Studies in Linguistics and Philosophy

Essays in Logical Semantics

Johan van Benthem

D. Reidel Publishing Company

JOHAN VAN BENTHEM
Mathematics Institute, University of Amsterdam

ESSAYS IN LOGICAL SEMANTICS

D. REIDEL PUBLISHING COMPANY

A MEMBER OF THE KLUWER ACADEMIC PUBLISHERS GROUP

DORDRECHT / BOSTON / LANCASTER / TOKYO

Library of Congress Cataloging-in-Publication Data

Benthem, J. F. A. K. van, 1949–
Essays in logical semantics.

(Studies in linguistics and philosophy ; v. 29)
Bibliography: p.
Includes indexes.
1. Semantics. 2. Semantics (Philosophy) I. Title. II. Series.
P325.B426 1986 412 86–417

ISBN-13: 978-90-277-2092-4 e-ISBN-13: 978-94-009-4540-1
DOI:10.1007/ 978-94-009-4540-1

Published by D. Reidel Publishing Company,
P.O. Box 17, 3300 AA Dordrecht, Holland.

Sold and distributed in the U.S.A. and Canada
by Kluwer Academic Publishers,
190 Old Derby Street, Hingham, MA 02043, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers Group,
P.O. Box 322, 3300 AH Dordrecht, Holland.

ESSAYS IN LOGICAL SEMANTICS

STUDIES IN LINGUISTICS AND PHILOSOPHY

formerly *Synthese Language Library*

VOLUME 29

TABLE OF CONTENTS

INTRODUCTION

Recent developments in the semantics of natural language seem to lead to a genuine synthesis of ideas from linguistics and logic, producing novel concepts and questions of interest to both parent disciplines. This book is a collection of essays on such new topics, which have arisen over the past few years.

Taking a broad view, developments in formal semantics over the past decade can be seen as follows. At the beginning stands Montague's pioneering work, showing how a rigorous semantics can be given for complete fragments of natural language by creating a suitable fit between syntactic categories and semantic types. This very enterprise already dispelled entrenched prejudices concerning the separation of linguistics and logic. Having seen the light, however, there is no reason at all to stick to the letter of Montague's proposals, which are often debatable. Subsequently, then, many improvements have been made upon virtually every aspect of the enterprise. More sophisticated grammars have been inserted (lately, lexical-functional grammar and generalized phrase structure grammar), more sensitive model structures have been developed (lately, 'partial' rather than 'total' in their composition), and even the mechanism of interpretation itself may be fine-tuned more delicately, using various forms of 'representations' mediating between linguistic items and semantic reality. In addition to all these refinements of the semantic format, descriptive coverage has extended considerably. Nowadays, we possess valuable (though by no means conclusive) formal semantic accounts of a wide variety of linguistic phenomena beyond Montague's original samples — in particular, many of them independent from the intensional preoccupations which are the philosopher's burden.

There is also another type of development. Exhaustive description of fragments is useful, all the more so because of increasing contacts with computer science, trying to implement the above theories computationally. But, there remains the more global aim of understanding broad patterns in natural language, both within specific languages and across different ones. At this level too, logic and linguistics can meet and interact. Notably, in between general concerns of category/type fit and detailed semantic description of single lexical items, one can study the behaviour of various specific categories of expression in their entirety. For instance, the theory of 'generalized quantifiers' has inspired a logico-linguistic investigation of the category of determiner expressions in natural language, attempting to find out precisely the range of admissible semantic denotations for these expressions. Montague Grammar would allow, in principle, just any denotation of the appropriate type; but the above research has produced powerful and natural constraints. Although this goes beyond mere fit of syntactic category and semantic type, such an investigation does not go into complete lexical detail about any specific determiner expression. It is just this middle road which gives one a handle on important questions about natural language which have seemed hitherto rather metaphysical than scientific. For instance, given certain independently motivated constraints, which of the semantic possibilities left are actually realized by natural language expressions? If all of them are, this may be interpreted as the statement that natural language attains some optimum of expressibility — and we have realized part of the philosopher's dream, to explain *why* there are the things there are.

By itself, the generalized quantifier framework is just a convenient semantic format for determiner and quantifier expressions. But, it has proven very fruitful both as a medium of semantic description and a vehicle for semantic theorizing. The latter will be demonstrated in the first six chapters of this book.

Both general and special model-theoretic constraints are studied for *determiners* in Chapter 1, leading to several definability theorems. The notions and proof techniques obtained in this way are then used to evaluate recently proposed 'semantic universals', i.e., general regularities, of determiner meanings across all human languages. For instance, natural language contains 'systematic gaps': determiners with certain combinations of features are missing — and we want to understand why. (Thus, with one bold leap of the imagination, we double traditional areas of research, studying both what does and what does not occur in natural language.) Among the determiners, there is a distinguished class of 'logical' items, and such *quantifiers* are the topic of Chapter 2. Here, new themes arise for logical research. For instance, various intuitions of 'logicality' are developed, leading to more sophisticated hierarchies of logical constants than the usual set. Moreover, earlier conditions on determiner denotations can now be re-interpreted as possible patterns of inference — and we arrive at a study of 'inverse logic', classifying possible quantifiers validating given clusters of inferences. This, of course, is the mirror image of the usual Aristotelean mode of logical research, which describes inferential behaviour of already given quantifiers. The notions and results of the first two chapters can be generalized to arbitrary types of expression, in line with current tendencies to take a broader *cross-categorial* view of linguistic denotations. This is the theme of Chapter 3, which investigates the various manifestations of similar or related constraints across such categories as determiners, noun phrases, adjectives and connectives.

Up till this point, only extensional denotations have been considered. But, our type of investigation can also be transferred to an intensional setting. Chapter 4 demonstrates this for the case of *conditionals*, viewed as generalized quantifiers involving sets of possible worlds for their antecedents and consequents. One conspicuous topic here is to develop general intuitions of 'conditionality' in a more systematic fashion than is usually done in philosophical logic. Moreover, some unity of perspective results for the multitude of existing 'conditional logics' infesting the latter discipline. Afterwards, a similar road takes us into the traditional heartland of intensionality: the area of *tense and modality*. Thus, in Chapter 5, a reasonable hierarchy of denotational constraints provides a new classification of linguistic tenses.

Finally, in Chapter 6, another aspect of this enterprise is highlighted. These latter-day semantic trends are actually reminiscent of traditional logic, in particular the Syllogistic. Some connections between the two are explored, and especially, an outline is given of a *natural logic*, being a system of logical inference based directly on grammatical form, without any artificially created 'logical form' level.

It should be stressed again that these are theoretical questions, be it often with a direct descriptive motivation. Given the results obtained in these chapters, it seems that the present simple generalized quantifier perspective represents some optimum on the curve of compromises between faithful description and elegant general theory of natural language.

Next, the book turns to questions concerning the mechanics of interpretation. An interest in an account of semantic interpretability independent from syntactic grammaticality leads us to consider a more flexible *categorial grammar* allowing various rules of type change for expressions, as required by the varying needs of interpretation. Such a system of rules is gaining attention from a growing community of linguists these days, reviving the old Ajdukiewicz/Bar-Hillel framework. In particular, in Chapter 7, we shall provide a semantics for a system of type change rules essentially due to Lambek in the fifties, which has had to wait for recognition until the transformational juggernaut had passed.

Then, in Chapter 8, another more dynamic aspect of interpretation is considered. There is an attractive, though slightly marginal folklore idea that certain types of expression should be given 'procedural' denotations, i.e., procedures for computing suitable values. For the special case of quantifiers, and later on for other categories too, we find *semantic automata* doing just this. Surprising analogies then come to light with the Chomsky Hierarchy of grammars and automata, both in its coarse and its fine-structure. Thus, what used to be viewed as a stronghold of pure syntax, now becomes an asset of semantics too. By this road, the usual concerns of learnability and computability then also enter the semantic realm.

Finally, we ascend to our highest level of abstraction, asking various methodological questions about the semantic enterprise — using some of the apparatus of contemporary philosophy of science. As it turns out, semantic theories may be viewed as *empirical theories* in a standard sense, and Chapter 9 shows how central questions in the philosophy of science correspond to standard logical concerns. Notably, the usual industry of proving completeness theorems can now be motivated as a search for 'eliminability' of theoretical terms, such as accessibility or similarity relations in possible worlds semantics. Still, there arises a Popperian worry, viz. that the latter research program might be 'irrefutable', in the sense of being able to semanticize any kind of data. Fortunately, an example can be presented which is provably beyond the resources of the possible worlds machinery. This result has a wider significance: similar suspicions of 'infallible success', and hence lack of real explanatory achievement, surround the Montagovian paradigm.

More systematically, Chapter 10 is devoted to an ascending ladder of goals for a semantic theory, viz. providing a faithful (compositional) account of denotations, accounting for given (non-)inferences, suggesting global regularities in languages and even semantic universals. A *logic of semantics* will then consist of a multitude of questions concerning the prospects at each level; several examples of which are given.

Notably, there remains a need for a better understanding of linguistic 'information processing', enabling us to make more concrete logical sense of various intuitions of 'stability', 'minimal complexity' and 'efficiency'. And so, we have arrived at the Last Questions concerning natural language, which all discerning semanticists share and treasure.

The various chapters in this book are revised and expanded versions of a sequence of papers, many of which changed beyond recognition. I would like to thank the following institutions for their permission to use this material. D. Reidel Publishing Co. for 'Determiners and Logic' (*Linguistics and Philosophy* **6**, 1983, 447—478) [Chapter 1], as well as 'Foundations of Conditional Logic' (*Journal of Philosophical Logic* **13**, 1984, 303—349) [Chapter 4] — The Association of Symbolic Logic for 'Questions about Quantifiers' (*Journal of Symbolic Logic* **49**, 1984, 443—466) [Chapter 2] — North-Holland Publishing Co. for 'A Linguistic Turn: New Directions in Logic' (in P. Weingartner, ed., *Proceedings of the 7th International Congress in Logic, Methodology and Philosophy of Science, Salzburg 1983*, Amsterdam, 1986) [Chapters 3, 6] — The Center for the Study of Language and Information for 'A Manual of Intensional Logic' (CSLI Lecture Notes 1, Stanford, 1985) [Chapter 5] — Foris Publishing Co. for 'Themes from a Workshop' (in J. van Benthem and A. ter Meulen, eds., *Generalized Quantifiers in Natural Language*, GRASS series 4, Dordrecht, 1985, 161—169) [Chapters 3, 6], 'Semantic Automata' (in J. Groenendijk, D. de Jongh and M. Stokhof, eds., *Information, Interpretation and Inference*, GRASS series 5, Dordrecht, 1986) [Chapter 8] as well as 'The Logic of Semantics' (in F. Landman and F. Veltman, eds., *Varieties of Formal Semantics*, GRASS series 3, Dordrecht, 1984, 55—80) [Chapter 10] — John Benjamin Co. for 'The Semantics of Variety in Categorial Grammar' (in J. van Benthem, W. Buszkowski and W. Marciszewski, eds., *Categorial Grammar*, Amsterdam, 1986) [Chapter 7] — and the Polish Academy of Sciences for 'Logical Semantics as an Empirical Science' (*Studia Logica* **42**, 1983, 299—313) [Chapter 9] as well as 'Possible Worlds Semantics: A Research Program that Cannot Fail?' (*Studia Logica* **43**, 1984, 379—393) [Chapter 9].

And finally, I would like to thank all my colleagues in the Groningen circle of logic and linguistics — in our venerable free city at the cross-roads of semantic traffic from Poland, Scandinavia, America and Holland.

PART I

CONSTRAINTS ON DENOTATIONS

CHAPTER 1

DETERMINERS

A small group of linguistic categories forms the backbone of elementary sentence formation, as summed up in the following rewrite rules:

$$\begin{aligned} &S \Rightarrow NP\ VP \\ &NP \Rightarrow PN \\ &NP \Rightarrow Det\ N; \end{aligned}$$

involving the notions of Sentence, Noun Phrase, Verb Prase, Proper Name, Determiner and Common Noun. In what follows, we shall focus on the Noun Phrases in this scheme. Some of their parts seem to be 'interpretatively free', in the sense of allowing any available denotation in a model. Thus, in principle, proper names can denote arbitrary individuals, and common nouns can assume arbitrary extension sets of individuals. In contrast, determiner expressions ('determiners') form a more structured class, which is reflected in certain constraints on their semantics. Thus, the latter category of expression has been at the centre of attention in recent studies of possible denotations for natural language items. A related, more intrinsic reason is this: determiners provide the 'conceptual glue' with which we express basic relations between predicates (denotationally: $[\![Det]\!]\ ([\![N]\!], [\![VP]\!])$). Accordingly, determiners will be the first topic in this book, though by no means the last.

1.1. DETERMINERS IN LANGUAGE

To begin with, here are some actual determiner expressions. An extensive list for the case of English may be found in Keenan and Stavi (1982), both simplex (*all, some, two, both, most, few, enough, which,* etcetera) and complex. The latter comprise rather tightly knit compounds such as *almost all, all but two, at least three, the five, more than one, some of the four, too many,* but also broader categorial combinations, notably

— Boolean compounds:	Det $\Rightarrow$ *not* Det	(*not all*)
	Det $\Rightarrow$ Det $\left\{\begin{matrix} \textit{and} \\ \textit{or} \end{matrix}\right\}$ Det	(*all or some*)
— Adjectival restrictions:	Det $\Rightarrow$ Det Adj	(*no sane*)
— Possessives:	Det $\Rightarrow$ NP *'s*	(*some girl's*)

Actually, even such an empirical list already incorporates decisions of classification. For instance, it has also been proposed to analyze, e.g., *two, many* as adjectives rather than (or: as well as) determiners. Also, adjectival restriction is not equally acceptable to all linguistic observers. Such issues will not be investigated here. In any case, the general drift of our study tolerates a certain latitude.

Another syntactic restriction to be made was already implicit in the above. In accordance with the earlier rules, we shall focus on such schemata as $(all\ X)Y$, or rather the non-hierarchical *all* XY. Occurrences of determiners in non-subject position, such as direct object NP's or so-called 'floated quantifiers' (*the boys all volunteered*) will be largely ignored, except for some remarks at the end of this chapter. Generalization to these cases seems straightforward.

Also, the above expressions show varieties of meaning, not all of which can be studied here. In particular, we shall restrict attention to determiners that are

— *total* (no presupposition-bearing cases, as with *the* or *both*);

— *extensional* (no intensional phenomena, as in *all alleged*); and

— *discrete* or *countable* (no continuous uses, as in *some, much, all tea*).

None of these are essential limitations though, and some will be reconsidered in later chapters.

Another semantic issue down-played in our formulation is the variation of singular and plural forms for the predicates X, Y. We shall read the latter as standing for collections of individuals. (A semantics for plurals will be touched upon in Section 2.10., however.) But eventually, one will also want a systematic account of the semantic parallels and differences in such pairs as *all birds/every bird, many birds/many a bird, some birds/some bird, no birds/no bird.* It may not even be possible to handle this purely denotationally in the end, and some more 'representational' account may have to be super-imposed, in the style of Kamp (1981).

Let us now take a more systematic view of syntactic constructions creating new determiners. For a general setting, we use a categorial grammar, with basic types e (*entity*) and t (*truth value*) and functional combinations (a, b). In chapters 3, 7, this system will be studied in much more detail.

Determiners combine with (simple or complex) nouns to form noun phrases; the semantic net effect of which is displayed in the category assignment

$$((e, t), ((e, t), t)); \quad \text{or} \quad (p, (p, t)),$$

where $p = (e, t)$ stands for 'property'. Of the various possible questions, here is the most obvious one. Which categories can combine with determiners to form new determiners? More technically, what are the categorial solutions to the equation

$$x + \text{det} = \text{det}?$$

In strict categorial grammar, only one type of solution is forthcoming, most obviously $x = (\text{det}, \text{det})$. This accounts for the above Boolean negation on determiners. A little higher up, conjunction and disjunction may be accounted for in the same manner. But, one also wants to consider solutions where the left-most 'det' represents the function rather than the argument.

Now, any more realistic (and interesting) categorial grammar will allow something like the expansion rule of Geach (1972), combining

$$(a, b) + (b, c) \text{ to } (a, c).$$

Such more flexible grammars will be studied in Chapter 7. In the present case, this gives us the additional solution

$$(p, p) + (p, (p, t)) = (p, (p, t)).$$

As (p, p) is the category of adjectives, this accounts for the above adjectival restriction.

1.2. DETERMINERS AND GENERALIZED QUANTIFIERS

The well-known semantic treatment of noun phrases and determiners in Montague Grammar implicitly presupposes the logical 'generalized quantifiers' proposed in Mostowski (1957). The importance of the latter notion for natural language was brought out explicitly in Barwise and Cooper (1981). Basically, the idea is to let a noun phrase DX (*all* *women, most children, no men*) refer to a set of sets of individuals, viz. the denotations of those Y for which $(DX)\,Y$ holds. Thus, e.g., in a fixed model with universe E,

$$\begin{array}{lll} \textit{all } X & \text{denotes} & \{A \subseteq E \mid [\![X]\!] \subseteq A\}, \\ \textit{most } X & \text{denotes} & \{A \subseteq E \mid |[\![X]\!] \cap A| > |[\![X]\!] - A|\}, \\ \textit{no } X & \text{denotes} & \{A \subseteq E \mid [\![X]\!] \cap A = \emptyset\}; \end{array}$$

where $[\![X]\!]$ is the extension of the predicate X in the model. This point of view permits a uniform treatment of the subject/predicate form that pervades natural language.

Such denotations of noun phrases exhibit familiar mathematical structures. For instance, *all X* produces 'filters', and *no X* 'ideals'. The denotation of *most X* is neither; but it is still *monotone*, in the sense of being closed under supersets. Mere closure under subsets occurs too, witness *few X*. These structural properties are at present being used in organizing linguistic observations, and formulating hypotheses about them. In addition to the above-mentioned paper, one may mention the work of Ladusaw (1979) and Zwarts (1981) on 'negative polarity' and 'conjunction reduction'. In the course of such originally descriptive studies, several methodological issues of a wider logical interest arose, and these have inspired the present investigation.

In order to present these issues, we shift the above perspective, placing the emphasis on determiners per se — viewed as denoting *relations D* between sets of individuals. Thus,

$$DAB \quad \text{rather than} \quad B \in DA.$$

Even more generally, determiners pick out a binary relation among sets of individuals on arbitrary universes E, thus

$$D_E AB.$$

We shall picture this in the Venn Diagrams of classical logic (see Figure 1).

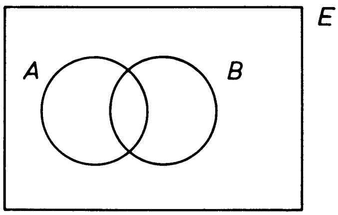

Fig. 1.

This point of view is not that of Mostowski but is related to the more general proposals in Lindström (1966), who introduced binary relations between arbitrary predicates (not just unary A, B).

A point of notation. We shall write DXY when thinking of a determiner expression followed by suitable linguistic items. When thinking about a determiner relation with sets (denotations for these items), we will write DAB. This abuse of notation is to be preferred to an abundance of denotation brackets.

The parameter E has its uses for context-dependent determiners. For instance, one of the meanings of *many* AB is that the relative frequency of B in A exceeds that of B in the whole universe E ('relatively many'). Another example is the compound determiner *all tall*, whose adjective may vary its denotation depending on the 'reference group' E.

Finally, an important decision remains to be made about these universes E. Our general feeling is that natural language requires the use of *finite models* only. Infinite models only arise (out of the former) through philosophical or scientific reflection. This restriction still leaves us a 'potential infinity' of arbitrarily large finite universes, with which to model linguistic phenomena — without sailing away into the 'actual infinity' of higher set theory. Perhaps, eventually, an actual infinity may be admitted, through reifying certain limits out of the finite realm — provided that some specific semantic motivation can be given. This strategy, which tends to make semantic modelling a lot more interesting, has turned out to be a bit of a shock to many addicts of Cantor's Paradise. For these, it should be pointed out that only few results in this book depend essentially on the finiteness restriction — and we shall occasionally note possible 'infinite' generalizations.

1.3. THE SEMANTIC RANGE OF DETERMINERS: CONSERVATIVITY

Which binary relations on a universe E are to count as admissible determiner denotations? There are two strategies of description here. One approaches from the outside, so to speak, accumulating global conditions, so as to fit to size. The other builds up from the inside, starting from evident cases, and giving an inductive generating procedure. We start with the former.

Global Conditions

One recurrent constraint accounts for the privileged role of the first argument in a determiner statement: it 'sets the stage':

CONS $\quad D_E AB \quad \text{iff} \quad D_E A(B \cap A)$ (*Conservativity*)

There is even a familiar idea that the common noun of an NP restricts the domain of evaluation for the determiner. Its full force is to declare everything outside of A irrelevant:

CONS$^+$ $\quad D_E AB \quad \text{iff} \quad D_A A(B \cap A)$.

The latter is a 'cross-contextual' constraint, relating various universes. Upon closer inspection, it combines Conservativity with a principle of 'context-neutrality' which may be stated separately:

EXT $\quad$ if $A, B \subseteq E \subseteq E'$, then
$D_E AB \quad \text{iff} \quad D_{E'} AB$ (*Extension*)

The latter principle holds for most determiners, exceptions being the above context-dependent cases.

In a sense, Extension plus Conservativity express the intuitive notion of 'aboutness'. Being 'about' A means remaining true for A, no matter how its context changes. For instance, in the famous 'Paradox of the Ravens' in the philosophy of science, *All ravens are black* is a rule about ravens, whereas its logical equivalent *All non-black things are non-ravens* is not about ravens in this sense, not being context-neutral.

