is (V aux + takes cnp vform tnsd agr sing distr X
          (lambda (_np)
              (lambda (_x)(_np (lambda (_y) (equal _x _y))))))
    (V aux + takes nbar vform tnsd agr sing distr X emptystring))


]
```

D.3 Some Comments on the Translation Relation of Appendix D

D.3.1 Background

The grammar of **appendix D** represents a naïve amalgamation of the fragments of English covered by the previous treatments of Link's system and Pelletier & Schubert's system, respectively. The amalgamation is naïve in the following sense:

- It is assumed that the two axiomatisations can coexist without trouble; in particular it is assumed that expressions such as

  $^{\bullet}(\beta\ claret)$ i.e. $(plural\ (convkind\ claret))$

  are meaningful and well-defined. That is, the "kind"-entities are dealt with by the closure- and the plural-operators just like all other entities. A necessary condition of this being meaningful without further ado is the fact that my implementation of Pelletier & Schubert has in a sense done away with their "semilattice of kinds". For instance, "the kind claret"—(μ *claret*)—will not denote the "individual sum of all the kinds of claret", but simply an atomic entity which enters into certain relations with other "kind"-entities in virtue of its translation (and the axiomatisation).

- No attempt at simplifying the amalgamated system has been made. Since the treatment of mass expressions stems from a "semi-lattice approach" just like the treatment of plurals, this might be possible. Obviously, there is a strong resemblance between the "is-part-of"-relation (Π) on one hand and the relations "is-a-kind-of" (ς) and "is-a-conventional-kind-of" (β) on the other. (But if the original semi-lattice of kinds is really a too weak construction, as I suspect, it may be quite difficult to simplify the amalgamated system after all).

The grammar given here is subject to a rather sharp restriction: it relies on all nouns being subclassified as either *count+* or *count–* (i.e. mass). If a noun is subclassified as (exclusively) *count+*, it cannot be interpreted as *mass* in any context—and likewise the other way round. The only way to make it possible for a noun to act in both capacities (together with the current grammar) is to give it dual lexical entries, and that certainly should be done for obviously ambiguous nouns such as "chicken"; but for many other nouns such duality would give rise to some implausible ambiguities. For instance, if "student" and "master" are entered as both *count+* and *count–*, then the sentence

"some students are masters"

would besides its count-interpretations have some mass-interpretations, which in my opinion would be highly implausible.

One way of overcoming this sharp division between *mass-* and *count-* nouns is to provide syntactic rules which "stretch" count nouns into mass nouns, where no other analysis will yield a grammatical sentence (see also [PS86b, pages 82-83]). That will have to await further work, though.

The present grammar allows us to study sentences with

- "pure" mass terms constructions in the singular and in the plural (e.g. "a wine is liquid", "clarets and wine are liquid");
- "pure" count noun constructions in the singular and in the plural (e.g. "John dies", "John and some students build a college and the chapels");
- constructions involving both count and mass expressions, in the plural as well as in the singular
(e.g. "John and the students drink wines").

D.3.2 Various Remarks

Apart from the general remarks above, I wish to draw attention to the following features of the grammar and lexicon:

Comprehensive extension and the γ-operator:

In this amalgamation I have done away with those two notions. All predicates P which occur unmodified in the output logical forms can be regarded as equivalent to $(\gamma\ P)$—that is, they are true of *individuals*, but not *kinds* or *quantities*.

The background of this simplification is partly the decision *not* to follow Pelletier in making all nouns capable of both a *mass* and a *count* interpretation, and partly the definition of *comprehensive extension* itself:

Definition 2 (Comprehensive Extension)
For all P in $\mathrm{ME}_{\langle a \rangle}$*, for all x:*
$P(x) \leftrightarrow_{\mathrm{def}} [(\gamma\ P)(x) \vee (\delta\ P)(x) \vee (\varsigma\ P)(x)]$

It follows from this definition that whenever we know an entity α to be e.g. a *kind*-entity, we can replace $P(\alpha)$ by $(\varsigma\ P)(\alpha)$—and analogously with the other sorts of entities. With the present fragment and its distinction between *mass*- and *count*-nouns, we always have this knowledge, and so we can dispense with the notion of *comprehensive extension*[132].

But then the next obvious step is to get rid of the γ-operator, which only blurs the connection between traditional logic and the logic used here.