Most of the proposed counter-examples to Conservativity have remained controversial, witness the following three cases. *Only willows weep* is a contingent statement, non-equivalent to the tautology *only willows are weeping willows*. But then, it may be argued that *only* is an adverb rather than a determiner, because of its syntactic distribution. Another case has occurred before: in the 'relatively'-reading, *many girls are giggling* may be false, while *many girls are giggling girls* is true. But here again, a different categorization has been proposed for independent linguistic reasons. (E.g., *many* is an adjective in Hoeksema, 1982.) Finally, there are intensional counter-examples. *All alleged males are females* may be true, even though *all alleged males are male females* is an analytic falsehood. But then, such intensional locutions call for an enriched semantic picture anyway, having families of possible universes — where an appropriate generalization of Conservativity will in fact be forthcoming. Thus, we shall stick with such prime examples as the equivalence between *all Dutch are morose* and *all Dutch are morose Dutch.*

Further general constraints on determiner denotations have been difficult to find. One candidate, perhaps, is the feeling that also the second argument of a determiner relation should 'matter':

VAR if A is non-empty, then there exist $B, B' \subseteq E$ such that $D_E AB$, not $D_E AB'$ (*Variety*)

Simplex determiners seem to satisfy this criterion, except for numerals, such as *two* in a universe with one single object. Complex cases also provide counter-examples. Accordingly, several weaker variants of Variety have been proposed (cf. Westerståhl, 1982). Nevertheless, we shall often use Variety to obtain a more surveyable field when proving results. Lifting this restriction is usually a matter of mere additional combinatorics. And in fact, further conditions, whether generally valid or not, still carve out interesting special classes of determiners, as will be seen below.

For the moment we are left with CONS as the only really obvious constraint on determiner denotations. In this connection, it is worth noting that this condition is preserved by the main closure conditions of Section 1.3., viz. Boolean combination and adjectival restriction (at least, with respect to *intersective* adjectives, obeying the scheme $[\![\text{Adj N}]\!] = [\![\text{Adj}]\!] \cap [\![\text{N}]\!]$). Such behaviour is not always displayed by VAR, or other special purpose conditions. Indeed, Keenan and Stavi have proposed a 'semantic universal' to the effect that

all human languages have only conservative determiners.

That one can do no better than this, in a sense, is the contention of the next section.

How much of a constraint is Conservativity? One way of answering such a question is by means of actual *counting*. On a fixed universe E, with n elements, there are 2^{4^n} generalized quantifier relations, being all sets of ordered couples of subsets of E. To explain the number 4, one views these couples (A, B) as functions from individuals in E to couples of truth values: yes/no (in A), yes/no (in B). Now, conservative determiners are completely specified by their pairs (A, B) with $B \subseteq A$, as is easily established. So, the value combination 'no', 'yes' drops out, and only three possibilities remain. Hence, there are 2^{3^n} conservative determiner denotations — a considerable reduction in size, though not in order of magnitude.

Inductive Conditions

Now, let us start from the inside, generating admissible determiner denotations. That the two principles of description coincide in the end is expressed in the main result of Keenan and Stavi (1982), of which a simplified, non-algebraic proof is given here.

Fixing some universe E, we start from some initial class of basic determiner relations, here: just *inclusion* (*all*) and *overlap* (*some*), allowing some reasonable constructions, here: *Boolean combination* and *restriction to intersective adjectives*. These generators plus operations create a class of determiner relations D-GEN. In a meandering fashion, using a different principle of generation, Keenan and Stavi prove their 'Definability Theorem':

$$\text{CONS} = \text{D-GEN}.$$

What does this equation mean? The number of conservative determiners increases with the size of E, as we have seen. Therefore, one cannot expect definitions for all determiners from some fixed finite stock, as its Boolean compounds will still be finite in number, up to logical equivalence. Thus, any definition obtained is bound to depend essentially on the universe E. Indeed, by comparing the counting formula for conservative determiners with a suitable one for Boolean definition from a fixed set of parameters (involving common noun and adjective denotations), Keenan and Stavi have also shown that, in the long run, the former will outrun the latter. Therefore, no global version of the Definability Theorem is possible.

Accordingly, we stay within a fixed E. Assume that individuals have been introduced only when they are distinguishable from all others already present in E through predicates definable in the language. By standard reasoning it follows that every set of individuals will be definable as the denotation of some (possibly complex) predicate of the language. Thus, every subset of E has a name in our language. Under these circumstances, we have the above identity, as will be shown now.

First, inclusion and overlap are indeed conservative relations, and the latter property is preserved under Boolean operations and intersective adjectives. E.g., if D is conservative on E, then, for any fixed set C, $D_E(C \cap A)B$ iff $D_E(C \cap A)(C \cap A \cap B)$ iff $D_E(C \cap A)(A \cap B)$, since $(C \cap A) \cap (A \cap B) = C \cap A \cap B$.

Conversely, let D_E be an arbitrary conservative relation. Observe first that $D_E AB$ iff $D_E A(B \cap A)$ iff $\exists X \subseteq A(D_E AX \;\&\; X = B \cap A)$. Therefore,

$$D_E AB \Leftrightarrow \bigvee_{\substack{D_E XY \\ Y \subseteq X}} (A = X \;\&\; B \cap A = Y),$$

and the problem is reduced to defining the latter conjunction in admissible terms. Here is one solution: take the conjunction of

— *all* $(Y \cap A, B)$
— *no* $(-Y \cap A, B)$
— *all* $(-X \cap A, B)$
— *some or not all* $(\{x\} \cap A, B)$ (for all $x \in X$).

Explanation: To begin with, note how this may be read as one long determiner, applied to arbitrary A, B. As for particulars, the relativizations are allowed since all sets mentioned (Y, $-Y$, $-X$, $\{x\}$ $(x \in X)$) were definable, while *no* merely abbreviates the admissible combination *not some*. Now, the first two clauses express that $B \cap A = Y \cap A$. In combination with the third, and the fact that $Y \subseteq X$, this yields $A \subseteq X$. Finally, by 'brute force', the last clauses ensure that $X \subseteq A$. (B is not really being used here, as the disjunction amounts to stating the fact that $\{x\}$ and A intersect.) □

The theorem should not be read as stronger than it is. For instance, although *most XY* (being conservative) has now been shown to be locally definable in terms of *some, all,* this is a matter of mere enumeration. After all, we know that no global first-order definition for *most* exists; witness Chapter 2. Still, one interesting interpretation of the theorem remains ('effability'). Under optimal local circumstances (when no anonymous individuals inhabit the universe), natural language has the resources for expressing every possible determiner denotation.

1.4. SPECIAL CONDITIONS: MONOTONICITY

Even without being generally valid, further conditions on determiners may be useful, both descriptively and theoretically. One prominent example is the following

if $D_E AB$ and $B \subseteq B'$, then $D_E AB'$ (*Monotonicity*)

Monotone determiners have a certain stability: when $D_E AB$ holds, even upon the basis of partial knowledge B about the denotation of the second predicate, this statement will remain true as more members of that denotation are discovered. Both reality and our information about it are in constant flux, but our language has to provide some more stable means of description. Hence, it is no surprise to find that basic determiners such as *all, most, some* are monotone.

Next, on the above scheme, there may be monotonicity in the left argument as well:

$$\text{if} \quad D_E AB \quad \text{and} \quad A \subseteq A', \quad \text{then} \quad D_E A'B.$$

This is the 'Persistence' of Barwise and Cooper (1981); valid for *some*, but not, e.g., for *all* or *most*. Moreover, these 'upward' versions also have obvious downward duals. Notably, the four resulting types of *Double Monotonicity* are exemplified in the traditional logical 'Square of Opposition'. With an obvious notation:

$\downarrow$*all*$\uparrow$ $\qquad\qquad\qquad$ $\uparrow$*some*$\uparrow$

$\downarrow$*no*$\downarrow$ $\qquad\qquad\qquad$ $\uparrow$*not all*$\downarrow$

And this connection is quite intimate:

THEOREM. The four determiners in the Square of Opposition are precisely those satisfying Variety and Double Monotonicity.

Proof. Here is a sample argument. Suppose that D is doubly monotone, of type $\downarrow$MON$\downarrow$. Then D must be the relation of disjointness (*no*):

— If $A \cap B = \emptyset$, then choose some non-empty $A' \supseteq A$. For some X, $DA'X$ (VAR), and hence $DA\emptyset$ ($\downarrow$MON$\downarrow$). Therefore also, DAB (CONS).

— If DAB, then $D(A \cap B)B$ ($\downarrow$MON), and so $D(A \cap B)(A \cap B)$ (CONS). By MON$\downarrow$ then, $D(A \cap B)X$ for all X, and hence $A \cap B = \emptyset$ (VAR). □

How restrictive is, say, upward Monotonicity by itself? Again, such a question may be approached by means of counting denotations. In Thijsse (1985), upper and lower bounds are obtained for the number of CONS, MON$\uparrow$ determiners (there are more than ${}_2 2^n$). An exact solution is still open — as it has been ever since Dedekind posed an essentially equivalent combinatorial question in the last century.

Unlike Conservativity, Monotonicity is not preserved under Boolean operations. For instance, *all or no* is not monotone. And even a simplex determiner such as *one* is not monotone either way. Yet, the latter does satisfy the weaker condition of *Continuity*:

$$\text{if} \quad D_E AB_1, D_E AB_2, B_1 \subseteq B \subseteq B_2, \quad \text{then} \quad D_E AB.$$

Essentially the following semantic universal may be found in Barwise and Cooper (1981):

every simplex determiner in natural language is continuous.

Note that the continuous determiners are exactly the conjunctions of an upward and a downward monotone one. (For instance, *one* is *at least one and at most one*. Cf. Thijsse, 1983.)

As with Monotonicity, Continuity can also occur in the left-hand argument:

$$\text{if}\ D_E A_1 B,\ D_E A_2 B,\ A_1 \subseteq A \subseteq A_2, \quad \text{then} \quad D_E AB.$$

For instance, *one* is both right- and left-continuous. But, e.g., *most* is neither left-monotone nor left-continuous. A counter-example is the case of $A_1 = \{1\}$, $A = \{1, 2\}$, $A_2 = \{1, 2, 3\}$ and $B = \{1, 3\}$. Due to the privileged role of the argument A, the left-hand versions of Monotonicity and Continuity are stronger than their right-hand ones — a phenomenon to be studied in Chapter 2.

Thus, special purpose conditions provide interesting classifications of determiners.

1.5. INFERENTIAL PATTERNS

The relational view of determiners invites the introduction of well-known conditions on binary relations from other areas of semantics and mathematics. For instance, *all* is *reflexive* and *transitive, some* and *no* are *symmetric, not all* is *connected*, etc. Such relational conditions may interact with the earlier notions:

THEOREM. All reflexive transitive determiners have Monotonicity type ↓MON↑.

Proof (cf. Zwarts, 1981). ↓MON: If $D_E AB$ and $A' \subseteq A$, then

$D_E A'A'$ (reflexivity), $D_E A'A$ (CONS), $D_E A'B$ (transitivity). MON↑: If $D_E AB$ and $B \subseteq B'$, then likewise, $D_E BB$, $D_E BB'$, $D_E AB'$. □

Incidentally, *at most one X is not Y* is a ↓MON↑ determiner which fails to be transitive.

Another kind of example is the following. In a study of 'definite determiners', Higginbotham has introduced so-called "properties of concepts": determiners only dependent on the intersection of their arguments. Formally,

$$\text{if } U \cap V = A \cap B, \text{ then } D_E UV \quad \text{iff} \quad D_E AB.$$

Immediate consequences are Conservativity as well as symmetry for D. But also conversely, the latter two conditions add up to being a property of concepts; through their consequence

$$D_E UV \quad \text{iff} \quad D_E(U \cap V)(U \cap V), \quad \text{for all } U, V.$$

Proof. $D_E UV$ iff $D_E U(U \cap V)$ iff $D_E(U \cap V)U$ iff $D_E(U \cap V)(U \cap V)$. □

A deeper interpretation of such relational conditions is possible. Essentially, they express *patterns of inference*, which may or may not be validated by certain determiners. Thus, we are now classifying determiners by their inferential potential.

A more systematic survey of possibly relevant relational conditions, then, starts with a catalogue of possible inference schemata. For instance, among the two-premise *syllogisms*, prominent examples are

$$\underset{(\textit{transitivity})}{\frac{DXY \quad DYZ}{DXZ}} \qquad \underset{(\textit{circularity})}{\frac{DXY \quad DYZ}{DZX}} \qquad \underset{(\textit{euclidity})}{\frac{DXY \quad DXZ}{DYZ}}$$

Then, one may ask for a classification of all possibilities.

Up till now, definite answers have only been obtained for the rather more special class of *logical* determiners ('quantifiers') to be introduced in the next section. Here is one example, taken from Westerståhl (1984):

THEOREM. The transitive reflexive quantifiers are precisely those of the forms *there are at most n X, or all X are Y* $(n = 0, 1, 2, \ldots)$

Thus, essentially, reflexivity and transitivity are characteristic properties of the universal quantifier.

Some potential patterns of inference are not even realized at all; witness the following result from Chapter 2:

THEOREM. There are no circular quantifiers, except the empty and the universal one.

Thus, natural language provides no vehicle for muddle-headed reasoning.

Likewise, there are no euclidean quantifiers (a conjecture in Zwarts, 1981, proved in Van Benthem, 1984d). For classifications of other cases, notably that of symmetric quantifiers, see Westerståhl (1984).

A more serious interpretation for non-existence results like the above is that they provide an explanation for observed 'systematic gaps' in natural language; a phenomenon noted in Barwise and Cooper (1981) and Zwarts (1981). For instance, one semantic universal proposed in the latter paper is that

no human language has asymmetric determiners.

At least for quantifiers, there is logical, rather than empirical necessity behind this observation: see Section 1.6. below.

To conclude, it should be noted that there are other relevant patterns of inference than purely relational ones. For instance, upward Monotonicity itself amounts to the step

'from *DXY* to *DX*(*Y or Z*)'

or equivalently

'from *DX*(*Y and Z*) to *DXY*'.

These are cases of interaction between determiners and connectives. This richer logic will be taken up in Chapter 2.

1.6. LOGICAL DETERMINERS

Some of the basic determiners are precisely the central *logical constants* called 'quantifiers'. Which additional constraints set these apart from the family of all determiners? One general intuition, upon which most authors seem to agree, is 'topic-neutrality', or insensitivity to individual traits of objects:

QUANT $D_E AB$ depends only on the numbers of individuals in A, $A \cap B$, B and E. (*Quantity*)

Thus, the numbers a, b, c, e in Figure 2 determine whether $D_E AB$ holds.

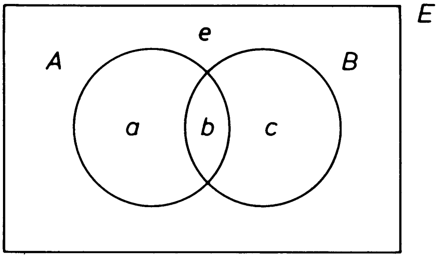

Fig. 2.

Here are some examples in this arithmetical setting:

all: $a = 0$, *some*: $b \neq 0$, *no*: $b = 0$, *all but one*: $a = 1$, *most*: $b > a$, *many*$_1$: $b > n(e)$ (where n is some norm function), *many*$_2$: $b/(a+b) > (b+c)/(b+c+a+e)$ (cf. Section 1.3.).

A detailed study of the *quantifiers* satisfying CONS, EXT and QUANT will be found in Chapter 2. For the moment, we just state how, on a universe with n elements, Quantity reduces the number of conservative relations by an order of magnitude: only $2^{(n+2)(n+1)/2}$ remain.

In addition to topic-neutrality, there are further broad intuitions concerning logicality. These seem to come in two strands.

Graduality. When the set A changes a little (resulting in a truth value change for $D_E AB$), the original truth value may be restored by means of some small corresponding change in B. This informal idea motivates various versions of the Continuity principle introduced in Section 1.4.

Uniformity. The behaviour of D should be regular ('the same') across all universes.

Actually, semanticists often speak about 'the model' for a language; and one might wish to implement the above intuition in some 'generic model', where a determiner receives its typical interpretation, once and for all. Lacking such a structure, we shall eventually present various formulations of Uniformity in terms of tables of behaviour across all finite universes (Chapter 2). As it turns out, the above informal idea then dissolves into a hierarchy of possibilities: how much uniformity we can 'see' depends on our conceptual apparatus (or, if one prefers, our metalanguage, where denotations are defined). Even without further technical details, the typical kind of result to come out of this can be stated thus:

among the quantifiers which are gradual, the first level of uniformity is exhausted by the Square of Opposition: *all, some, no, not all*; while the second adds a higher-order 'Square': *most, least, not least, not most.*

Beyond the second level, possibilities increase rapidly.

In this and earlier results, one obtains both 'positive' basic determiners (*all, some*) and 'negative' ones (*not all, no*); whereas natural language seems to favour the former at simplex level. Additional cognitive speculation may be in order here. Perhaps, the mind's eye has more difficulty in discerning absence than presence in our Venn Diagrams.

In the restricted area of logical determiners, effects of proposed semantic universals may be conveniently tested. For instance, the earlier-mentioned 'systematic gap' observed by Zwarts turns out to be a logical one:

THEOREM. There are no asymmetric quantifiers, except the empty one.

Proof. Suppose that $D_E AB$ holds anywhere. Recall that D satisfies CONS, EXT, QUANT. Now consider A, $A \cap B$, and add new individuals to create a symmetric situation (see Figure 3).

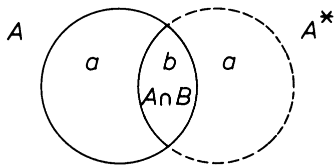

Fig. 3.

By the above three postulates, we have successively, DAA^*, $DA^*(B \cap A)$, DA^*A: and hence D has at least one symmetric pair; which refutes asymmetry. □

Thus, one can gauge the empirical content of proposed semantic universals. Always against the background of certain broad constraints, of course, which may be questioned themselves eventually.

1.7. BACK TO DETERMINERS IN GENERAL

Several of the intuitions presented in the preceding section seem equally attractive for (basic) determiners in general. For instance, the motivation for Graduality and Uniformity is not tied up with logical quantifiers — although some of their specific implementations may be. Indeed, in the perspective of this book, a demarcation of 'logic' has largely lost its interest, because there turns out to be so much logicality outside of logic proper.

The main principle of distinction between Section 1.6. and earlier ones was the use of Quantity. Pure numbers do not determine the meaning of the earlier adjectival restrictions, or of possessive determiners. We shall look into these cases now. But note that both concern complex determiner expressions. There do not seem to be non-quantitative simplex determiners, as is noted in Keenan and Stavi (1982). The adequacy of Quantity will be taken up again at the end of this section.

For present purposes, it is advantageous to reformulate Quantity to an equivalent, often encountered in the logical literature:

for every permutation π of the individuals in the universe E,
$$D_E AB \quad \text{iff} \quad D_E \pi[A]\pi[B], \quad \text{for all} \quad A, B \subseteq E.$$

Thus again, no individual feature of the objects involved is relevant to the truth of the determiner relation.

Let us now consider some other determiners, starting with adjectival restriction. An expression such as *all blond* satisfies Conservativity, but it lacks Quantity (see Figure 4).

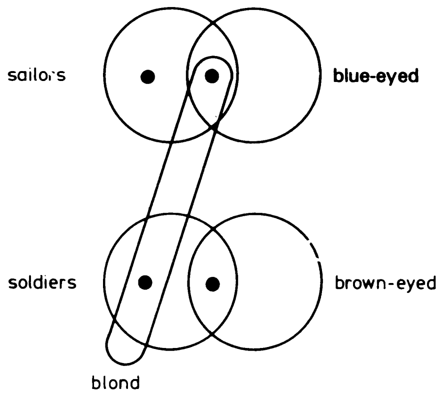

Fig. 4.

One can permute sailors and soldiers in vertical pairs. Yet *all blond sailors are blue-eyed* is true, while its permuted version *all blond soldiers are brown-eyed* is false.

But there is a remedy. Admissible permutations in this case ought to satisfy an additional feature, viz. respect of the predicate *blond*: $x \in [\![\text{blond}]\!]$ iff $\pi(x) \in [\![\text{blond}]\!]$. In other words, adjectival determiners are sensitive to additional structure of the universe. Thus, models will now become enriched structures $\langle E, \vec{P}\rangle$ with additional predicates $\vec{P}$: and Quantity becomes a maxim of *Quality*:

QUAL for every permutation of the universe E which is a *$\vec{P}$-automorphism* of $\mathscr{E} = \langle E, \vec{P}\rangle$, and all $A, B \subseteq E$, $D_{\mathscr{E}}AB$ iff $D_{\mathscr{E}}\pi[A]\pi[B]$.

Thus determiners have now become functors assigning binary relations among subsets to models of a given 'similarity type'.

In the new perspective, adjectival restrictions satisfy the obvious generalization of the Extension principle:

if $\mathscr{E} = \langle E, \vec{P}\rangle$ is a submodel of $\mathscr{E}' = \langle E', \vec{P}'\rangle$, then, $D_{\mathscr{E}}AB$ iff $D_{\mathscr{E}'}AB$, for all $A, B \subseteq E$.

Now, let us see how this works for possessive determiners. A basic case like *Mary's* satisfies CONS, but QUANT fails again (see Figure 5).

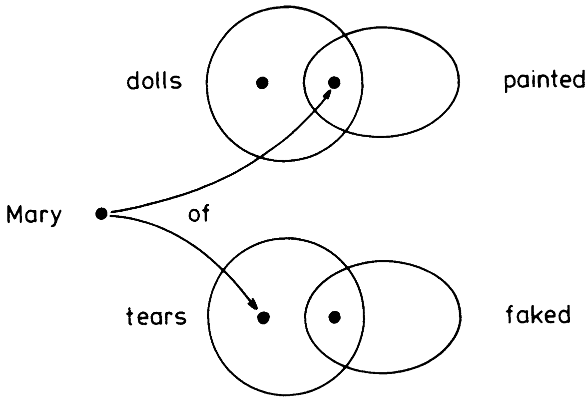

Fig. 5.

As with the sailor/soldier example, *Mary's dolls are painted* may be true, whereas the quantitatively similar *Mary's tears are faked* is false.

Evidently, possessives are sensitive to the underlying possession ties ('of') among individuals. But again, Quality holds with respect to a suitable similarity type: this time, 'individual constant (*Mary*), binary predicate (*'s*)'. And the same holds true of Extension. (For more complex possessives, such as *every girl's*, the latter principle needs more care.)

The general picture emerging is that of determiners satisfying a 'hidden variable' version of Quantity:

> for some suitable finite similarity type σ, D can be regarded as a functor on all models $\langle E, \sigma \rangle$, assigning a binary relation among subsets of E which is invariant for σ-automorphisms of E.

As we shall see later, this postulate allows us to transfer several results obtained for quantifiers to arbitrary determiners. How general is this strategy as a means of describing determiners? There are some mathematical questions in the background here concerning global definability of the permutation groups leaving a certain pre-assigned relation D_E invariant. But, the previous examples rather suggest that the relevant similarity type be read off directly from the linguistic items entering the description of the determiner, such as adjectives, proper names, affixes, etc.

With such additional structure present there also arise further reasonable constraints. For instance, one might desire *Non-Creativity*:

> a statement DXY implies no non-trivial facts about the underlying σ-structure.

In other words, determiners should not rule out any underlying patterns.

For an illustration consider the above example of *Mary's*. Which first-order definitions could it have in the language appropriate to its similarity type? First, the combined effect of Conservativity and strong Extension is to make $D_{\mathscr{E}}AB$ equivalent to $D_{A^+}A(B \cap A)$, where A^+ is the submodel of $\mathscr{E}$ with universe $A \cup \{[\![\text{Mary}]\!]\}$. Syntactically, this amounts to a relativization of all quantifiers in the defining formula to A^+, and hence to a formula involving only A-restricted quantifiers together with assertions about Mary.

Non-Creativity holds here, because the *reflexivity* of our example

(*Mary's XX*) even implies that, for every possession structure $\mathscr{E} = \langle E, M, of \rangle$, there exist $A, B \subseteq E$ such that $D_{\mathscr{E}}AB$.

Moreover, earlier special purpose conditions are still applicable. Notably, *Mary's XY* is upward monotone in Y — and so Y will only have 'positive' occurrences in the defining formula (cf. Section 2.5.). Also, downward monotonicity holds with respect to X — and even for the possession structure *of*: removal of *of*-ties from a model does not affect the truth of *Mary's XY*. By the earlier relativization condition then, it follows that our basic possessives are preserved under the transition to *submodels* of the original universe. By standard logic, then, the defining formula must be a purely *universal* predicate-logical sentence — in which, moreover, all occurrences of A, *of* will be syntactically 'negative' ones.

Thus, the following is the basic scheme for possessive determiners:

$$\forall^X \vec{y}(\gamma(of; x, \vec{y}) \rightarrow \Sigma Y\vec{y}),$$

and their conjunctions. Here, the superscript 'X' indicates relativization to the predicate X, γ is any positive quantifier-free condition, and $\Sigma Y\vec{y}$ some disjunction over a subset of $\vec{y}$. The simplest possible case is $\forall^X y(of\,xy \rightarrow Yy)$: which is precisely the meaning of the phrase x'sXY.

So, previous *notions* still apply to qualitative determiners. But what about earlier *results*? Most work remains to be done; but, at least, Chapter 4 contains an example of suitably generalized global intuitions of Graduality and Uniformity in a qualitative setting, with a corresponding classification theorem.

But, could one not take a short-cut, by 'immediate transfer'? There is a feeling, arising from practice, that for instance the non-existence results of Section 1.6. for logical determiners are quite characteristic for determiners in general. Here is one reason why.

Often, the relevant semantic regularities are of the form $\forall D\varphi$, where φ is some first-order assertion about the determiner D. For instance, the asymmetry example becomes $\forall D \neg \forall xy(Dxy \rightarrow \neg Dyx)$ — or, one can think of such valid connections as 'transitivity implies quasi-reflexivity ($\forall x \forall y(Dxy \rightarrow Dxx)$)' (cf. Zwarts, 1983). Our conjecture is that, for laws of this kind, validity for logical determiners implies validity for all determiners. (Thus, Quantity would be a 'conservative addition' to the universal theory of determiners.) Here is one case where this can be proved.

No new inferences. Let D be any qualitative determiner with respect to some similarity type σ (without individual constants or functions). Define a new determiner D^+ on universes E as follows. Set all predicates in σ equal to the *universal* relation (of the proper arity) on E. In this way, every permutation of E becomes a σ-automorphism of the enriched model $\mathscr{E}$. Then, setting D_E^+AB iff $D_{\mathscr{E}}AB$, makes D^+ a quantitative determiner. And evidently, D^+ will validate any inference schema which held for D. (For the converse question, see Section 4.10.)

But already with different similarity types, difficulties arise. For instance, the proof for the asymmetry universal does not go through for a possessive such as *Mary's*. The reason is that the necessary 'duplication' of A to A^* may not be possible. For, in case the proper name *Mary* denotes an object in $A - B$, it cannot also denote something in $A^* - B$. And in fact, if we could construe certain cases of *x's XY* as implying that x is in X, but not in Y, then there would be an asymmetric determiner after all. (Actually, this appears to be impossible — but the reason is not entirely clear.)

Still, whether by immediate transfer or through suitable generalization, there are good prospects for a general logical theory of determiners.

Even so, this direction of research may be misguided. Importing linguistic material from complex determiner expressions into their similarity types may not be superior to the obvious alternative: which is to bring in *explicit arguments* or parameters for determiner relations. Thus, adjectival restrictions (*all* Z)XY (meaning $[\![Z]\!] \cap [\![X]\!] \subseteq [\![Y]\!]$) could also be viewed as ordinary quantitative determiners in *three* variables. A similar observation holds for possessives — whose proper analysis is in dispute in any case, being tied up with the general semantics of genitives. And, once the influence of *case* is acknowledged, the latter had better receive a uniform treatment on top of our determiner account. Thus, it may be premature to give up Quantity at all.

1.8. FURTHER DIRECTIONS

As was stated at the outset, not all occurrences of determiner expressions have been treated here, nor even all possible uses of the subject position.

To begin with the latter, one obvious desideratum is an account of *plurality* extending the present one. Moreover, given the well-known analogies between plurals and *mass-terms*, the interplay between what may be called 'discrete' and 'continuous' quantification should be investigated. Some preliminary proposals on both scores will be found in the next chapter.

Then, there are the other occurrences of determiners in *direct objects, relative clauses*, in '*floated*' position, etc. Our prediction is that these will reduce to the present account, once an appropriate categorial perspective is adopted (cf. Chapter 3). For instance, although the clause *fears every wolf* seems to have the determiner connect a unary predicate with a *binary* one, there is an ordinary inclusion relation underneath — between $[\![\text{wolf}]\!]$ and $[\![\text{being feared by } x]\!]$, for some fixed object x. Likewise, floated occurrences are brought into the fold in Dowty and Brodie (1984).

More of a challenge is posed by ordinary subject occurrences in so-called 'donkey sentences'. Barwise and Fenstad have suggested that the familiar example *every farmer who owns a donkey beats it* expresses a determiner relation *every* between the two binary predicates $\lambda xy \cdot (farmer(x)\ \&\ owns(x, y))$ and $\lambda xy \cdot beats(x, y)$. More generally then, arbitrary determiners might have to relate pairs of predicates of higher arities. Technically, this is an interesting instance of the earlier-mentioned very generalized quantifiers of Lindström (1966). Nevertheless, there are some reasons for caution here. For instance, the meaning of these higher occurrences is often far from clear. Although the first example is usually read as inclusion of pairs, this format breaks down for other cases, such as *most farmers who* . . . . And even the meaning of the standard example is being debated. Thus, a proper generalization from the unary case needs a good deal of reflection.

Finally, an area which has been studied already to some extent is that of *determiners with more arguments*, such as *more X than Y*. As is shown in Keenan and Moss (1984), much of the preceding theory generalizes to this area, often in a surprising manner. One simple example is the new form of Conservativity:

$$DA_1 \ldots A_n B \quad \text{iff} \quad DA_1 \ldots A_n(B \cap (A_1 \cup \ldots \cup A_n)).$$

Summarizing, there is every reason to expect that the present approach will generalize to all types of determiners. Still, this does not mean that the generalized quantifier perspective is a unique best approach to noun phrases and determiners. In fact, there are important phenomena, such as *iteration* of determiners or *anaphoric* relations, about which it has little if anything to say. We have been throwing light upon natural language from a very specific angle.

CHAPTER 2

QUANTIFIERS

Ever since the days of Aristotle, quantifiers have occupied a central place in the logical study of reasoning. Traditionally, attention has been restricted to the four quantifiers in the Square of Opposition and their inferential behaviour. The more general perspective of our first chapter extends this field to arbitrary quantitative ties between predicates. As a result, various new directions arise for logical research.

2.1. WHAT ARE QUANTIFIERS?

Usually, the existence of a limited set of logical constants is taken as an ultimate fact. But, how is it that there are just these? Such metaphysical questions seem to go beyond the province of logic itself. As it turns out, however, exact answers may be obtained by formulating plausible intuitive constraints on 'logicality' of generalized quantifiers.

Recall that a *generalized quantifier* is any functor Q assigning, to each universe E, a binary relation Q_E between subsets of E. But, to qualify as a real logical 'quantifier', additional constraints are to be satisfied, some of which were already found in Chapter 1.

First, being determiners, quantifiers share the general property of *Conservativity*:

$$Q_E AB \quad \text{iff} \quad Q_E A(B \cap A) \qquad \text{(CONS)}$$

A more specific feature is their 'topic-neutrality': no individual plays a distinguished role. Mathematically, this becomes the invariance principle of *Quantity*:

for all permutations π of E, and all $A, B \subseteq \mathrm{E}$,
$Q_E AB$ iff $Q_E \pi[A]\pi[B]$.

Thus, the individuality of the members of A, B is discounted. But, one may go further than this local requirement, crossing boundaries between various contexts E:

for all sets E, E', all bijections π from E to E', and all A, B $\subseteq E$, $Q_E AB$ iff $Q_{E'} \pi[A]\pi[B]$ (QUANT)

'Pragmatically loaded' quantifiers such as *few* or *many* need not pass this test — but the ordinary ones, such as *all, most, some* and *no* do.

As we have seen already, once distinguished (groups of) individuals become important (as for *Mary's, no blue*), Quantity makes way for a weaker invariance principle of *Quality*, with respect to permutations respecting this additional structure.

Moreover, logical quantifiers are 'context-neutral', being invariant for *Extension* of the context:

for all E, E', and $A, B \subseteq E \subseteq E'$,
$Q_E AB$ iff $Q_{E'} AB$ (EXT)

This postulate allows us to drop the subscript 'E' whenever convenient.

The cumulative effect of these constraints may be pictured as in Figure 6.

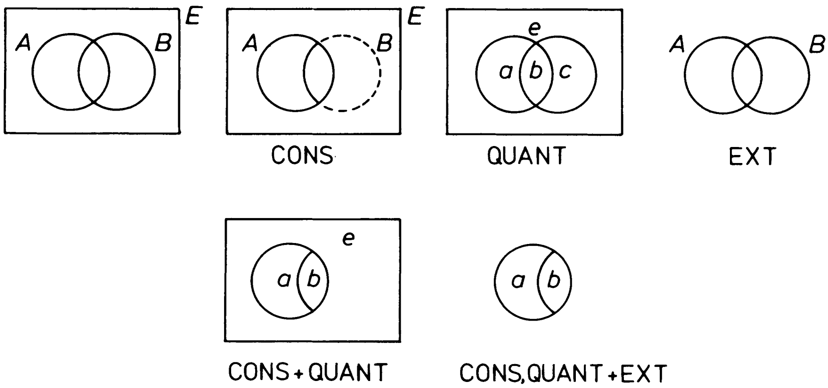

Fig. 6.

Finally, if the relation Q is really to depend on its second argument — or alternatively, if the logical constant is to do some work for us, it ought to exhibit some *Variety* of behaviour:

for all non-empty $A \subseteq E$, there exist $B, B' \subseteq E$ such that $Q_E AB$, not $Q_E AB'$ (VAR)

Weaker versions of this condition exist too, such as the following:

for every non-empty E, there exist $A, B \subseteq E$ with $Q_E AB$, but also $A', B' \subseteq E$ without $Q_E A'B'$.

Even the latter, more modest formulation still excludes numerical quantifiers such as *at least two* (consider a one-element universe).

Variety is not an essential postulate in our view. Nevertheless, it often facilitates exposition, allowing us to concentrate on essential cases first, postponing complicating combinatorics.

Next, in ordinary logical model theory, one would drift toward the usual score of questions concerning the present quantifiers; moving into the realm of infinite cardinalities (a, b), in order to apply current compactness or Löwenheim—Skolem arguments. But in the semantics of natural language, it may be argued that *finite* models are fundamental (cf. Section 1.2.). Therefore, we will usually avoid infinite cardinalities, even when this deprives us of slick logical methods. Indeed, several results in this chapter hold for finite models only. One would like to see more results in logical semantics where this characteristic assumption plays a crucial role.

2.2. THE TREE OF NUMBERS

The net effect of CONS, QUANT and EXT is to make a quantifier Q equivalent to the set of couples of cardinalities

$$(a, b), \quad \text{with} \quad a = |A - B|, b = |A \cap B|$$

which it accepts. A very convenient geometrical representation then arises, of quantifiers Q as subsets of the following 'Tree of Numbers' (or more prosaically, the north-eastern quadrant of the integer plane) (see Figure 7).

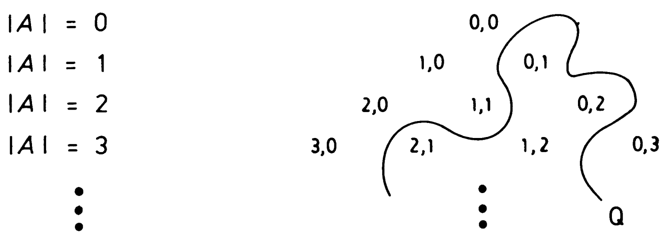

Fig. 7.

Conversely, representability in this tree implies the above three basic postulates.

Some examples of quantifier patterns are as follows, with markers + for Q, − for non-Q (see Figure 8).

all	at least two	half or more	all but an even number
+	−	+	+
− +	− −	− +	− +
− − +	− − +	− + +	+ − +
− − − +	− − + +	− − + +	− + − +
− − − − +	− − + + +	− − + + +	+ − + − +

Fig. 8.

Now, additional conditions on quantifiers will translate into geometrical constraints on quantifier sets, which are often easily visualized. For instance, VAR says that every row below the top must have occurrences of both + and −. Likewise, the earlier special purpose conditions of Monotonicity and Continuity (Section 1.4.) acquire concrete meanings. For instance,

— MON↑ expresses precisely that, if a point on the tree belongs to Q, then so do all points to the *right* of it on the same horizontal line;

— MON↓ is the analogous principle toward the *left*;

— ↑MON expresses that, if a point belongs to Q, then so do all points in the *downward triangle* generated by this point as a root (i.e., by successively adding units left and right); and

— ↓MON is the analogous principle in the *upward* direction.

One can easily draw pictures to illustrate this. Likewise, continuity conditions become *convexity* properties of quantifier sets. For instance, left-continuity corresponds to convexity in the natural geometry of the tree:

if $(a, b), (c, d) \in Q$, with $a \leqslant e \leqslant c$ and $b \leqslant f \leqslant d$,
then $(e, f) \in Q$ (see Figure 9).

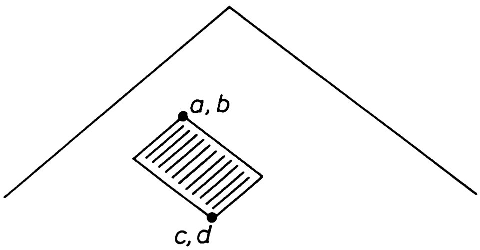

Fig. 9.

For proofs of these assertions, see van Benthem (1984d). But, they are really evident, once the principle of the tree representation is grasped.

Various technical applications of the above point of view will be found in the following sections and chapters. In addition, the tree picture itself suggests new conditions on quantifiers, as we shall see below.

2.3. MONOTONICITY

With the above point of view, more insight can be obtained into the central notions of Monotonicity, Persistence and Continuity, presented in Section 1.4.

To begin with, for logical quantifiers, Persistence is a very strong condition. Notice that all non-standard examples, such as *most, few* lack it, even though they do possess forms of Monotonicity.

THEOREM (VAR). The persistent quantifiers are precisely those in the Square of Opposition.

Proof. That all these quantifiers are persistent follows by inspection. Conversely, consider any persistent logical quantifier in the tree of numbers. There are only four possible top triangles (by VAR, the second row must already have $+-$ or $-+$). Each of these determines one quantifier in the tree, through the earlier geometric observations on monotonicity types. For instance, a top triangle ${}_{+}{}^{+}{}_{-}$ already violates $\uparrow$MON: whence the corresponding quantifier must be $\downarrow$MON. This again implies that $+$ can only occur on the left edge of the tree (otherwise, the indicated marker $-$ would have to be $+$); where indeed it must occur, by VAR. Thus, this case yields the quantifier *no*. $\square$

Without Quantity, this characterization fails. For instance, fix any object a. Set QAB if $A \cap B = \{a\}$ (when $a \in A$), or $A \cap B = \emptyset$ (when $a \notin A$). This quantifier satisfies all general postulates, except Quantity. Moreover, it is downward persistent — and yet outside of the Square. Note, however, that Q is not monotone, either way. With Double Monotonicity, the above result holds even in the absence of QUANT, witness Section 1.4.

Without Variety, many more patterns are available for double monotone quantifiers. Still, all of these can be classified geometrically.

For instance, consider the shapes of ↓MON↓ quantifiers in the tree. These are closed under 'upward trees' as well as 'left lines'; i.e., typically, each point (a, b) in the quantifier contributes the trapezoid (a, b), $(a+b, 0)$, $(0, 0)$, $(0, b)$ (see Figure 10).

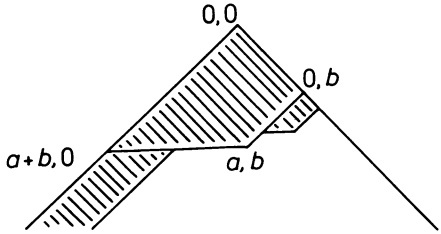

Fig. 10.

Geometric inspection of the possible shapes allowed by this closure property reveals a finite union of such trapezoids, possibly together with an infinite band along the left edge of the tree. Any such pattern may be viewed as an intersection of unions of regions of the types

— an infinite band (*at most k A are B*)
— a top triangle (*there are at most n A*).

Thus, every ↓MON↓ quantifier is logically equivalent (on finite universes) to a conjunction of sentences of the types

there are at most n A, or at most k A are B,

or equivalently, *at most k out of every n + 1 A are B*. Two more melodious special cases are $n = 1$, $k = 0$: *no A is B*, $n = k + 1$: *at most k A are B*. From this classification a description for the remaining three double monotonicity types is easily extracted.

Observe that all quantifier patterns mentioned are *first-order definable* in a monadic predicate logic with two unary predicates and identity. Indeed, the left-hand side is crucial here:

all persistent patterns are first-order definable.

Proof. This result has the following geometrical explanation. First, consider ↑MON quantifiers. Every point generates a downward triangle, and therefore, the whole quantifier must be a finite union of such triangles, as in Figure 11.

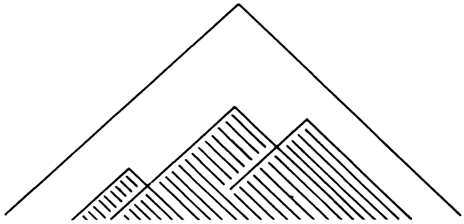

Fig. 11.

For, start from any triangle within the quantifier: only finitely many steps can be made toward the edges of the whole tree. Evidently, every pattern of this kind is first-order definable.

As for ↓MON cases, notice that their *negations* are ↑MON quantifiers. □

Using this method of shapes, Westerståhl has even shown the following, on finite universes:

THEOREM. All left-continuous quantifiers are first-order definable.

Thus, our constraints on quantifier denotations turn out to be related to more traditional logical ways of description.

Various further definability results using the Tree may be found in van Benthem (1984d) and Westerståhl (1984). Another type of application is made in Thijsse (1985), where counting formulas are obtained for an array of denotational conditions. One elegant example is this: assuming VAR, the number of left-continuous quantifiers is counted by (every second member of) the well-known *Fibonacci Sequence*. There must be some deep truth here about the semantic flora.

2.4. FURTHER INTUITIONS OF LOGICALITY

The three postulates QUANT, CONS, EXT, together with VAR, delimit an important, but still rather heterogeneous class of generalized quantifiers. It contains many of the usual first-order quantifiers, but also some highly artificial second-order ones. Can one single out a more homogeneous group? Indeed, there are various further intuitions concerning quantifiers (or logical constants in general) that may be brought to light. These additional postulates draw the dividing line in a way which does justice to our feeling that some higher-order quantifiers, such as *most* or *least*, are relatively natural ones, while most others seem unrealistic.