There still are a few cases for which one might think that *comprehensive extension* predicates are needed. For instance, it is plausible that

"all wine is liquid"

should indeed quantify over all sorts of entities. Its translation[133] with the grammar and lexicon here is

$$\text{(i) } \forall x[\mathit{atomic}(x) \wedge {}^*(\varsigma\ \mathit{wine})(x) \rightarrow {}^*(\varsigma\ \mathit{liquid})(x)]$$

so it would seem that some of the required readings are missed out. However, an appropriate set of meaning postulates (cf. section 6.4) will allow us to infer from (i)

$$\text{(ii) } \forall x[\mathit{atomic}(x) \wedge {}^*(\beta\ \mathit{wine})(x) \rightarrow {}^*(\beta\ \mathit{liquid})(x)]$$

and

$$\text{(iii) } \forall x[\mathit{atomic}(x) \wedge {}^*\mathit{wine}(x) \rightarrow {}^*\mathit{liquid}(x)]$$

[132] In the axiomatisation for mass terms, a similar point was expressed by the axiom

Axiom 5 *For all P, Q in* $\mathrm{ME}_{\langle a \rangle}$: $Q(\mu\ P) \leftrightarrow (\varsigma\ Q)(\mu\ P)$

[133] Recall that this translation is only tentative, since no natural language inferences with this type of sentence have been studied thoroughly—cf. section 4 of the report. In fact, the sentence has one more translation, arising from its *be-of-identity*-reading, but that interpretation is not to the point here.

In this case it would admittedly have been simpler to have the notion of *comprehensive extension* available in the logical language itself, but in general it is easier to do without it, in particular w.r.t. drawing the proper inferences from the translated sentences.

The verb "be": The grammar distinguishes between the "*be-of-identity*" and the "*be-of-predication*". It may be worth pointing out the effect of this upon translations of sentences with mass nouns:

Recall that a bare plural noun phrase "X" is regarded as equivalent with "some X"; then, a sentence like

(1) "clarets are wines"

has two readings:

(2) $\exists x[^{\bullet}(\beta\ claret)(x) \wedge \exists y[^{\bullet}(\beta\ wine)(y) \wedge x = y]]$
(the "identity reading");

(3) $\exists x[^{\bullet}(\beta\ claret)(x) \wedge {}^{*}(\beta\ wine)(x)]$ (the "predication reading");

Reading (2) arises from the rule

VP $\Longrightarrow$ Vaux+CNP

while reading (3) arises from the rule

VP $\Longrightarrow$ Vaux+massNbar

Absence of composite "Nbars": A sentence like

(4) "clarets are wines and liquids"

receives only one interpretation, namely

(5) $\exists x[^{\bullet}(\beta\ claret)(x) \wedge \exists y[^{\bullet}(\beta\ wine)(y) \wedge$
$\exists z[^{\bullet}(\beta\ liquid)(z) \wedge x = y \oplus z]]]$

This is due to the fact that there are no rules in the grammar to cover "composite Nbars"; the only possible parse of "wines and liquids" is as a CNP. Thus (4) does not receive the rather obvious interpretation analoguous to (3) above, i.e.

$$(6)\ \exists x[{}^{\bullet}(\beta\ claret)(x) \wedge {}^{*}(\beta\ wine)(x) \wedge {}^{*}(\beta\ liquid)(x)]$$

A fuller grammar should provide the latter reading, too.

massNbar vs. countNbar: we have to distinguish between these cases, because bare *count* noun phrases occur only in the plural ("student" can never be a full (count) noun phrase); see the rules

NP $\Longrightarrow$ countNbar

and

NP $\Longrightarrow$ massNbar.

Furthermore, in predicative positions mass-nouns in the singular may co-occur with the copular verb in the plural, cf.

"clarets are liquid".

But for count nouns that pattern is unacceptable, cf.

∗ "John and Mary are student", ∗ "the masters are student".

(See the rules

VP $\Longrightarrow$ Vaux+massNbar

VP $\Longrightarrow$ pluralVaux+singularmassNbar

VP $\Longrightarrow$ Vaux+countNbar.)

E Project Proposals

Mike Gordon and Steve Pulman's project proposal to the Science and Engineering Research Council of the United Kingdom, *Logic and Natural Language Understanding*; Per Hasle's project proposal to the Danish Research Council for the Humanities, *En undersøgelse af sammenhængen mellem sprogfilosofi, matematisk logik og logikprogrammering* (*An Investigation into the Relationship between Philosophy of Language, Mathematical Logic, and Logic Programming*).

F Inference Engine for Mass Terms Axiomatisation

Thomas Forster's computational implementation of the axiomatisation of Pelletier & Schubert's theory of mass terms.

G Inference Engine for Plurals Axiomatisation

Thomas Forster's computational implementation of the axiomatisation of G. Link's theory of plurals.

H The Original Paper of Pelletier & Schubert

The version sent to Per Hasle in 1986, upon which the investigation into the theory of mass terms is based.