Graduality. One persistent idea is that there should be a certain 'smoothness' in the semantic behaviour of basic quantifiers. The various continuity notions of earlier sections already captured several forms of this phenomenon.

The most innocent, and attractive formulation is surely the right-hand version:

$$QAB,\ QAC,\ B \subseteq D \subseteq C \text{ imply } QAD.$$

In the tree of numbers this means that the quantifier intersects horizontal rows in uninterrupted stretches. But, stated in this way, the principle is biased toward presence of the quantifier relation — unlike the earlier intuitions, which treated absence symmetrically. Therefore, it seems equally reasonable to demand continuity of absence:

$$\text{not } QAB, \text{not } QAC,\ B \subseteq D \subseteq C \text{ imply not } QAD.$$

Together, these principles enforce right-monotonicity (upward or downward) in each row of the tree. Their conjunction will henceforth be referred to as CONT.

But, there is more to the above intuition. Looking in a vertical direction, one also expects regular behaviour of a quantifier across rows. In particular, there should be smooth transitions between adjacent horizontal rows. That is, there should be no dead-lock. If QAB holds, and one adds a new individual to A, then at least one of the two options (enlarging $A - B$ or enlarging $A \cap B$) must again result in truth for Q; and similarly for falsity of Q. In terms of the number tree:

$$\begin{array}{llllll} \text{if} & (a, b) \in Q, & \text{then} & (a+1, b) \in Q & \text{or} & (a, b+1) \in Q, \\ \text{if} & (a, b) \notin Q, & \text{then} & (a+1, b) \notin Q & \text{or} & (a, b+1) \notin Q. \end{array}$$

This postulate will be called PLUS.

Together, CONT and PLUS express a strong form of continuity in all three main directions of the tree: $\leftrightarrow$, $\nearrow$ and $\searrow$. (This point of view will return in Chapters 8, 10.)

Finally, we come to the most esoteric, but perhaps also the most fundamental of our intuitions of logicality.

Uniformity. The above graduality theme is interwoven with a related idea, viz. that basic quantifiers should have a 'uniform' behaviour. No cardinality pair (a, b) should be special, so to speak. One dynamic way of implementing this idea is by identifying a quantifier with a (recursive) *procedure* for assigning truth values to cardinality pairs — which is the main topic of Chapter 8. Right now, we opt for a more static formulation, however.

The earlier addition of an individual to some 'situation' may be regarded as a typical *thought-experiment* for testing the behaviour of a quantifier. Starting from an arbitrary (a, b) (with Q true or false), one notes the truth values for $(a+1, b)$ and $(a, b+1)$. In all, there are eight possible truth value patterns for this experiment (of which PLUS rules out the outcomes $\substack{+\\-\;-}$ and $\substack{-\\+\;+}$). A straightforward version of uniformity is then the following:

> *for each truth value, the addition experiment has the same triangle of outcomes everywhere.*

Thus, it does not matter where we perform our test: Q will behave uniformly. This postulate will be called UNIF.

Thus, three additional postulates have been extracted from our intuitive ideas about the desired regularity of quantifiers. Which quantifiers (if any) are left by these, in combination with the earlier constraints?

THEOREM. On the finite sets, the only generalized quantifiers satisfying QUANT, CONS, EXT, VAR as well as CONT, PLUS, UNIF are *all, some, no* and *not all.*

Thus again, the Square of Opposition emerges.

Proof. First, these four quantifiers satisfy all seven constraints. Conversely, consider the tree of numbers. Which +/− patterns are allowed by these conditions? At the top, there may be either + or −. At the next row ($a + b = 1$), more possibilities appear, and hence we distinguish some cases. Case 1: + on top. By VAR, the second row must be +− or −+. Consider the former first. By UNIF, the third row will start with +− again: and then, by CONT, its last entry must be −. By UNIF once more, the pattern now extends downward, to form the quantifier *no*. Analogously, the other case becomes the quantifier *all.* In a similar manner, a top position − generates only the two possibilities *not all* and *some.* □

Given the small range of quantifiers remaining, it becomes of interest to re-examine the uniformity idea. Perhaps, this should be relaxed, so as to allow different outcomes for the thought-experiment. But then, where else should the desired regularity be located?

Let us conduct the experiment, first increasing *a* only, noting the outcome, then increasing *b* only, but finally adding a unit to *both*, thus restoring the original balance. Even allowing different patterns for the first two separate tests, one can at least demand *unicity of outcome* for the final test: *each truth value pattern for the addition experiment determines a unique outcome for the combined move.*

But our idea of uniformity goes further than this. As before, the particular place (a, b) where the experiment occurs should be immaterial; and hence we also require 'repetition' in the following sense: *if the combined addition experiment restores the original truth value, then it will repeat itself at the new location.* This new version of the third postulate will be called UNIF*. Here is the statement of its impact.

THEOREM. On the finite sets, the only generalized quantifiers satisfying QUANT, CONS, EXT, VAR as well as CONT, PLUS, UNIF* are *all, some, no, not all* together with *most, not most, least, not least.*

Thus, one next 'Square of Opposition' emerges.

Proof. Again, these eight quantifiers meet all seven requirements. Conversely, one checks possibilities in the number tree. Here is one typical case. Let the top position be +, followed by a second row +−. By PLUS and CONT, the third row can only be ++− or +−−. At this stage, the revised uniformity condition comes into play. With the former third row, the first combined experiment has been restorative $(+\,{}^{+}_{+}\,-)$, whence it will repeat itself. Consequently, two more restorative patterns appear on its sides, viz. $+\,{}^{+}_{+}\,+$ and $+\,{}^{-}_{-}\,-$. By UNIF* and CONT, then, the quantifier must be *not most.* With the latter third row, however, the first experiment $({}_{+}{}^{+}{}_{-})$ now has outcome −, and, by unicity of outcomes, this phenomenon extends downwards. Then, again by UNIF* in combination with CONT, the quantifier becomes *no.* The remaining cases are entirely analogous, the top triangle ${}_{-}{}^{+}{}_{+}$ producing the quantifiers *not least* and *all,* while the negative top position generates the remaining four quantifiers in a wholly symmetric fashion. □

Thus, the 'respectable' higher-order quantifiers have been found in the same boat with the basic first-order ones. This type of reasoning is quite flexible, yielding many additional insights. For instance, it may be seen that leaving out CONT altogether would still generate a *recursively enumerable* class of admissible quantifiers.

Obviously, the uniformity intuition has not received a final evident form here. There seems to be rather a whole hierarchy of uniformity notions, and this may conceivably give rise to a cumulative hierarchy of quantifiers in 'degrees of uniformity'.

Appendix: Variants of Uniformity

As the development of more volatile intuitions concerning denotations is an unfamiliar topic in semantics, we add one further illustration.

One attractive form of Uniformity requires endless repetition of truth value patterns, in the spirit of Mandelbrot's 'fractals':

> *equal truth values in the tree generate equal downward tree patterns.*

Its effect is to leave only the following possibilities:

no, an even number of, all, all but an even number of,
some, an odd number of, not all, all but an odd number of.

Proof. By Variety, there are only four possible top triangles. Each of these then splits up into two possibilities for the third row, after which all downward propagating patterns have become fixed. □

An obvious weakening of this condition would allow for some *finite variety* of tree patterns for each truth value. This lets in quite a few additional quantifiers, in a hierarchy following the number of distinct patterns allowed.

Again, these examples illustrate our two main concerns: *definability* results for quantifiers satisfying new intuitive constraints, as well as *hierarchy* results classifying our quantifiers in layers of denotational complexity. Thus, we obtain 'natural kinds' of quantifier beyond the traditional realm of logical constants.

A final remark. The above notions of Uniformity are connected with the earlier-mentioned procedural point of view. For instance, the requirement of equal generated triangles for all + positions (and likewise, for all − positions) allows for a representation of the relevant quantifiers by means of two-state *finite state machines*. (Here, + becomes an accepting state, − a rejecting one; with transition arrows for scanning further *a* or *b* individuals read off in the tree.) In Chapter 8, this machine perspective will be developed further, with several pertinent results in Section 8.2. In general, the above hierarchy of patterns of 'finite uniformity' corresponds exactly to a hierarchy of quantifiers computed by finite state machines with varying numbers of states; as finitely homogeneous trees may be contracted to finite transition graphs.

2.5. FIRST-ORDER DEFINABILITY

The preceding sections have been mainly devoted to notions arising out of current semantics of natural language. But a more traditional logical question has also appeared at times. How are the quantifiers studied here related to those expressible in the usual logical languages, from monadic first-order logic upward? And furthermore, are any new questions generated concerning the latter?

Starting with the simplest case, the class of first-order definable quantifiers is wider than just the monotone ones. For instance, *precisely one* is nonmonotone, yet first-order definable. Let us now consider the latter class in the light of the preceding.

A first goal is to find a semantic characterization of first-order definable quantifiers on finite models. Now, in model theory, the answer for the general case is provided by the Keisler—Shelah theorem, in terms of *isomorphs* and *ultraproducts*. But the latter construction has no significance in the finite realm. There is also the *Fraïssé back-and-forth characterization*, however, which does go through in this restricted area. For the monadic predicate language, this amounts to invariance for models that are alike up to some fixed threshold. More precisely, set

$$X \sim_n Y \text{ if either } |X| = |Y| = k < n, \text{ or } |X|, |Y| \geqslant n.$$

By extension, set $\langle E, A, B\rangle \sim_n \langle E', A', B'\rangle$ if the relevant four monadic slots stand in the $\sim_n$-relation to their primed counterparts. The relevant characterization then becomes as follows:

THEOREM. On the finite models, a quantifier Q is first-order definable if and only if, for some fixed n,

$$\langle E, A, B\rangle \sim_n \langle E', A', B'\rangle \quad \text{implies} \quad Q_E AB \quad \text{iff} \quad Q_{E'} A'B'.$$

Again, the tree of numbers suggests a geometric way of viewing the characteristic behaviour of first-order quantifiers. In the light of the above theorem, these are the ones that, after an initial 'Sturm und Drang' phase, reach a 'Fraïssé threshold', i.e., a line $a + b = 2n$ such that

— the truth value at (n, n) determines that of its generated downward triangle,

— all truth values at $(n+k, n-k)$ are propagated along their downward left lines (parallel to the edge), and

— all truth values at $(n-k, n+k)$ determine that of their downward right lines; as in Figure 12.

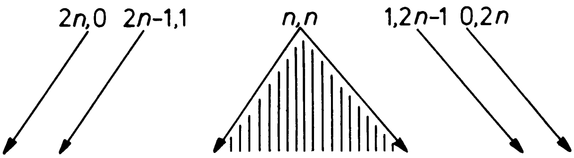

Fig. 12.

Thus, on finite sets, the first-order quantifiers are essentially just *finite unions of convex* (*and hence left-continuous*) *quantifiers*. This is a kind of converse to the left-continuity theorem of Section 2.3. By way of contrast, note how a quantifier like *most* fails to be first-order definable on finite universes: as its pattern lacks a characteristic triangle of the above kind.

The second main goal is again one of explicit classification. Here is an early description of monadic first-order logic with identity: *all sentences are logically equivalent to Boolean compounds of the types*

at most k (*non-*)*A are* (*not*) *B* and
there are at most k (*non-*)*A*.

Using the tree representation, such classifications are easily verified by geometrical inspection. Thus, monadic first-order logic merely adds some book-keeping devices to the simple quantifiers in the Square of Opposition.

The duality between a structural notion (threshold invariance) and its syntactic counterpart (first-order definability) is typical of logical model theory. Many more questions concerning this interplay arise from the preceding sections. Here are two examples.

— To find a preservation result characterizing Conservativity of first-order sentences. The obvious conjecture is that $\varphi(A, B)$ is (strongly) conservative in A iff φ is logically equivalent to some sentence with all its quantifiers *A-restricted.* Kit Fine has observed that this follows, indeed, from the work of Feferman (1969).

— To find a preservation result for monotonicity of first-order sentences. This time, the obvious conjecture is the following. $\varphi(A, B)$ is upward monotone in A iff φ is logically equivalent to some sentence whose only occurrences of A are syntactically 'positive' (in the usual sense). Again, Kit Fine has shown how this follows, by a simple deduction, from the Lyndon version of Craig's interpolation theorem.

Another type of question was suggested by Jon Barwise. Instead of searching for definability, one may consider what happens when a certain kind of generalized quantifier is *added* to the first-order language. For instance, let us add Q in the form of assertions $Qxy \cdot \varphi(x)$, $\psi(y)$ — or $Q\lambda x \cdot \varphi(x)$, $\lambda y \cdot \psi(y)$. What about the logic of such an enriched language? Evidently, first-order logic remains valid, and so does any condition imposed on Q that can be expressed in the language. But, will there be additional 'mixing principles'? For the case where Q is *monotone*, Barwise has shown that no such new principles appear. Without proof, we state a similar result for one of our key notions:

> *predicate logic with an added conservative generalized quantifier has its universally valid principles axiomatized by the usual predicate-logical axioms and rules of inference plus the conservativity axiom.*

Perhaps surprisingly, predicate logic with an added *quantitative* generalized quantifier has a non-recursively axiomatizable logic (cf. Väänänen, 1980).

But actually, all these questions may be too traditional. Having gained the generalized quantifier perspective, as a better mirror of natural language, we should be wary of the usual formalisms.

Query: Can one do preservation and completeness results (and model theory in general) directly in terms of the generalized quantifier framework — and that in an enlightening way?

A preliminary attempt at developing logical theories of inference in this way will be found in Chapter 6. One important step on that road will be taken right now.

2.6. INFERENTIAL CONDITIONS

The above questions take us back to an earlier theme (cf. Section 1.5.) of inferential conditions on determiners and quantifiers. Quantifiers exhibit familiar relational properties, such as

transitivity	: $\forall XYZ(QXY \,\&\, QYZ) \rightarrow QXZ)$	: *all*
reflexivity	: $\forall XQXX$	: *all, most*
symmetry	: $\forall XY(QXY \rightarrow QYX)$	: *no, some*
antisymmetry	: $\forall XY((QXY \,\&\, QYX) \rightarrow X = Y)$	: *all*
irreflexivity	: $\forall X \neg QXX$	: *not all*
linearity	: $\forall XY(X = Y \vee QXY \vee QYX)$	: *not all*

Several of these play a role in the semantic literature. Plain reflexivity and irreflexivity are prominent in Barwise and Cooper (1981) (under the names of 'positive strong' and 'negative strong'), symmetry was important in Section 1.5. Even less common properties are exemplified, such as 'quasi-reflexivity' ($\forall XY(QXY \rightarrow QXX$), which holds for *some*. Likewise, *no* is 'quasi-universal': $\forall XY(QXX \rightarrow QXY)$. (Curiously, these properties can often be found as conditions on alternative relations in possible worlds semantics — a connection whose explanation is obscure.) Yet more exotic properties may be realized through Boolean compounds of the above simplex quantifiers.

Of course, such common conditions on binary relations need not be the most appropriate ones in the present area. But the above list at least suggests a certain relevance, that will become clearer below.

Combinations of relational conditions can be used to classify 'natural kinds' of quantifiers. But one should be careful here. For instance, many beautiful results can be proved about 'the partially ordered quantifiers': but it turns out that one is describing only one single specimen:

THEOREM. Inclusion (*all*) is the only reflexive antisymmetric quantifier.

Proof. If $A \subseteq B$, then QAB (by reflexivity and Conservativity, as before). If QAB, then $QA(B \cap A)$ (CONS), but also $Q(B \cap A)A$, just as before: and hence $A = B \cap A$, (by antisymmetry), i.e. $A \subseteq B$. □

More positively, one can think of this result as a characterization of the universal quantifier by its inference patterns.

COROLLARY. *Not all* is the only quantifier that is irreflexive and linear.

Proof. The characteristic properties for *all* produce those for *not all*, as the former may also be regarded as the negation of the latter. $\forall XQXX$ goes to $\forall X \neg QXX$. $\forall XY((QXY \;\&\; QYX) \rightarrow X = Y)$ goes to $\forall XY((\neg QXY \;\&\; \neg QYX) \rightarrow X = Y)$, which is equivalent to linearity. □

Similar characterizations may be proved for *some, no*; be they of a rather more artificial nature. In the presence of VAR, these results become more elegant, as we shall see below.

To obtain larger classes, one must relax requirements. Generally, mere *transitivity* is very restrictive already:

THEOREM. If Q is a transitive quantifier, then, on finite models, QAB implies $A \subseteq B$ or $QA\emptyset$.

Proof. Suppose that QAB without $A \subseteq B$. By CONS, $QA(B \cap A)$, with $B \cap A$ properly contained in A. Now, let B' be a *minimal* set properly contained in A such that QAB'.

Claim: $B' = \emptyset$.

For, otherwise, choose A' such that $|A'| = |A|$ and $A' \cap A = B'$. (Here, CONS and EXT are presupposed.) Therefore, since QAB', also QAA'. Now, consider any permutation π leaving B' as well as possible individuals outside of $A \cup A'$ fixed, while interchanging $A - B'$ and $A' - B'$. By QUANT, it follows that $Q\pi[A]\pi[B']$, i.e., $QA'B'$.

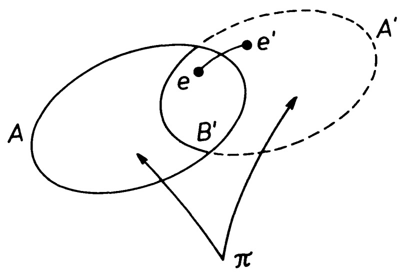

Fig. 13.

Next, choose any $e \in B'$ and $e' \in A' - B'$. Let the permutation π' interchange only e and e', leaving all other individuals fixed. Again by QUANT, $Q\pi'[A']\pi'[B']$, i.e., $QA'\pi'[B']$. By transitivity then, QAA' and $QA'\pi'[B']$ imply $QA\pi'[B']$: whence $QA(\pi'[B'] \cap A)$ (by CONS). But the latter intersection is properly contained in B', contradicting the latter's minimality. □

This type of argument is typical for the use of Quantity.

The conclusion of the theorem cannot be improved to read inclusion only. E.g., 'all A are B or there is at most one A, is a transitive quantifier admitting a non-inclusion case. Also, the theorem fails for infinite sets: 'A is infinite and $A - B$ is finite' is a transitive quantifier.

Using these proof techniques, one can classify various natural sets of quantifiers by their inferential patterns. For instance, we recall the earlier result (Section 1.5.) that all *reflexive transitive* quantifiers are precisely the forms *there exist at most n A, or all A are B.* This result may also be verified in the tree of numbers, using a transcription from relational into geometrical or numerical conditions. A systematic use of this method is found in Westerståhl (1984).

Once we assume Variety, all these cases collapse into one.

THEOREM (VAR). *All* is the only reflexive transitive quantifier.

Proof. This follows from two results in Chapter 1. Reflexive transitive quantifiers have monotonicity type ↓MON↑, and, modulo Variety, there is just one of the latter kind, viz. inclusion. □

Again, the earlier negation transform implies a similar result for noninclusion:

COROLLARY. *Not all* is the only irreflexive and almost-connected quantifier.

Here, *almost-connectedness* is the following basic property of comparatives: $\forall XYZ(QXY \rightarrow (QXZ \vee QZY))$.

On the other hand, completely negative results are also of interest, showing 'systematic gaps' of unrealizable conditions. As we have seen in Sections 1.5./1.6. already, such results confirm the suspicions of certain linguists, who had formulated 'semantic universals' to this effect. For instance, we have shown already that *there exist no asymmetric quantifiers* (except for the empty one). And again by the negation transform, it follows that *there are no strongly connected quantifiers* (except for the universal one). Further reflections on the interplay between logical and linguistic (im)possibility will be found in Chapter 10.

2.7. INVERSE LOGIC

Relational conditions on quantifiers are universal first-order sentences which may be regarded as expressing *patterns of inference*. More systematically, here are some 'pure' patterns, involving only a single quantifier Q:

0-premise: $\dfrac{\quad}{QAA}$ (reflexivity) $\qquad \dfrac{\quad}{QAB}$ (universality)

1-premise: $\dfrac{QAB}{QBA}$ (symmetry) $\qquad \dfrac{QAB}{QAA}$ (quasi-reflexivity)

$\dfrac{QAA}{QAB}$ (quasi-universality) $\qquad \dfrac{QAB}{QBB} \quad \dfrac{QAA}{QBA}$

2-premise: $\dfrac{QAB \qquad QBC}{QAC}$ (transitivity), etc.

Here and in the sequel, attention will be restricted to inferences where all statements are of the atomic form QAB. (Thus, e.g., disjunctive conclusions '$QAB \vee QCD$' are not considered.)

The third case, with two premises sharing a 'middle term', is the area of the Aristotelean *Syllogistic*. Apart from transitivity, the following interesting patterns are found:

$$\frac{QAB \qquad QBC}{QCA} \qquad \frac{QAB \qquad QCB}{QAC} \qquad \frac{QBA \qquad QBC}{QAC}$$

(circularity) (anti-euclidity) (euclidity)

Now, what Aristotle did was to take specific logical constants Q, and ask which patterns were validated by them. This has been the dominant question in the logical tradition ever since. What we have been doing, however, is a converse activity: *given a set of inference patterns, to determine the range of logical constants realizing them*. This change in perspective might be called another 'Copernican Turn' — this time, challenging deep Aristotelean presuppositions inside logic itself.

Some answers have been found in the preceding sections already, be it in a nonsystematic fashion. There are few transitive quantifiers, and no euclidean ones, as we have seen. Moreover, the earlier methods of proof yield many further answers. For instance, there are no symmetric-reflexive or symmetric-transitive quantifiers. Instead of exhaustive cartography, here is a representative example.

THEOREM. There are no nontrivial circular quantifiers.

Proof. By earlier types of argument (see Section 1.6., Figure 3), a quantifier with QAB anywhere has QAA^*, QA^*A for some A^* — and hence QAA, by Circularity. So, circular quantifiers are *quasi-reflexive*. Hence, if QAB, then also QAA, and, again by Circularity: QAA, QAB imply QBA: circular quantifiers are also *symmetric*. But then, they are *transitive* as well. Now, consider any non-empty circular quantifier, with QAB somewhere. Either $A \subseteq B$ (i) or it does not (ii). Case (ii): by an earlier result on transitive quantifiers (Section 2.6.), it follows that $QA\emptyset$. But then, Q will hold for arbitrary pairs of sets C, D (and hence be trivial after all). For, consider any C, D. We have: $Q\emptyset A$ (symmetry), $Q\emptyset\emptyset$ (CONS), $Q\emptyset(C \cap D)$ (CONS), $Q(C \cap D)\emptyset$ (symmetry), $Q(C \cap D)(C \cap D)$ (quasi-reflexivity); and hence QCD, as in the characterization of symmetric determiners given in Section 1.5. Finally, case (i) may be reduced to case (ii). Let QAB with $A \subseteq B$. Choose B' properly containing B: QAB' (CONS). Then consider $QB'A$ (symmetry). □

Other types of logical question arise too. For instance, the earlier characterizations of the universal quantifier suggest the following. 'Holistically', the meaning of a logical constant is given by the sum total of its valid inference patterns. Taking the pure case first, it may be asked *whether the usual quantifiers are uniquely determined by their valid inference patterns*. In subsequent chapters, we shall see how this query generalizes to other logical constants as well.

Reflexivity plus transitivity turned out to characterize *all* (modulo Variety). What about its dual?

THEOREM (VAR). *Some* is the only quantifier that is both symmetric and quasi-reflexive.

Proof. Let Q be an arbitrary symmetric quasi-reflexive quantifier. Then Q must be the overlap relation. First, suppose that $A \cap B \neq \emptyset$. By VAR, there exists $X \subseteq A \cap B$ such that $Q(A \cap B)X$. Then

$Q(A \cap B)(A \cap B)$ (quasi-reflexivity), $Q(A \cap B)A$ (CONS), $QA(A \cap B)$ (symmetry) and hence QAB (CONS). Next, assume that QAB. Suppose that $A \cap B = \emptyset$. Then $QA\emptyset$ (CONS). As in the preceding proof, it follows that QCD for all sets C, D: which contradicts VAR. □

By the negation transform, there is an immediate

COROLLARY. *No* is the only quantifier that is both symmetric and quasi-universal.

Summarizing the previous results then, modulo Variety, the Square of Opposition is characterized by the following pure inference patterns:

all:	transitive, reflexive	*some*:	symmetric, quasi-reflexive
not all:	almost-connected, irreflexive	*no*:	symmetric, quasi-universal

Of these four quantifiers, *not all* is the only one without valid syllogisms of the original kind. (Its inferences all involve 'meta' negation and disjunction.) Perhaps, this accounts for the empirical fact that no human language seems to have found it necessary to contract it to a smoother simplex form.

Removing the condition VAR allows more quantifiers. Indeed, there arises a noticeable underdetermination:

THEOREM. *At least two* validates the same pure inference patterns as *some*.

Proof. First, suppose that some Q-inference is refuted in a model where Q is non-empty intersection. Then this model can be inflated to an *at least two*-counterexample, simply by adding new individuals $e_{X,Y}$ to X and Y, whenever such a couple of sets overlaps. Next, if some Q-inference is refuted in a model where Q is *at least two*, singleton intersections are to be removed in order to obtain a *some*-counterexample, without changing the relational pattern. This time, the procedure is this: as above, add new $e_{X,Y}$ and $e'_{X,Y}$ to X and Y, for each pair X, Y with $|X \cap Y| \geqslant 2$. Then, just strike out all old individuals. □

By the same reasoning, *some* turns out to have the same syllogistic theory as *at least n* (for any $n = 1, 2, 3, \ldots$). The theorem does not disclose the form of that theory, however. Here it is.

THEOREM. Symmetry and quasi-reflexivity comprise the complete pure inferential theory of *some*.

Proof. Both are valid, of course: the converse is the crux. Suppose that $\varphi(Q)$ is any universal first-order sentence about Q which does not follow from the above two principles. By Gödel's Completeness Theorem, φ is falsified in some symmetric quasi-reflexive order. We shall be done if this can be transformed into a counterexample for φ where the relation is overlap between sets.

First, contract all (possible) isolated points to a single one. This contraction is a so-called *strong homomorphism*, leaving (at least) all universal first-order sentences about Q invariant. Next, the resulting model may be represented as a set of doubletons (and singletons) with the overlap relation, by setting

$$x \overset{F}{\longmapsto} \{\{x, y\} \mid Qxy\}.$$

It may be checked that, successively, F is one-one, F preserves Q, and F preserves overlap. □

Similar results may be obtained for the family of quantifiers *at most n* $(n = 0, 1, 2, \ldots)$.

Finally, characterization of logical constants may also be viewed from another angle. Instead of imposing strong conditions such as Variety, one may also increase the number of inferences involved, by considering schemata in which several quantifiers occur at the same time. (An algebraic analogy may be helpful here. In the pure case, one is searching for a unique solution to a system of equations with one variable Q. In the mixed case, one searches for simultaneous solutions of the form (Q_1, Q_2); etc.)

A partial result in this direction was obtained in van Benthem (1985a) (see also Westerståhl, 1985 for a strengthening):

THEOREM: The complete syllogistic theory of *some* and *all* is satisfied by precisely all couples

at least n X are Y / *there are at most n* $-$ *1 X or all X are Y* (with $n = 1, 2, \ldots$).

Thus, in a sense, traditional logic fails to enforce its intended interpretation, at least inferentially.

Yet, there are also mixed inferences involving both quantifiers and connectives, such as forms of monotonicity, persistence, and even conservativity itself. Also, linguists have been interested in NP-denotations that are *filters*, being sets of sets closed under supersets (i.e., being MON$\uparrow$) and also under the formation of *intersections*: from QXY and QXZ to $QX(Y \textit{ and } Z)$. In the finite case, the latter possess one smallest 'generator', which can be recovered — something which has been used in the analysis of so-called 'definite' noun phrases. Our previous considerations are fully applicable to such new notions. One sample result is that, using VAR, *all* is the only 'filtrating' quantifier. Other, dual notions occur too, such as 'idealizing' quantifiers (in their left-hand argument), such as *all* and *no*. (Cf. van Benthem, 1984d.)

Thus, there arises the truly holistic question *whether the (pure and mixed) valid inferences of predicate logic determine precisely the usual interpretation of the logical constants*. An answer to a question of this kind would be a deep type of completeness theorem.

An Afterthought

The preceding sections have illustrated two approaches to the notion of logicality for quantifiers. One proceeds by way of broad semantic constraints on denotations, such as Quantity or Uniformity. This makes quantifiers rich in 'semantic transfer': if they hold somewhere, they will hold in many similar situations. The other route, logically the more standard one, makes the relevant quantifiers rich in 'inferential potential': logical constants are those key words and phrases oiling the wheels of reasoning. These two points of view are not necessarily co-extensive — and their connection ought to be clarified.

2.8. INFINITY

In the remaining three sections, possible extensions of the above 'standard theory' will be considered, all of them rather tentative.

In Chapter 1, a 'finitizing program' was advocated for semantics, re-admitting the usual infinite models only when some plausible reconstruction for them can be found. This is clearly a debatable position, and hence nothing prevents us from having a look at the infinite realm already.

In van Deemter (1985), a survey is made of some central notions and results in the above, with an eye toward reducing or removing the finiteness restriction. As it turns out, many results go through at once, or may be modified to do so. The only more intrinsically finitistic subjects are certain definability theorems (e.g., the classifications of monotone quantifiers) as well as the earlier uniformity hierarchy.

With quantifiers on both finite and infinite universes, three broad groups emerge. There are obvious 'extrapolations' of the earlier examples, there are also some natural essentially infinite ones (such as *finitely many, infinitely many*), and there is any number of more remote mathematical possibilities. Two important questions then arise. One is to formulate suitable notions of *extrapolation* from patterns in the finite tree of numbers to patterns in the 'Infinite Tree', having rows at each infinite cardinality.

For instance, one might adopt a 'Stabilization Principle':

Let $n \in N$,

— if $(a, m) \in Q$ for all $m \geqslant n$, then $(a, \aleph_0) \in Q$

— if $(a, m) \notin Q$ for all $m \geqslant n$, then $(a, \aleph_0) \notin Q$

— and likewise for the middle column (m, m) $(m \geqslant n)$ toward $(\aleph_0, \aleph_0)$.

For example, in this way, the finite pattern of *all* will extend to the first infinite row — and beyond, once the above is suitably generalized. Likewise, the finite quantifier *at least nine tenths* will produce the infinite pattern of *almost all*, in the sense of 'all, but for finitely many exceptions'.

These ideas do not always suffice, however. For instance, *most* only determines the first infinite row up to its middle position $(\aleph_0, \aleph_0)$. In such a case, two strategies are possible. One is to allow undefined cases (which leads us to a *three-valued* approach, which may be attractive for presupposition-bearing quantifiers in any case), another is to introduce more sophisticated limit rules.

Another interesting question is to characterize the 'natural' infinite quantifiers by our earlier method of denotational constraints. Here is one example, having to do with the central notion of monotonicity.

Consider the following class of quantifiers Q, satisfying the properties of

D-MON: monotonicity (upward or downward) in both arguments;

PLUS: if $(a, b) \in Q$, and k is any cardinality (finite or infinite), then there exist k_1, k_2 with $k = k_1 + k_2$ such that $(a + k_1, b + k_2) \in Q$.

And likewise for non-Q.

The latter condition is van Deemter's generalization of the 'smoothness' intuition PLUS of Section 2.4.

To exclude presently irrelevant considerations of higher infinite cardinality, we shall demand that Q be non-trivial on *countable* cardinalities (i.e., not always true or always false there): a kind of 'Löwenheim condition'.

THEOREM. The only quantifiers satisfying D-MON and PLUS under the Löwenheim restriction are those in the 'finite Squares of Opposition':

> *at most n not, at least n + 1, at least n + 1 not, at most n* ($n = 0, 1, \ldots$) together with the infinite Square
> *at most finitely many not, at least infinitely many,*
> *at least infinitely many not, at most finitely many.*

This certainly delineates a very natural class.

Proof. All quantifiers mentioned satisfy the above conditions. Conversely, we shall consider one typical example. Suppose that Q has type ↑MON↑. Well-known arguments give its pattern in the finite tree of numbers: either it is empty, or it has a shape like that in Figure 14.

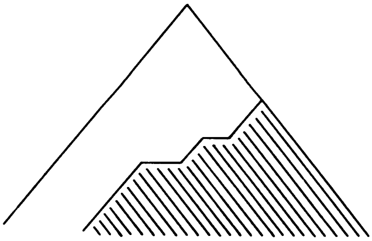

Fig. 14.

Consider the latter case. PLUS forbids irregular kinks in the boundary, and hence the shape must be that of *at least* k, for some $k \geqslant 0$. Actually, $k \geqslant 1$, since otherwise the first infinite row would be all + (by ↑MON), and Q would become countably trivial. But then, the infinite rows of the tree are fully determined through the following observations. At infinite level a,

— all positions (a, s) with $s \geqslant k$ get + (apply ↑MON to $(k, k) \in Q$ and $k \leqslant a, k \leqslant s$),

— the position $(a, k-1)$ gets $-$ (since $(k, k-1)$ has $-$, using PLUS, adding a entities. Notice that the $-$ cannot be put any further toward the right.),

$-$ all positions (a, s) with $s < k-1$ then get $-$ (by MON↑).

Thus, this case becomes *at least k* all the way through.

In addition, the former case with a negative finite tree yields one infinitary quantifier. For, consider the position $(\aleph_0, \aleph_0)$. If it is $-$, then so are all $(n, \aleph_0)$ and $(\aleph_0, n)$ for finite n (by ↑MON), making Q countably trivial. Therefore, $(\aleph_0, \aleph_0)$ must have $+$, and hence so do all nodes (a, b) with $a \geqslant \aleph_0$, $b \geqslant \aleph_0$ (by ↑MON), as well as all nodes towards their right (by MON↑). Finally, a suitable use of PLUS with respect to the $-$nodes (n, n) (n finite) gives $-$ for all nodes (a, n) ($a \geqslant \aleph_0$). Thus, the quantifier becomes *at least infinitely many.*

The other three double monotonicity cases are similar. □

2.9. DISCRETENESS AND CONTINUITY

Of the several uses of quantifiers ignored at the outset, one seems particularly intriguing, viz. the 'continuous' one (*some water, much wine*, . . .). The preceding sections have been devoted to 'discrete' or 'countable' collections, rather than continuous chunks of the world — a neglect which is customary in modern semantics. But the duality between these two perspectives seems a basic fact of human thought, and hence the question arises how far the preceding investigation is tied up with the discrete world view.

Fortunately, the earlier notions and results may be viewed from a more comprehensive standpoint, embracing both discrete and continuous entities. Via the well-known analogy between mass nouns (*water, wine, time*) and collections (*willows, girls*), one may think of general semantic models as 'inclusion structures'

$$\mathfrak{I} = \langle I, \sqsubseteq \rangle;$$

of the form $\langle \mathscr{P}(E), \subseteq \rangle$ in the earlier cases, possibly atomless in the continuous case. These models may carry additional structure. Specifically, inclusion may give rise to a *lattice* structure for join ($\sqcap$) and meet ($\sqcup$), or even a Boolean Algebra with complements. Henceforth, a generalized quantifier will be a functor assigning, to each inclusion structure $\mathfrak{I}$, some binary relation $D_{\mathfrak{I}}$ on its universe I. (We will usually think of $\mathfrak{I}$ as a lattice.)

Earlier general constraints are easily formulated in this new setting. Notably, *Conservativity* becomes

$$D_{\mathfrak{I}}xy \quad \text{iff} \quad D_{\mathfrak{I}}x(y \sqcap x), \qquad \text{for all } x, y \text{ in } I.$$

And indeed, this is equally plausible in the continuous case. *Most wine was drunk last night* means the same as *most wine is wine drunk last night.*

Quantity requires some reformulation, as it cannot be assumed any longer that there are individuals (atoms) underlying our inclusion structure. But in its earlier permutation version, the induced map $\pi[X]$ on subsets X of E may be characterized independently as follows. A permutation π of the power-set of E is induced by an underlying permutation of indviduals if and only if π respects inclusion:

$$X \subseteq Y \quad \text{iff} \quad \pi(X) \subseteq \pi(Y).$$

Proof. Two observations suffice. First, such *inclusion automorphisms* π map singletons (inclusion atoms) to singletons, enabling one to recover an individual permutation. Also, such functions are *continuous*, in the usual mathematical sense of preserving arbitrary unions of subsets:

$$\pi\Big(\bigcup_i X_i\Big) = \bigcup_i \pi(X_i). \ \square$$

Now, Quantity may be given the general formulation that determiner relations be invariant for inclusion automorphisms of $\mathfrak{I}$.

Perhaps surprisingly, the effect of Quantity is more severe in the continuous case than in the discrete one — where it left standard quantifiers such as *all, some,* but also *most,* or *many.*

EXAMPLE. Let $\langle I, \subseteq\rangle$ consist of all open real intervals, with set inclusion. The only conservative, quantitative determiner relations are *all, some* and their Boolean compounds.

That any other possibility is excluded may be seen as follows. Consider an open interval A with a proper open subinterval B of any size. There is always some topological transformation of the reals which is a bijective inclusion automorphism in the above sense, fixing A, while sending B to some pre-assigned portion of A (large or small). So, one can only distinguish the above coarse possibilities. $\square$

Thus, the traditional logical quantifiers can be viewed as the '*topological* invariants'. The others, such as *most, much* or *little*, are only invariant for those automorphisms which also preserve some additional *metric* structure. After all, this is not completely alien to the discrete world picture either, witness earlier discussions of 'qualitative' determiners (Section 1.7.).

Finally, other general intuitions, such as the Graduality of Section 2.4., remain plausible in the continuous realm too. In fact, they may even receive more appealing mathematical formulations there, in the familiar continuity terminology of 'small changes — small effects'.

In general, such principles only unfold their meaning on infinite structures — which is yet another motive for pursuing the topic of infinity after all.

2.10. PLURALITY

The syntactic diversity of quantifier expressions is immense. For instance, numerals (*one, ten, . . .*) may also be classified as adjectives, and other quantifiers (*all, only, . . .*) may occur in adverbial positions. Then also, the compositional mechanism of complexes such as *at least two, at most two* may be studied in finer detail. Most striking among these syntactic phenomena is the ubiquitous connection between determiner expressions and plural forms: *at least two girls, all girls, not many girls*, etc. Should not this play some role in our formal theory?

In general, the semantics of plurality is a mine-field of conflicting intuitions. Nevertheless, it is also an important phenomenon, reflecting our ability as language users to organize the world in collective terms. Moreover, given the treatment of continuous quantification in the above, and the analogies between mass terms and plurals, some sort of extension ought to be feasible. In this final section, we shall steer a conservative course into this dangerous area.

Consider a simple sentence such as *three toddlers were sitting on a fence*. Its abstract form, until now, was taken to be QXY. But of course, in the actual sentence, X, Y are plural forms of singular predicates 'toddler', 'be sitting on a fence'. On general Fregean grounds then, one might wish to bring in a semantic account for this plural formation. A very simple pilot example is this:

$$[\![P_{\text{plural}}]\!] = \text{POW}^{+}([\![P]\!]);$$

where POW^{+} sends a set to the collection of its non-empty subsets. Many variations on this theme are possible; in particular, 'groups' may eventually replace mere subsets.

Thus, quantifiers (and determiners in general) will start expressing binary relations between sets A^+, B^+ (being short for $\mathrm{POW}^+(A)$, $\mathrm{POW}^+(B)$). What are the appropriate truth conditions? One popular strategy is to copy earlier explications:

$$[\![all]\!](A^+, B^+) \quad \text{if} \quad A^+ \subseteq B^+,$$
$$[\![some]\!](A^+, B^+) \quad \text{if} \quad A^+ \cap B^+ \neq \emptyset.$$

This fails, however, for cases like *most*. With $A = \{1, 2, 3\}$, $B = \{1, 2\}$, most A are B, and yet only three out of the seven sets in A^+ belong to B^+. And indeed, there is something counter-intuitive here in letting our quantifiers range over groups rather than individuals. The correct account, therefore, is more reductive:

$$QA^+B^+ \quad \text{if} \quad QA \bigcup_{C \in B^+} \{a \in C \mid a \in A\};$$

where the second quantifier compares A with all its members 'participating' in sets in B^+, in the old style.

At this new level, the earlier constraints are quite easy to state. For instance, Conservativity remains the same. (Eventually, a suitable notion of 'restriction' may have to replace mere intersection, however.) Quantity now becomes invariance for inclusion automorphisms, as in the preceding section. Moreover, special purpose conditions such as Monotonicity remain in force; witness the plurals in, say, $\downarrow$MON$\uparrow$: 'if all toddlers are sitting, then all female toddlers are sitting or standing'. Old results, such as the Double Monotonicity characterization of the Square of Opposition, will then go through.

But of course,there is more to plurality than this uneventful lifting. To proceed, here is another common error in setting up truth conditions for plural terms. It has often been proposed to read numerals as adjectives, producing the following reading for, say, *three toddlers* viewed as a complex noun:

$[\![\textit{toddlers}]\!] = \mathrm{POW}^+([\![\textit{toddler}]\!])$,
$[\![\textit{three toddlers}]\!] =$ all sets of precisely three toddlers.

Then, an example like the original one above would be read as follows, with *three toddlers* now raised to NP-position:

> some element of $[\![\textit{three toddlers}_{\mathrm{N}}]\!]$ belongs to POW^+ ($[\![$*be sitting on a fence*$]\!]$).

This produces a reading *at least three*, rather than exactly three;

because holding for at least one group of three does not exclude holding for more than three toddlers. The problem becomes more acute with *at most three*: holding for some group of at most three toddlers is not at all the same as holding for at most three toddlers. These observations point at the difficulties of an overly collective view of plurals. Nevertheless, they also suggest an important distinction to be made.

There is a way of justifying the rejected account after all. Like singular NP's, plural NP's can have both 'general' and 'specific' (or 'referential') uses. In the singular case, one has such well-known anaphoric differences as that between '*A girl* came in, and *she* . . .' versus the unacceptable '*Every girl* came in, and *she* . . .'. In the plural case, there is a similar distinction, between the admissible '*Some girls* came in, and *they* . . .' versus the unacceptable '*Most girls* came in, and *they* . . .'. One possible explanation here is that certain noun phrases, due to some special mathematical structure of their denotations, admit this kind of anaphora, whereas others do not. But, it is simpler to assume that *girls* has a definite contextual reference here, with the whole sentence making a statement about that contextually given set. Accordingly, the proper anaphoric relation would be as indicated: 'Some *girls* came in, and *they* . . .'. The latter 'referential' readings may be at the back of the earlier-mentioned existential truth clauses. (A similar account can be given for *all, most*, etc.)

This point of view raises interesting questions concerning the semantics of the latter readings, and their connection with our standard approach. Notice, e.g., that on the referential reading, earlier forms of monotonicity may fail: *at most three babies cried*, taken in the latter sense, need not imply *at most three healthy babies cried*. Another theme which comes to the fore here is the semantic role of bare plurals and *the*-phrases.

The latter issues lead one naturally to the topic of *collective* predication. Obviously, not all VP-denotations can be taken to be of the form $\mathrm{POW}^+(A)$ for some set of individuals A. Some predicates are themselves collective (*gather, quarrel*), others become so, either implicitly (*weigh three tons*) or explicitly (*hated each other*). Our previous discussion has been restricted to so-called *distributive* predicates P of sets, satisfying the reduction

$$\text{for all } A,\ A \in P \quad \text{iff} \quad \{a\} \in P \text{ for all } a \in A.$$

And then, $\{a\} \in P$ amounts to having $a \in P^-$, for some individualized version P^- of P. Put differently, distributive predicates satisfy the following two conditions:

— if $A \in P$, $B \subseteq A$, then $B \in P$ (*Heredity*)
— if A, $B \in P$, then $A \cup B \in P$ (*Homogeneity*)

(Mathematically, they are 'ideals of sets'.) These conditions are ubiquitous in the literature on plurals, as well as many areas of 'partial logic' (cf. van Benthem, 1985a).

Other kinds of collective predication may satisfy just one of these two conditions, or none at all. Evidently, all this deserves thorough exploration. But, are quantifiers in our sense essentially involved here? For instance, many of them do not go well with collective predication: *all boys lifted the piano* (?), *at most two girls quarrelled* (?). Such expressions seem felicitous only (if at all) in the above referential readings. In view of all these uncertainties, we have refrained from 'collectivizing' the theory of quantifiers and determiners in this book — even though there seem to be no difficulties a priori in doing so.

ALL CATEGORIES

Our investigation up till now has been concerned with the linguistic category of determiners, including quantifiers. But, the techniques developed can be brought to bear upon arbitrary grammatical types: adjectives, connectives, adverbs, etc. To broaden our scope, we will work against the background of an extensional *categorial grammar*, with basic types e ('entity') and t ('truth value') — as in Section 1.1. — allowing formation of functional types (a, b) ('from a-denotations to b-denotations'). Models will then be semantic structures $(D_a)_{a \in \text{TYPE}}$, with base domains D_e, D_t and a recursive construction rule $D_{(a, b)} = D_b{}^{D_a}$. Chapter 7 is devoted to further theoretical study of this mechanism as such.

As was observed in Chapter 1, some categories of expression a seem to be 'interpretatively free', allowing arbitrary denotations in their associated domains D_a — whereas others are more constrained. Naturally, we shall be especially interested in the latter, searching for enlightening constraints on denotations. In principle, these may be quite different from those encountered for determiners — but, there is certainly no harm in looking for relatives of the earlier basic conditions across all categories. Before proceeding to the most general case, we consider some specific examples.

3.1. CONNECTIVES

One basic example to start from is that of connectives, the other main type of traditional logical constant. In natural language, the usual Boolean connectives occur primarily as operators on predicates, with less frequent inter-sentential uses. (Eventually, they may be viewed as operations on all types resulting in a truth value; cf. Keenan and Stavi, 1982.) For instance, the former use underlies the very linguistic formulation of Conservativity:

$$QXY \quad \text{iff} \quad QX(Y\,and\,X).$$

Likewise, Monotonicity inferences involved predicate conjunction and disjunction.

Earlier notions may be transferred to this new category. Thus, *Quantity*, the earlier hallmark of logicality, now acquires the obvious sense that

> for all permutations π of E, and all $A \subseteq E$,
> $f(\pi[A]) = \pi[f(A)]$,
> for unary connectives f.
> And likewise for higher ones.

I.e., quantitative connectives commute with permutations. The force of this requirement is expressed in the following

EXAMPLE. Every Boolean expression in A, B, $\cup$, $\cap$, $-$ (with complement taken with respect to the universe E) defines a quantitative connective. This is so because permutations π commute with unions, intersections and complements:

$$\pi[C \cup D] = \pi[C] \cup \pi[D],\ \pi[C \cap D] = \pi[C] \cap \pi[D],$$
$$\pi[E - C] = E - \pi[C].$$

But also conversely, every quantitative connective has a Boolean definition for each tuple of arguments. For, consider $A, B \subseteq E$. For each of the four zones in the Venn diagram: $A \cap B$, $A - B$, $B - A$ and $E - (A \cup B)$, $f(A, B)$ either contains it or avoids it altogether. (For instance, if $f(A, B)$ were to non-trivially intersect $A \cap B$, one could permute two individuals in $A \cap B$ across the boundary — leaving all other objects in E fixed — which would leave A, B invariant, while changing $f(A, B)$: a contradiction.) So, $f(A, B)$ must be some disjunction of the above four sets: i.e., a Boolean compound of A, B. □

Thus, Quantity again expresses an entirely natural form of logicality in this setting, linking quantifiers to truth-functional connectives. Evidently, the above argument goes through for arbitrary k-ary connectives as well. Still, it does not enforce one uniform definition for all argument tuples. To get the latter, postulates are needed relating values at different arguments. One example will be presented below.

The other general conditions on quantifiers are less plausible in this case. Notably, the complement operation (*not*) is universe-dependent; whence it fails to satisfy *Extension*. Likewise, *Conservativity* has no direct appeal here. But then, new conditions in the same spirit may be forthcoming for special categories. For instance, for many operations C, a principle of *Restriction* is plausible:

$$\begin{aligned}&\text{if } E \subseteq E', \text{then, for } A_1, \ldots, A_k \subseteq E',\\ &C_E(A_1 \cap E, \ldots, A_k \cap E) = C_{E'}(A_1, \ldots, A_k) \cap E \quad \text{(REST)}\end{aligned}$$

This condition holds for all Boolean set connectives; and that is no coincidence:

THEOREM: QUANT, REST characterize the Boolean operations uniformly.

Proof. By a judicious use of Restriction, arbitrary values for f may be written down in terms of its 'E, $\emptyset$-truth table'. For instance,

$$\begin{aligned}f_E(A, B) = {}&((A - B) \cap f_E(E, \emptyset)) \cup ((A \cap B) \cap f_E(E, E)) \cup\\ &((B - A) \cap f_E(\emptyset, E)) \cup ((E - (A \cup B)) \cap f_E(\emptyset, \emptyset)). \quad \square\end{aligned}$$

In addition, various special purpose conditions arise here. One useful property, already encountered in Section 2.9., is the well-known mathematical condition of *Continuity*, in the sense of preserving unions:

$$\begin{aligned}&\text{for all families } \{A_i \mid i \in I\} \text{ of subsets of } E,\\ &f_E\Big(\bigcup_i A_i\Big) = \bigcup_i f_E(A_i).\end{aligned}$$

Actually, Continuity makes most sense when applied to unary adverbs and adjectives, rather than connectives (see Section 3.2. below).

Moreover, earlier family patterns also emerge in this new setting. For instance, there are *Squares of Opposition* for connectives, of which the following is a simple illustration.

EXAMPLE. Binary truth functions.

In the area of truth tables, connectives f may be regarded equivalently as relations between subsets of some set $\{x\}$, being $\emptyset$ (0) and $\{x\}$ itself (1). *Conservativity* then means that $f(0, 1) = f(0, 0)$. *Variety* requires that $f(1, 0) \neq f(1, 1)$. Now, inspection of truth tables shows that, modulo these two conditions (cf. Section 1.4.):

the doubly monotone connectives form precisely the Square

$$\begin{array}{ccc} p \to q & & p \to \neg q \\ p \wedge q & & p \wedge \neg q. \end{array}$$

For instance, suppose that f has type ↓MON↑. Case (1): $f(0, 1) = f(0, 0) = 1$. Then $f(1, 0) = 0$. (Otherwise, $f(1, 0) = 1$, $f(1, 1) = 1$ (MON↑), contradicting VAR.) So, again by VAR, $f(1, 1) = 1$. I.e., f is material implication. Case (2): $f(0, 1) = f(0, 0) = 0$. As before, $f(1, 0) = 0$, $f(1, 1) = 1$. But then, by ↓MON, $f(0, 1) = 1$: a contradiction. So, this case cannot occur. □

A much more general survey of such 'Squares' across natural languages may be found in Löbner (1984).

In addition to such denotational constraints, the earlier inferential concerns still apply in this setting (cf. Sections 1.5., 2.6.). We may classify common algebraic conditions on connectives (idempotence, commutativity, associativity, and the like) as patterns of inference. And then, the earlier question may be raised which classes of connectives are determined by their characteristic patterns of reasoning. Here is one illustration.

Query: Which triples of QUANT operations on sets satisfy the complete set of Boolean identities in $\neg$, $\wedge$, $\vee$?

Even the earlier observed phenomenon of 'multiple solutions' has been known for a long time in this particular area, as the 'duality' of $\wedge$, $\vee$. Here is one answer for a special case.

THEOREM. In propositional truth value semantics, the complete set of Boolean identities has exactly two solutions:

$$\neg, \wedge, \vee \quad \text{and} \quad \neg, \vee, \wedge.$$

Proof. First, there are at most these solutions. Various valid identities narrow down the range of possible candidates:

$\neg\neg X = X$: $\neg$ can only be identity or value reversal.
$X \wedge X = X$: $\wedge(0, 0) = 0$, $\wedge(1, 1) = 1$.
$X \wedge Y = Y \wedge X$: $\wedge(0, 1) = \wedge(1, 0)$.
$(X \wedge \neg X) \wedge Y = X \wedge \neg X$: if $\neg$ were the identity, then $X \wedge Y = X$ would be valid — and a contradiction arises: choose $X = 0$, $Y = 1$: $\wedge(0, 1) = 0$/choose $X = 1$, $Y = 0$: $\wedge(1, 0) = 1$. Therefore, $\neg$ denotes value reversal.

For disjunction, similar observations imply $\vee(0, 0) = 0$, $\vee(1, 1) = 1$, $\vee(0, 1) = \vee(1, 0)$. Moreover, because of the mixed principle $X \wedge (Y \vee X) = X$: $\vee(0, 1) \neq \wedge(0, 1)$. (Otherwise, if $\vee(0, 1) = \wedge(0, 1) = 0$, then set $X = 1$, $Y = 0$: $\wedge(1, \vee(0, 1)) = \wedge(1, 0) = 0 \neq 1$: the X-value.

Or, if $\vee(0, 1) = \wedge(0, 1) = 1$, setting $X = 0$, $Y = 1$ leads to a similar problem.) In all, this leaves just two possibilities for $\wedge$, $\vee$.

That there are at least these two solutions follows by any one of the familiar duality arguments. □

In this inferential perspective, the two main kinds of logical constant may also be combined. This may cut down the number of possibilities on either side, as 'mixed inferences' form additional constraints. Thus, we have a question (cf. Section 2.7.):

modulo Quantity, does the logic of all, some, or, and, not *determine the interpretation of these logical constants uniquely*?

Appendix: Quantifiers Revisited

Within a general categorial framework, quantifiers may also be regarded, not primarily as determiners or noun phrases, but as *reducers of argument places*. Thus, e.g., the phrase *someone* turns the binary relation *loves* into the unary property *love someone*. This point of view is the basis of predicate logic as developed in Quine (1966). Its major operations on predicates are 'permutation', 'identification' and 'projection'; of which the latter two are relevant here:

$$id(R) =_{\text{def}} \{x \mid \langle x, x\rangle \in R\}, \quad proj(R) =_{\text{def}} \{x \mid \exists y \langle x, y\rangle \in R\}.$$

There is a connection here with the linguistic phenomenon of 'argument drop', as displayed in the following examples. *Mary washes the clothes* can go to *Mary washes* [something] ('right projection'), to [someone's] *washing the clothes* ('left projection'), or to *Mary washes herself* ('reflexivization'). In a sense, the following argument explains this scarcity of argument drop mechanisms.

Let us restrict attention to relation-reducing functions sending binary relations to subsets of their domains. The above examples satisfy Quantity as a general postulate. Moreover, their special distinguishing feature turns out to be an earlier special purpose condition:

THEOREM. The Quine operations id, proj are essentially the only relation-reducing functions satisfying Quantity and Continuity.

Proof. Let E be any universe, with a binary relation R on it. Any function f satisfying the above two conditions will map R to $\bigcup\{f(\{\langle d, e\rangle\}) \mid \langle d, e\rangle \in R\}$. So, it is to be determined what can be assigned in these singleton cases.

— $\langle d, d\rangle \in R$: f assigns either $\emptyset$ (1), or $\{d\}$ (2).

By Quantity, if case 2 occurs for any object d, it will occur for all $d' \in E$. (Consider a permutation interchanging d, d', while leaving all other objects fixed.)

— $\langle d, e\rangle \in R$, $d \neq e$: f assigns either $\emptyset$ (3), or $\{d\}$ (4).

Again by Quantity, case 4 either occurs for all couples, or for none.

Summing up, there are four possible cases:

case 1 + 3: $f(R) = \emptyset$, case 1 + 4: $f(R) = \mathrm{proj}(R) - \mathrm{id}(R)$,
case 2 + 3: $f(R) = \mathrm{id}(R)$, case 2 + 4: $f(R) = \mathrm{proj}(R)$. □

This second view of quantifiers is also relevant when determiners are considered, not in subject, but in direct *object* position (*Mary loved every lamb*). In the latter position, quantifiers stand for functions on sets and binary relations, yielding sets of individuals. For instance,

$$[\![\mathit{every}]\!]([\![\mathit{love}]\!], [\![\mathit{lamb}]\!]) = \{e \in E \mid [\![\mathit{lamb}]\!] \subseteq \{d \mid \langle e, d\rangle \in [\![\mathit{love}]\!]\}\}.$$

In a relational perspective, there is a ternary relation involved here, between an individual, a binary relation and a property: $Q_E(e, R, A)$. What we want, then, is a reduction to the original treatment of Chapter 2. One strategy is to use a principle of *Locality* (with $R_e =_{\mathrm{def}} \{d \mid \langle e, d\rangle \in R\}$):

$$Q_E(e, R, A) \quad \text{iff} \quad Q_E(e', R', A), \quad \text{whenever} \quad R_e = R'_{e'}.$$

Then, given suitable versions of QUANT, EXT, CONS, the binary relation $\{(A, R_e) \mid Q_E(e, R, A)\}$ can be treated exactly like before. This ad-hoc move becomes justified in Chapter 7, as an instance of a general type shift rule inducing a canonical change of meaning.

3.2. GENERALIZATIONS

Once certain analogies have been observed between different categories, a bolder leap becomes possible, to obtain denotational conditions that make sense in all categories. The prime example is again *Quantity*:

Let π be any permutation of the universe $E = D_e$. Setting π equal to the identity on the other base domain D_t, this function may be lifted recursively to all domains D_a by stipulating, for $f \in D_{(a, b)}$:

$$\pi(f) = \{\langle \pi(x), \pi(y)\rangle \mid \langle x, y\rangle \in f\}.$$

An object x in the domain of any category will now be said to satisfy Quantity if, for all such permutations π, $\pi(x) = x$.

For determiners and connectives, this produces the earlier formulations. But now this logicality constraint also has effects in other categories. Indeed, some categories, such as D_e itself, are entirely without logical items. (How does this tie in with their 'interpretative freedom'?) In others, only very few qualify. Thus, among binary relations (type $(e, (e, t)))$, Quantity leaves only *identity* and its Boolean compounds as logical constants. Likewise, in type $(e, ((e, t), t))$, only *elementhood* ($\in$) and its variants remain. All this is exactly as it should be.

The preceding examples illustrate the effect of Quantity in various special categories. In fact, one can find explicit counting formulas for the logical items in all these cases. One obvious general question then is to find a *universal counting formula*, which tells us, for each type a (and universe of size n) how many logical items occur in the domain D_a. For instance, exactly which categories have only logical items? And which categories lack logical items?

Next, one can speculate a bit about the other basic constraints. *Extension* becomes less plausible now, because of context-dependent items such as negation. E.g., when $A \subseteq E \subseteq E'$, *not-A* will not denote the same predicate in E as in E'. Still, there remains an obvious stability, in that, on individuals in E, the new set makes the same decisions as the old one. Formally, then, we have again a principle of *Restriction*:

$$\begin{array}{ll} \text{if} & A_1, \ldots, A_k \subseteq E \subseteq E', \\ \text{then} & f_{E'}(A_1, \ldots, A_k) \cap E = f_E(A_1, \ldots, A_k). \end{array}$$

If this is to be stated in a form suitable for all categories, intersection will have to be replaced by a suitable notion of *restriction*. Generally, each choice of a universe $E = D_e$ generates a model structure $(D_a(E))_a$; which changes to $(D_a(E'))_a$ when $E \subseteq E'$. Note that not all levels $D_a(E)$ need be included in $D_a(E')$: there are some well-known pitfalls here. The proper view is rather that items in $D_a(E')$ can be restricted to items in $D_a(E)$. We shall not pursue the technicalities of this notion here. But its intuitive intent is clear.

The above condition also suggests a stronger version, already employed in Section 3.1.:

$$\begin{array}{ll} \text{if} & E, A_1, \ldots, A_k \subseteq E', \\ \text{then} & f_E(A_1 \cap E, \ldots, A_k \cap E) = f_{E'}(A_1, \ldots, A_k) \cap E. \end{array}$$

The latter holds of Boolean connectives; but not in general of all determiners (now viewed as functions). In the latter area, it implies that $D_E(A \cap E)(B \cap E)$ iff $D_{E'}AB$: whence D would be downward persistent.

Incidentally, the present *functional* type theory may not always be the most advantageous setting for our study. For an equivalent, but often more convenient *relational* format, cf. van Benthem (1983a).

Finally, *Conservativity* would express in general that the left-hand argument of our functions always plays some special 'covering' role. As a rule, this need not be the case. But there are certainly many instances of this phenomenon. Our final example illustrates this.

EXAMPLE. Recall the earlier restriction with respect to intersective adjectives (Section 1.1.):

$$f_E(B) = B \cap A, \quad \text{for some fixed set } A \subseteq E.$$

Here, A is the E-extension of the adjective — and, if the latter is suitably 'absolute', f_E will satisfy the strong version of the *Restriction* principle. Moreover, it is *introspective*, in the sense that $f_E(B) \subseteq B$ for all arguments B. Together, these two properties are characteristic for intersective adjectives. Each introspective strongly restrictive operation on E can be represented as an adjectival restriction to some suitable set A, viz.

$$A =_{\text{def}} \bigcup_{X \subseteq E} f_E(X).$$

Claim. $f_E(B) = B \cap A$, for all $B \subseteq E$.

Proof. '$\subseteq$': $f_E(B) = B \cap f_E(B) \subseteq B \cap A$. '$\supseteq$': for arbitrary $X \subseteq E$, $f_E(X) \cap B = f_{X \cap B}(X \cap B)$ (by Restriction) $= f_E(B) \cap X \cap B$ (again by Restriction). Hence $f_E(X) \cap B \subseteq f_E(B)$, $B \cap A = B \cap \bigcup f_E(X) = \bigcup (B \cap f_E(X)) \subseteq f_E(B)$. □

This example will revisited in Section 3.3.

Finally, earlier special purpose conditions can be generalized too. One prominent example is *Monotonicity*. First, introduce a suitable notion $\sqsubseteq$ of *inclusion* in all types as follows:

on D_t, $\sqsubseteq$ is $\leqslant$; on D_e, it is the identity $=$;
on $D_{(a,b)}$, $f \sqsubseteq g$ if, for all $x \in D_a$, $f(x) \sqsubseteq g(x)$.

This coincides with set inclusion for $D_{(e,t)}$. In general, $\sqsubseteq$ may be thought of as a universal relation of *logical consequence* in all categories. With respect to it, functions can be 'directly' or 'inversely' monotone in their argument in obvious ways. (These will be the denotations of expressions which 'respect logical consequence', one way or another.)

As before, one special case deserves attention. Notice first that all functional types can be written in one of two forms: $(a_1, (a_2, \ldots (a_n, e) \ldots))$ or $(a_1, (a_2, \ldots (a_n, t) \ldots))$ — i.e., as functions of n arguments going to either entities or truth values. Then, such functions can be monotone in all these arguments. For instance, special logical interest would attach, in each category, to the *quantitative multimonotone* functions. Can we extend our previous characterization theorems for doubly monotone determiners to this general case?

3.3. CROSS-CATEGORIAL CONNECTIONS

There is another reason for investigating categorial structure beyond one single type. Up till now, we have considered conditions valid for one category, and their possible generalizations elsewhere. But, there is also a different type of connection, even between categories with quite different kinds of basic constraints. As categories are intertwined, semantic constraints do not live in isolation: they may interact.

Propagation of constraints. For instance, consider the basic triad of Chapter 1: noun $((e, t))$, noun phrase $(((e, t), t))$ and determiner $(((e, t), ((e, t), t)))$. As has been observed in Chapter 2, the first seems to be interpretatively free, whereas the third is subject to Conservativity. But then, ipso facto, the second may have been constrained already: if it turns out that not all possible noun phrase denotations can be obtained in the form DX for some conservative determiner D and arbitrary noun X. In this particular case, Keenan and Moss (1985) point out that, in fact, Conservativity is still quite generous: every noun phrase denotation is thus obtainable. But Conservativity in combination with Quantity becomes restrictive: not all possible noun phrase denotations can be obtained in this way, as we shall see below.

In this perspective, various questions arise. For instance, fixing one constraint on a functional category (say, Conservativity), one may study its 'transmission' behaviour. For various semantic conditions on its argument category, what are the resulting restrictions (if any) on its value category? And in any case, will such restrictions tend to 'die out' for categories which lie more application steps away? Another example is that of our earlier universal constraints, applicable to all categories, such as Quantity. Ought not these to be 'self-propagating', in the sense that, if a function and its argument obey the constraint already, then so do its values? (For Quantity, the answer happens to be affirmative.)

Here is a concrete example illustrating the original case. We shall characterize those noun phrase denotations X which can be obtained in 'Det N' form for some logical determiner D (or Q). First, define

$$c(X) =_{\text{def}} \bigcap \{U \mid \text{for all } Y,\ Y \in X \text{ iff } Y \cap U \in X\}.$$

If U_1, U_2 have this 'witness' property, then so does their intersection $U_1 \cap U_2$ (cf. Thijsse, 1985). For finite domains then, the above definition yields a smallest witness set for X. Next, define

$$c^*(X) =_{\text{def}} \{U \in X \mid U \text{ is contained in } c(X)\}.$$

Now, the desired characterization is this.

THEOREM. X is logically det N-representable if and only if

$$c^*(X) \text{ is closed under permutations of } c(X).$$

Proof. First, suppose that $X = \{B \mid DAB\}$ for some suitable D, and π is some permutation of $c(X)$. Now, $c(X) \subseteq A$, by Conservativity of D and the definition of c. Also, π can be extended to a permutation π^+ on the whole universe by adding the identity outside of $c(X)$. Then, if $U \in c^*(X)$ – i.e., $U \in X$, $U \subseteq c(X)$ – we have DAU, $D\pi^+[A]\pi^+[U]$ (by Quantity for D), whence $DA\pi[U]$, $\pi[U] \in X$, $\pi[U] \in c^*(X)$.

Conversely, let $c^*(X)$ have the stated closure property. Then, set $A = c(X)$, and define a determiner D as

$$\{(\pi[A], \pi[B]) \mid \pi \text{ is some permutation of the universe, and } B \in X\}.$$

It may be checked that D is conservative and quantitative. Moreover, the closure property guarantees that $\{B \mid DAB\}$ remains just X. □

This result gives a criterion for checking whether a given set of sets X is a possible denotation of the above form.

This general perspective on the contents of categories invites many further questions. For instance, a notable empirical phenomenon in natural languages is the limited range of types occupied by actual expressions. The following table illustrates this:

proper names: e,
common nouns: (e, t)
transitive verbs: $(e, (e, t))$
adjectives: $((e, t), (e, t))$
determiners: $((e, t), ((e, t), t))$
modifiers: $(((e, t), (e, t)), ((e, t), (e, t)))$
sentences: t,
intransitive verbs: (e, t)
adverbs: $((e, t), (e, t))$
noun phrases: $((e, t), t)$
prepositions: $(((e, t), t), ((e, t), (e, t)))$

This list is not exhaustive; but it does illustrate the kind of categories to be expected.

One measure of semantic complexity here is the notion of *order*, assigning natural numbers $o(a)$ to types a as follows:

$$o(e) = o(t) = 0$$
$$o((a, b)) = \max(o(a) + 1, o(b)).$$

The above expressions have orders 0, 0, 1, 1, 1, 2, 2, 2, 2, 3, 3, respectively. Various observers have noted that third-order seems to be a threshold for natural languages. There may be a syntactic explanation of this phenomenon, having to do with the complexity of type assignments in categorial grammars needed for linguistic description (cf. Buszkowski, 1982). But it would also be of interest to find some more semantic explanation — perhaps in terms of some expressive completeness result for these lower types with respect to the higher ones.

The frequent occurrence of denotational restrictions may also indicate the need for a change in the usual model theory for our type system. Instead of the full function domains $D_{(a, b)}$, certain smaller sets might suffice, as long as these contain all relevant maps. Technically, then, the background model could be any *Cartesian-closed category* (this time, in the sense of mathematical 'Category Theory'!); perhaps provided with additional *products* and *sums* of type domains.

Boolean structures. A pioneering approach toward charting this full categorial area is that of Keenan and Faltz (1985). These authors emphasize the *Boolean Algebra* structure found in many domains D_a; viz. in all those whose type ends in a 'final' t. (Incidentally, they also make an interesting case for having a different basic type structure, starting from p $(=(e, t))$; t rather than e; t.) This viewpoint does not necessarily lead outside of the full function domains — but it certainly suggests admitting arbitrary Boolean algebras as well.

The Boolean structure of many categories manifests itself in the relative freedom with which *not, and, or* can attach to the most diverse expressions. It is also involved in many of the preceding topics. For instance, the earlier general inclusion relation $\sqsubseteq$ is nothing but the basic Boolean inclusion $\leqslant$ on domains which are Boolean algebras. One Boolean notion of immediate interest is that of a $\leqslant$-*automorphism* on (suitable) domains D_a. The most robust items in a category, from this perspective, would be those objects in D_a which are fixed points of all its $\leqslant$-automorphisms. As has been observed before (cf. Section 2.9.), such automorphisms on the special domain $D_{(e,t)}$ correspond exactly to underlying permutations of the individual domain D_e. And in general, all permutations π as defined in Section 3.2. induce $\leqslant$-automorphisms. But the converse fails: e.g., there are $\leqslant$-automorphisms in $D_{((e,t),t)}$ which are not generated by any underlying permutation of individuals. In fact, Boolean invariance soon becomes an excessive strengthening of our earlier notion of logicality.

One central topic in this area is the variety of *morphisms* connecting algebras D_a, D_b (whether or not occurring 'encoded' in $D_{(a,b)}$). As Keenan and Faltz point out, these occur in various mathematical sorts. Notably, there are *Boolean homomorphisms* preserving the Boolean operations — such as proper names in the NP category, satisfying the equivalences

Olga doesn't drink $\leftrightarrow$ *not Olga drinks*
Olga drinks and smokes $\leftrightarrow$ *Olga drinks and Olga smokes.*

Another major example are transitive verbs:

kill not every wasp $\leftrightarrow$ *not kill every wasp*
kill three flies or one mosquito $\leftrightarrow$ *kill three flies or kill one mosquito.*

Other expressions show only part of this behaviour. Notably, certain adjectives and prepositions are *continuous* items, respecting arbitrary disjunctions:

with a pen or a knife $\leftrightarrow$ *with a pen or with a knife*
wicked (mother or daughter) $\leftrightarrow$ *wicked mother or wicked daughter.*

Finally, some expressions may only preserve inclusion $\sqsubseteq$; or even lack mathematical extras altogether.

Keenan and Faltz exploit all these structures for the purpose of linguistic subcategorization, as well as the study of structural analogies across various types of expression. It remains to be seen, however, how much of all this mathematical structure is real, and how much is an artefact of our description.

EXAMPLE. Homomorphisms reduced.

Let f be a homomorphism in a category $((a, t), (b, t))$. As f respects arbitrary unions, its image for any X in $D_{(a, t)}$ (viewed, for convenience, as a subset of D_a) is equal to $\bigcup_{x \in X} f(\{x\})$. Now, since homomorphisms send unit elements to unit elements, $f(D_a) = D_b$. So, $D_b = \bigcup_{x \in D_a} f(\{x\})$. Moreover, as all these atoms are disjoint (another relation preserved by homomorphisms), their f-images will be disjoint subsets of D_b. (Some $f(\{x\})$ may be empty.) So, a reverse function $\bar{f}$: $D_b \to D_a$ can be defined unequivocally, by setting

$$\bar{f}(y) \text{ is that } x \text{ in } D_a \text{ with } y \in f(\{x\}).$$

Then, in general, the following reduction obtains, for all $X \subseteq D_a$,

$$f(X) = \{y \in D_b \mid \bar{f}(y) \in X\} \quad (= \bar{f}^{-1}[X]).$$

So, f can be retrieved from $\bar{f}$. Conversely, any function g from D_b to D_a will induce a unique homomorphism $\tilde{g}$ from $D_{(a, t)}$ to $D_{(b, t)}$ with this scheme. Therefore,

> *the homomorphisms in $D_{((a, t), (b, t))}$ correspond exactly to arbitrary functions in $D_{(b, a)}$.*

As a special case, homomorphisms in $D_{((a, t), t)}$ correspond to arbitrary objects in D_a.

This reduction makes sense, e.g., for transitive verbs as treated by Keenan and Faltz, in type $((e, t), t), (e, t)$. Applying the above perspective, these are 'really' of type $(e, (e, t))$ — which is indeed their intuitive habitat. Conversely, let g be the verb *see* in the latter type. The corresponding function $\tilde{g}$ in $((e, t), t), (e, t)$ will assign, to each NP-type object X, the set of all individuals y in D_e whose $g(y)$ belongs to X. This will produce the correct meaning at the higher level, provided that we read $g(y)$ as the set of all individuals *seen by y*. There is also a connection here with the earlier-mentioned type shift for determiners in direct object position, which will not be spelled out.

On the other hand, there are several categories which in principle could harbour Boolean homomorphisms, but in fact do not. The determiner category $D_{(e, t), ((e, t), t)}$ is a notable example: there seem to be no non-trivial homomorphic determiners. (In the light of the above, these ought to be derived from 'choice functions' in $D_{(e, t), e}$.) Is there any explanation for this phenomenon, here and elsewhere?

Keenan and Faltz themselves treat another interesting case of 'type reduction' (cf. Section 3.2.).

EXAMPLE. Intersective adjectives.

Expressions such as *blond, wicked* have the following two formal characteristics: (i) they respect unions, as noted above, and (ii) they are 'introspective': e.g., *blond beast* → *beast, wicked queen* → *queen.* Here again, a reduction is possible, to an underlying 'absolute' notion. Let f in $D_{(a, t), (a, t)}$ have these two properties. Define $F \subseteq D_a$ as $\{x \in D_a \mid f(\{x\}) = \{x\}\}$. By a simple calculation, then,

$$f(X) = X \cap F, \text{ for all } X \subseteq D_a.$$

This explains how such adjectives can also serve as (and indeed are) properties of individuals.

The principles behind these examples will be discussed at the end of this section. For the moment we present one final cross-categorial topic, due to Partee (1985), another source of examples in the present general spirit.

Type change. Consider the triad of categories e (proper names), (e, t) (nouns) and $((e, t), t)$ (noun phrases). Partee observes that various natural operations run between these categories. For instance, individuals can become noun phrases via Montague's celebrated lifting rule

$$x \mapsto \lambda Y \cdot Y(x) \qquad \text{(M);}$$

or, they can become predicates via their singletons

$$x \mapsto \lambda y \cdot y = x \qquad \text{(I).}$$

Then, it may be asked which special operation from NP-denotations to N-denotations makes the diagram of Figure 15 commute (see Figure 15).

One natural answer is this:

there is a unique Boolean homomorphism from $D_{((e, t), t)}$ to $D_{(e, t)}$ making the *M, I*-diagram commute, viz. Montague's well-known explication for the verb *be*:

$$\lambda Y \cdot \lambda x \cdot \{x\} \in Y.$$

Incidentally, this rule is also what would be obtained by lifting ordinary equality (type $(e, (e, t))$) to the transitive verb type as indicated in the example of homomorphism reduction.

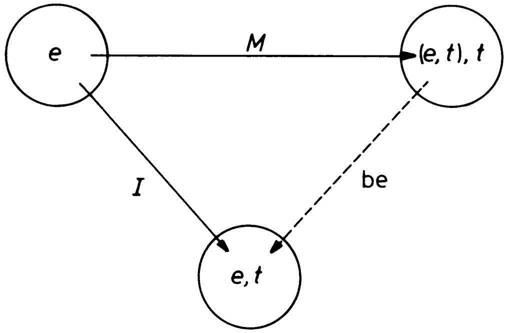

Fig. 15.

Going in the opposite direction, Partee has also asked for linguistically natural operators (say, syncategorematically encoded) sending N-denotations to NP-denotations. For this issue, our earlier perspective becomes relevant. The most obvious way of getting from (e, t) to $((e, t), t)$ is by prefixing a determiner. Which determiners D are reasonable here? In general, given any set of individuals A, D will produce a set of sets $\{B \mid DAB\}$ — in which process the original A may get lost. But, there is a canonical procedure for recovering something like A: viz. our earlier formation of a smallest witness set $c(\mathcal{X})$, for any set of sets $\mathcal{X}$. Thus, we might restrict attention to the determiners D 'encoding' their arguments in the following sense:

$$\text{for all } A,\ c(D(A)) = A \qquad (+)$$

Is this a significant restriction? As it happens, many candidates still pass this test — and additional considerations are needed to arrive at Partee's own proposed solution (being *a* and *the*). But at least, one suggestive result is worth stating:

THEOREM. The quantifiers satisfying (+) are precisely those obeying Conservativity and Variety.

Proof. First assume (+). Conservativity follows at once. Also, if A is non-empty, then let A' be an arbitrary proper subset of A. Since A' is not a witness set for $\{B \mid DAB\}$, there must be some B such that DAB (iff $DA(B \cap A)$) not: iff $DA(B \cap A')$: and hence $B \cap A$, $B \cap A'$ are two sets as required for Variety.

Conversely, assume these two postulates. Let $A \neq \emptyset$. Assume also

DAA. (The other case is similar.) I.e., $(0, n) \in D$, where $n = |A|$. Now, let A' be any proper subset of A, say with size $m < n$. Moreover, by Variety, let k be the maximal size of a subset B of A without DAB. We compare m and k. If $k \geqslant m$, then some B of size k contains A': and we have DAA, not DAB, yet $A \cap A' = B \cap A' = A'$: and A' cannot be a witness set. If $k < m$, then some B of size k lies within A'. Let $B^+ = B$ united with some element of $A - A'$. By the maximality of k, we have DAB^+ — although still not DAB. But, $B \cap A' = B^+ \cap A' = B$: and again, A' is not a witness set. □

These various examples point at general *type-change mechanisms* in natural language. Many expressions do not stay within one single category: they can travel, within certain constraints. For instance, in the Partee example, the Montagovian transition for proper names occurred, from e to $((e, t), t)$. Essentially, the same rule is invoked by Keenan and Faltz, when arguing for the adequacy of a basic type structure p, t. This transition was only a special case of the rule encountered in 'homomorphic inflation', which went from (b, a) to (a, t), (b, t). This again may be viewed as an instance of the 'Geach Rule', used already in Section 1.1. to account for the various uses of negation: sentential, predicative, and otherwise. The mechanics of this kind of type change will be studied in Chapter 7. In particular also, the associated 'recipes' will be generated for obtaining the 'lifted' meanings in the new types. As may be guessed from the preceding examples, these will come in the form of *lambda/application*-terms in a type-theoretical language. Given the concerns of this chapter, it will be of special interest to trace the metamorphoses of our denotational constraints through such type jumps.

The preceding examples exhibit a 'combinatorial' kind of type change, where expressions do not 'really' change their meaning: they merely adapt opportunistically to their environment. More substantial 'mathematical' examples are the earlier cases of individuals going to their singletons, or adjectives reducing to properties. (Another example would be the extraction of a predicate from an NP denotation which is a principal filter generated by that predicate.) In a first approximation, these would be the transitions whose associated recipes employ *identity*, in addition to lambdas and application. (Thus, the type-theoretical language acquires the full power of higher-order logic; cf. Gallin, 1975.) Finally, there are also type changes with more conceptual content than the above, such as nominalizations turning predicates into individuals. The latter phenomenon will not be investigated here.

The three kinds of type change thus identified share one common feature: one single expression adapts by itself to various linguistic contexts. But, the Partee example also introduced *lexicalized* type changers, such as determiners changing predicates into NP denotations in a special formal way. Now, in a sense, every functional expression is a lexicalized type changer, when applied to expressions of its argument type. But in general, this is only one, combinatorial aspect of its meaning. There are quite a few modest, but crucial expressions, however, whose *sole* purpose seems to be to effect type changes beyond what would be permitted by the 'self-changing' mechanisms developed in Chapter 7. Examples are the reflexivizer *-self*, turning binary predicates into unary properties, or the lambda-extractor *such that*, turning sentences into predicates (cf. Keenan and Faltz, 1985). A systematic study of such particles would be very useful.

With these themes, we conclude our survey of a general categorial approach to the study of admissible denotations for linguistic expressions.

CHAPTER 4

CONDITIONALS

Recently, the 'extensional fragment' of natural language has come into its own as a source of important semantic concerns. The previous chapters bear witness to this tendency — postponing the well-known puzzles of intensionality. Nevertheless, our investigation is not at all restricted to the extensional realm — and we shall study some central intensional notions in the next two chapters. The first topic is that of *conditional statements*, perhaps the principal concern of logic, which shows striking parallels with the earlier area of determiners.

General questions concerning the existing multitude of semantic accounts for conditionals are the following. About the language of such statements, should conditionality be an operation upon propositions, or rather a relation between these? As for the semantic apparatus, how can one judge the need for, or the relative merits of the various types of model and truth definition proposed in the literature? Finally, as to the 'logical evidence', what is the status of the intuitions of validity, often invoked as a touchstone for the conditional logic resulting from some particular analysis in this argumentative area?

These are issues which may give rise to lively, but also inconclusive philosophical debate. For instance, operational and relational views of conditionals both have their adherents; and some people even entertain both, to the point of confusing object-language and meta-language of their formalization. To mention another example, the validity of a principle such as Conditional Excluded Middle (*if X, then Y, or, if X, then not Y*) has strong intuitive support, but also provokes grave doubts . . . sometimes within the same observer. What we need, then, is a general unifying perspective, enabling us to arrive at more definite issues and results.

In this chapter, conditional statements *if X, then Y* will be analyzed in the generalized quantifier perspective of the preceding chapters. That is, the conditional particle *if* will denote a *relation between sets of antecedent and consequent occasions.*

It is not claimed that this approach is the uniquely correct one for the study of conditionals. Broadly speaking, there are two major directions in logical studies of conditionality: one 'vertical', having to do with iterations of conditionals and the resulting implicative relations, another 'horizontal', concerned with the interaction between single conditionals and the Boolean connectives *and, or, not*. The former direction is most prominent in the modal 'entailment' tradition (cf. Hughes and Cresswell, 1968), the latter in the study of counterfactuals and related topics (cf. D. Lewis, 1973). Our approach is partial to the horizontal direction, for reasons explained below — although not irrevokably so.

In developing this special theme at length, we also hope to provide a new set of questions for philosophical logic in general, beyond the usual score of completeness theorems and more completeness theorems.

4.1. CONDITIONALS AS GENERALIZED QUANTIFIERS

Throughout the subject of logic, one finds two views of conditionals. Sometimes, implication is a mere connective, then again it is taken to express a relation between propositions. The tendency is of long standing. Thus for instance, Immanuel Kant listed 'hypothetical propositions' under the heading of 'Relation' in his famous Table of Categories. Again, both points of view occur intermingled in C. I. Lewis' account of his intended 'strict implication'. If this is a confusion, as canonical textbook wisdom has it, it is a remarkably tenacious one — a phenomenon which itself requires explanation.

In order to arrive at a perspective doing justice to both viewpoints, one can take a cue from natural language. Unlike coordinating connectives such as *and, or*, the conditional particle *if* functions in subordinate constructions ($if\ X$)Y; where, categorially, $if\ X$ is a sentence modifier. As usual in such linguistic contexts, the full denotation of the expressions 'X', 'Y' may be involved — i.e., the ranges of occasions (worlds, situations, models) where these are true, not just a truth value on one specific occasion. Thus, the force of a conditional particle may be compared with that of ordinary determiners, such as *all, most*, which exhibit similar linguistic behaviour. (The precise nature and extent of this analogy need not be explored here, as no claims will be staked on it. To mention just one more possibly fruitful parallel, the particle *then* seems to function much like an anaphoric pronoun.)

Accordingly, we shall read *if X*, (*then*) Y as expressing some semantic relation $[\![if]\!]([\![X]\!], [\![Y]\!])$ between the sets of antecedent and consequent occasions. Again, our principal task will be to delimit a range of suitable conditional relation between such sets. Further questions will then arise in due course.

More generally, we will have, for any universe E of 'relevant' occasions, and $A, B \subseteq E$,

$$if_E AB,$$

meaning that the conditional relation holds in E between A and B. Examples are inclusion (*all*), majority (*most*) or overlap (*some*). A context-dependent example, where E is essential, is *relatively many*. Eventually, we shall restrict this general pattern to the context-neutral case.

The view of conditionals as relations between sets of occasions would seem to favour what are called *generic* conditionals over *individual* ones. The former refer to sets of events, as in *if* (i.e., whenever) *she comes, she quarrels*. The latter are about single events, as in *if he came, he cried*. Our view is that both statements presuppose variety of occasions. The first is about several events in one world, the second about one event in several possible worlds. The present concept of 'occasion' is meant to include both, as well as their combinations. Against this general background, specific choices of universes E may account for particular kinds of conditionals. For instance, the location of some distinguished 'actual world' in E may be important when treating the contrast between indicative and subjunctive conditionals. (Indeed, E itself may consist of some set of worldlines connected with that actual world.) Henceforth, the abstract common pattern is our central concern.

Another notable aspect of the generalized quantifier approach is that iterated conditionals become awkward to handle. (The reflects the fact that natural language has no direct means of iterating determiner expressions.) In itself, this need not be a defect. There is a well-attested danger of facile logical formalism leading us into iteration that just is not there in ordinary speech. For instance, on the causal reading of conditional relations between (sets of) events, iterated conditionals do not make sense; unless higher layers are interpreted in a different spirit. Such a meaning shift is also discernible in standard examples of iteration in the literature, such as *if this glass breaks if hit, it will break*.

Our notation, differing as it does from the usual arrows, stresses all these points. As in earlier chapters, we shall stick with atomic statements *if* XY, where 'X', 'Y' may be complex Boolean terms. Most fundamental types of conditional inference can already be expressed by these means — although one may eventually enrich the language somewhat, say by adding outer disjunctions (as in Conditional Excluded Middle). (From a more general logical point of view, one might even study the entire range of first-order assertions about the conditional relation, of course.)

Finally, iterations may be brought in after all, once the conditional relation is provided with an additional parameter: $if_{E,\,w}AB$. I.e., *if* AB holds as seen from the vantage point of some particular world w in E. Through lambda abstraction, conditional statements can then be made to correspond to sets of worlds, which may again be used as arguments A, B. (Compare the treatment in Chapter 3 of determiners in direct object position.) Much of what follows can be transferred to this parametrized setting without major changes.

4.2. INTUITIONS OF CONDITIONALITY

What kind of a generalized quantifier is a conditional? Before passing on to the usual display of paradigmatic (non-)inferences, let us reflect. Our intuitions come in various kinds, and it is important to consider the more volatile ones first, concerning the *kind* of notion that we are after, before these are drowned in a list of very specific desiderata. Only in the light of such background intuitions, one can take a proper look at more concrete claims of validity or invalidity of conditional inferences.

The difference may also be illustrated by an example from a different field of semantics. In the logical study of Time, attention is often restricted to the choice of specific axioms for the temporal precedence order, matching certain desired validities in the tense logic. But, there as well, there exist preliminary global intuitions, such as 'isotropy' or 'homogeneity', constituting the texture of our idea of Time, constraining rather than generating specific relational conditions.

Indeed, the ill repute of the term 'intuition' may be due to a misapplication. It is highly unlikely that intuition would settle such specific issues as the validity of concrete inference schemata. An appeal to intuitions at the latter level often amounts to a refusal to argue about the evidence. On the other hand, the proper place for intuition would seem to be at the level of the general structure of our concepts — in the spirit of Kant's philosophy. To paraphrase this great philosopher, we have certain a priori intuitions concerning the basic logical notions, and no human mind is entirely without them.

Global intuitions themselves come in various kinds, having different levels of generality. The first example to be presented here is very specific for conditionals, the next holds for determiners in general, and finally some constraints are reviewed on logical constants as such.

Confirmation. A conditional statement *if* XY claims that 'significantly many' ('enough') X-occasions are Y-occasions. As such, it is tied up with what might be called 'positive' or 'negative' evidence; i.e., cases where both X and Y hold, or cases where X holds without Y, respectively (see Figure 16).

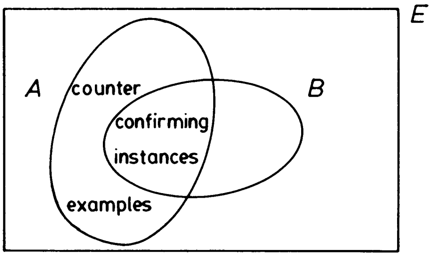

Fig. 16.

Briefly, our intuition is that addition of positive evidence, or removal of negative evidence will not affect a true conditional.

Out of the various ways of making this idea more concrete, here is one suggestive formulation. Suppose that one decides that *if* XY is true in E on the basis of partial information about the extensions $[\![X]\!]$, $[\![Y]\!]$, say $[\![X]\!] = A$, $[\![Y]\!] = B$. Now, further information may tell us that these estimates should be revised to $A' \supseteq A$, $B' \supseteq B$. Then the above intuition says that, if no counter-examples are added in this way, the conditional relation will continue to hold. Formally, for $A \subseteq A'$, $B \subseteq B'$,

$$\text{if} \quad if_E AB \quad \text{and} \quad A' - A \subseteq B', \quad \text{then} \quad if_E A'B'.$$

In practice, it is more convenient to split this up into

(1) fixed A, growth of B:
if AB implies *if* $A(B \cup C)$, for any set C;

(2) simultaneous growth of A, B:
if AB implies *if* $(A \cup C)(B \cup C)$, for any set C.

The first of these principles is known as 'weakening of the consequent' (in our terms, 'upward right monotonicity').

Next, suppose that there are errors in judgments already made about $[\![X]\!]$, $[\![Y]\!]$, and one retreats to $A' \subseteq A$, $B' \subseteq B$. At least, the above intuition tells us that mere removal of counter-examples will not affect the conditional assertion:

$$\text{if } if_E AB \text{ and } A - A' \subseteq A - B, \text{ then } if_E A'B.$$

More elegantly, this becomes the implication

(3) $\quad$ *if* $A(B \cap C)$ implies *if*$(A \cap B)C$.

What about a stronger principle, dual to the above (1), stating that 'strengthening the antecedent' will do no harm?

(3′) $\quad$ *if* AC implies *if*$(A \cap B)C$.

This would mean that possible removal of confirming instances does not affect the conditional either. Except for the extreme case where A is included in C to begin with, such a principle has little to recommend itself.

A fourth and final aspect of Confirmation would seem to be that 'optimal' evidence should verify a conditional:

(4) $\quad$ *if* AB whenever $A \subseteq B$.

This completes the exploration of the most distinctive feature of conditionality.

The next intuition is more general. A conditional statement invites us to take a mental trip to the land of the antecedent. Thus, the assertion of the consequent is only relevant in as far as it holds among the antecedent occasions:

$$if_E AB \text{ iff } if_E A(B \cap A).$$

And so, we again endorse the principle of *Conservativity* from earlier chapters. Most current accounts of conditionals obey this restriction.

Next, being a logical constant, the conditional obeys some of the general features of logicality formulated in Chapters 1, 2. Conditionals should be 'context-neutral': the relation between antecedent and consequent sets involves no more than these sets themselves. Therefore, we assume *Extension* (cf. Section 1.3.). Accordingly, we shall drop the subscript 'E' henceforth. Moreover, there is also a case to be made for

Variety; for essentially the same reasons as in Sections 1.3., 2.1. Together with clause (1) of Confirmation, the latter is equivalent to requiring 'reflexivity': *if* AA, as well as 'consistency': *if* $A\emptyset$ for no A except the empty set. (The latter condition need not be satisfied on, say, a probabilistic approach, where A might be a non-empty set of measure zero. Still, see Veltman, 1985, for a purely 'conditional' defense of Variety.)

Also, the less immediate logical intuitions of Section 2.4. are relevant here. Notably, there should be *Uniformity* in the range of a conditional relation. Accidental features, such as the size of the antecedent set, should not matter in its truth value behaviour. Typically, then, we are invited to make comparisons across different antecedent sets. Thus, recall the earlier 'thought-experiment', involving addition of one element to $A - B$ and $A \cap B$ separately, and then combining the two actions. The outcomes are 'confirmation patterns' of truth values, of the form

$$\begin{array}{ccc} & \text{old situation} & \\ \text{add counter-example} & & \text{add confirming instance} \\ & \text{add both} & \end{array}$$

A priori, sixteen possible outcomes can occur for our thought-experiment. Of these, Confirmation allows only those displayed in Figure 17.

$$\begin{array}{ccc} & + & \\ + & & + \\ & + & \end{array} \qquad \begin{array}{ccc} & + & \\ - & & + \\ & + & \end{array} \qquad \begin{array}{ccc} & + & \\ - & & + \\ & - & \end{array} \qquad \begin{array}{ccc} & - & \\ - & & + \\ & - & \end{array} \qquad \begin{array}{ccc} & - & \\ - & & + \\ & + & \end{array} \qquad \begin{array}{ccc} & - & \\ - & & - \\ & - & \end{array}$$

Fig. 17.

Now, Uniformity will typically constrain the occurrences of such outcomes. The experiment must exhibit certain regularities, independent from the particular location A, B where it is performed. One obvious requirement, then, is *uniqueness of outcomes*: the combined addition should always have its truth value determined by the results of the separate experiments. Thus, the second and third (or fourth and fifth) patterns in Figure 17 cannot occur together for the same conditional. There is room for a whole hierarchy of additional uniformity conditions here; but the present minimal kind of regularity suffices.

Finally, one very conspicuous logical constraint from the preceding will be imposed: for the moment, conditionals are to satisfy *Quantity* (cf. Section 1.6.). This principle of stark austerity may need some clarification here. Quantity may be viewed as a form of Occam's Razor: there should be no more to conditionals than meets the eye. That is, no semantic constructs should be relevant but those sanctioned explicitly by the syntactic material in a conditional sentence. Now, the proper denotations of the antecedent and consequent sentences are bare sets of occasions — and the particle *if* has to denote a relation between these. Violations of Quantity must then always result in 'hidden variable' theories, postulating additional semantic structure among occasions beyond what meets the eye. The latter procedure is quite respectable, of course, in the progress of science. But here, we want to explore the limits of the former, more austere realm of conditionals — if only to see just where this is to be transcended, and what the options are.

4.3. A TRILEMMA

Even if all our general intuitions about conditionals are plausible by themselves, their combined effect may be surprising. After all, the problem with intuitions is usually not their availability or vitality, but rather their *consistency*. Thus, we want to determine which generalized quantifiers are left by the combined postulates of the preceding section.

In a first analysis, there are good reasons for restricting attention here to *finite universes*. It is in accord with the intuitive semantics of natural language (as explained in Section 1.2.) — and it is also the area where proposed explications for conditionals work most smoothly (cf. Lewis, 1981).

Then, we may use the Tree of Numbers (Section 2.2.) to find the conditionals left by our postulates: essentially, two 'democratic' cases and one 'anarchistic' one:

THEOREM. The only conditionals in the present sense are those defined by *all, not least* and *some*.

Here, 'not least' is short for 'half or more', and 'some' stands for 'some or all'.

Proof. That these three satisfy all earlier postulates follows by geometric inspection of their tree patterns. The key observation here is that Confirmation (1), (2), (3) amount to the requirement that, whenever (a, b) belongs to the conditional, then so does the area $(0, \infty), (0, b), (a, b), (a, \infty)$ — as depicted in Figure 18.

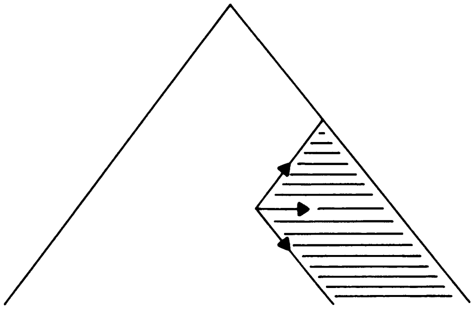

Fig. 18.

Moreover, Variety (including Confirmation (4)) says that the right edge of the tree lies within the +region, while the left edge (minus the top) lies outside of it.

Conversely, consider the pattern for any conditional satisfying our postulates. The top node has +, by Variety. The next row gets −+, for the same reason. The third row leaves a choice in the middle, its boundaries being fixed by Variety again. One possibility is −−+, in which case Uniformity produces a −diagonal along the right edge. By Confirmation then, the tree becomes that of *all* (see Figure 19(i)).

(i)	(ii)	(iii)
+	+	+
− +	− +	− +
− − +	− + +	− + +
− − − +	− + + +	− − + +
⋮	− + + + +	− − + + +
	⋮	− − − + + +
		⋮

Fig. 19.

The other possibility for the third row is −++. This fixes three positions in the fourth row, as before, leaving the second position open. Case 1: −+++. Then, the experiment $-^{-}_{+}$ has produced the outcome +, and will continue to do so (Uniformity). By the other postulates then, the tree becomes that of *some or all* (Figure 19(ii)). Case 2: −−++. Then, four positions in the fifth row are fixed. The remaining one (in the middle) is determined by Uniformity: $-^{+}_{+}$ will invariably produce the combined truth value outcome +. By a similar observation concerning the experiment $-^{-}_{+}$, the tree pattern will be that of *half or more* (Figure 19(iii)). □

So, the minimal analysis of Section 4.2. still leaves three main types of conditionality: at least one confirming instance, at least enough confirming instances to outweigh the counter-examples, and confirming instances only. If the first of these is to be excluded, then further intuitions of conditionality are to be produced excluding it. For instance, one might require 'consistency':

for no A, B: both if AB and if A$(E - B)$.

This particular condition would also rule out the second conditional; but a more discriminating approach is certainly possible.

The above scarcity of admissible denotations depends heavily on the Uniformity condition. Without it, (uncountably) many other conditional patterns pass the test; of which the following still show a good deal of 'regularity' in their truth value tree:

all A except at most n are B,
at least k/n-ths of the A are B.

Even so, there remain forms of conditional reasoning which are not captured by any of these modellings. Notably, in the study of counterfactual conditionals, a *basic subjunctive logic S* has arisen (cf. Burgess, 1981) with the following peculiarities. On the one hand, it differs from the 'classical' inclusion conditional, in that, e.g., strengthening of antecedents is not allowed. On the other hand, it supports the principle of *Conjunction*:

CONJ *if AB, if AC* imply *if A*$(B \cap C)$.

(Note that this is our first example of a *two-premise* principle.) Now, it will be shown in Section 4.10. that, on finite universes, CONJ forces a conditional to become inclusion after all: whence S cannot be modelled in the present setting.

The logical escape routes from here will be charted presently.

4.4. THREE WAYS OUT

The purely numerical approach to conditionals has reached its limits in the above trilemma. There are three main options for transcending this approach, encompassing, between them, most examples of conditionals studied in the literature.

One immediate option which violates none of the earlier intuitions is to lift the restriction to finite sets. Essentially, we have been studying conditionals as binary relations between finite sets of natural numbers — and we might now pass on to the full power set of IN, or even larger infinite sets. One attraction of this *infinitary* approach is the following. In doing semantics on finite models, one is typically concerned with an arbitrary, but presumably 'large' number of occasions. This spirit is sometimes better captured by infinite sets, abstracting from all peculiarities of particular finite sizes. And indeed, Section 4.5. presents some new attractive infinitary conditionals.

The predominant tendency in the area seems to be, not to exploit additional resources of infinity, but to enrich the old (finite) models by imposing additional semantic structure. The intuition to go then is, not a typical conditional principle such as Confirmation, but the general logical one of Quantity itself. (Quantitative violations of, say, Extension or Uniformity will not be pursued here.)

In principle, there are many ways of proceeding here. One is to assign 'weights' to different occasions, by introducing some *probability measure* P on subsets of E. Conditionals then arise such as the following notion of 'relative likelihood':

$$\mathit{if}\,AB \quad \text{if} \quad P(A \cap B) > P(A - B).$$

Note that this inductive approach reduces to the earlier numerical one when P is the equiprobability measure. The inductive approach is the topic of Section 4.6.; where a connection is found with earlier work in the foundations of probability.

Nevertheless, the main thrust in possible worlds semantics for conditionals has been rather to differentiate between individual occasions through patterns of 'accessibility' and 'similarity'. For instance, Quantity is violated in the counterfactual semantics of Lewis (1973), witness Figure 20. (Comparative similarity is just relative distance here; and truths are as indicated.)

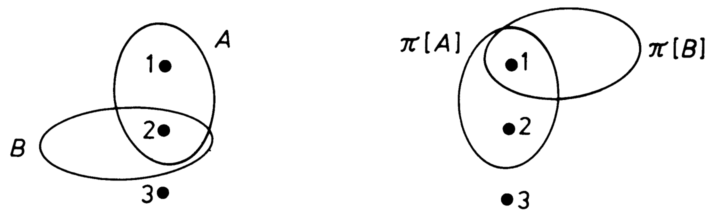

Fig. 20.

Let the function π permute the worlds 1 and 2, leaving 3 in its place. In the left-hand picture, *if AB* holds in world 3 on the Lewis account (all 'closest' A-worlds are B-worlds). But in the right-hand picture, *if* $\pi[A]\pi[B]$ fails in 3; even though the two situations depicted are numerically indistinguishable. So, the underlying *pattern* of the worlds is crucial (cf. Section 1.7. for the similar case of non-quantitative determiners). This hierarchical or *intensional* approach will be studied in Section 4.7.

Further connections between these three ways out of austerity will not be pursued here, nor any alternatives to them.

4.5. CONDITIONALS AT INFINITY

With conditionals viewed as relations between arbitrary sets, finite or infinite, the earlier tree of numbers receives an infinitary superstructure, as in Figure 21.

0, 0

1, 0   0,1

2, 0   1, 1   0, 2

⋮

$\aleph_0$, 0 $\aleph_0$, 1   ... $\aleph_0$, $\aleph_0$ ...   1, $\aleph_0$ 0, $\aleph_0$

$\aleph_1$, 0 $\aleph_1$, 1   ... $\aleph_1$, $\aleph_0$ $\aleph_1$, $\aleph_1$ $\aleph_0$, $\aleph_1$ ...   1, $\aleph_1$ 0, $\aleph_1$

⋮

Fig. 21.

We shall be careful not to become entangled in infinite cardinal arithmetic in these higher regions.

One new kind of conditional connection which can be expressed now is admitting a 'negligible' number of exceptions:

$$\textit{if}\,AB \quad \text{if} \quad \begin{cases} A \subseteq B & \text{, when } A \text{ is finite} \\ A - B \text{ is finite,} & \text{when } A \text{ is infinite} \end{cases} .$$

We shall return to this example presently.

Of the earlier intuitions, Conservativity, Confirmation and Variety remain equally plausible. Uniformity becomes less attractive, however, as one does not expect infinite sets to behave like finite ones in all respects (cf. the above example). Still, one might develop a viewpoint of admissible *extrapolations* from (possibly uniform) conditionals in the finite realm to those on arbitrary sets.

Regardless of the precise reasoning employed, the only really interesting conditionals in the infinite realm are those showing already on the first infinite row of the number tree (Figure 22):

$- - - \ \cdots\blacktriangleright \ - \ \blacktriangleleft\cdots \ - - +$	all
$- - - \ \cdots\blacktriangleright \ - \ \blacktriangleleft\cdots \ + + +$	most
$- - - \ \cdots\blacktriangleright \ + \ \blacktriangleleft\cdots \ + + +$	not least

Fig. 22.

Two of these are straightforward extrapolations of the main examples in Section 4.3., the one in the middle is the above case.

Do these infinite extrapolations differ significantly from their finite counterparts, notably in the inferences they support? *All* (inclusion) has precisely the same logic as on the finite universes; as is easy to see. (In any case, our only worry would be that one might lose inferences, having added so many models.) Our conjecture is that the same holds for *not least*.

The logic of the above example is more intriguing, however. It validates Conjunction (from *if* XY and *if* XZ to *if* $X(Y \wedge Z)$), which sets it apart from the *not least* logic. Also, it fails to validate transitivity, which distinguishes it from the logic of *all*. (To see this, consider some infinite A with finite $A' \subseteq A$ outside of B: *if* AB, *if* $A'A$, yet not *if* $A'B$.) Its exact behaviour is rather mysterious (cf. van Benthem, 1984b).

Still, as Frank Veltman has pointed out, there is another plausible way of extrapolating from the middle row of Figure 22:

$$
if AB \quad \text{if} \begin{cases} A \subseteq B & \text{, when } A \text{ is finite} \\ |A \cap B| > |A - B| & \text{, when } A \text{ is infinite.} \end{cases}
$$

As it turns out, the latter conditional yields the basic subjunctive logic S. (Cf. Veltman, 1985. The idea is this. All S-principles (cf. Section 4.9.) are valid. Conversely, one takes a finite S-counter-example for a conditional inference (cf. Burgess, 1981), converts this into a connected Lewis counter-model (see Section 4.9.), and then adds infinitely many copies of worlds in the concentric spheres, starting from the outside, with stepwise increasing infinite cardinalities. This has the effect of simulating the clause 'in all closest A-worlds' by 'for most A-worlds'.)

So, there is a purely numerical modelling of the basic subjunctive logic after all.

4.6. INDUCTION

Universes of occasions may be weighed by a probability measure P assigning real numbers $P(A)$ to sets $A \subseteq E$, subject to the following conditions:

$$
\begin{array}{ll} 0 \leqslant P(A) \leqslant 1;\ P(\emptyset) = 0,\ P(E) = 1 & \text{(normality)} \\ \text{for disjoint } A, B,\ P(A \cup B) = P(A) + P(B) & \text{(additivity).} \end{array}
$$

In this section, only finite universes will be considered.

Several 'inductive' conditionals may be defined employing such a measure. One says that $B \cap A$ fails to cover A by a zero-set:

$$
if AB \quad \text{if} \quad P(A \cap B) = P(A).
$$

The logic of this conditional is just the classical one of entailment. And in fact, when P is the equiprobability measure on E, this conditional reduces to *all*. Analogously, a generalization of *not least* arises, as a merest tip of the balance:

$$
if AB \quad \text{if} \quad P(A \cap B) \geqslant P(A - B).
$$

This stipulation is related to an idea in Lenzen (1980), who reads A-conditional belief of B as 'B is more likely than not, in the realm of A-worlds'.

In the spirit of this book, at issue are not so much special examples as general constraints on admissible probability measures, and conditionals based upon them — within and across various universes. Such a probabilistic picture suggests intuitions of its own, that would not come to the fore in the purely numerical setting of Section 4.2. For instance, conditionals $if_{E,P}$, now viewed as functions from $\mathscr{P}(E) \times \mathscr{P}(E)$ to $\{0, 1\}$, may be required to satisfy various 'smoothness' properties involving P. E.g., they might be in 'equilibrium', in the sense that P-small shifts in argument sets ought to leave truth values unperturbed. Actually, this is better stated in a more general perspective of 'fuzzy conditionals', assuming truth values in the real interval $[0, 1]$: where *if* may become *continuous* in the usual sense of analysis.

To conclude this short foray, here is a parallel between the generalized quantifier approach and an earlier historical one.

In his foundational studies of probability, Bruno de Finetti introduced a notion of 'relative probability' between sets of outcomes: $A \leqslant B$, meaning that 'A is at most as plausible as B'. (Cf. Lenzen, 1980, Chapter 4, for further details on the following topic.) He then produced a list of intuitive desiderata, including

(1) $\emptyset \leqslant A \leqslant E$, (2) $\emptyset \ngeqslant E$, (3) if $A \leqslant B \leqslant C$, then $A \leqslant C$, (4) $A \leqslant B$ or $B \leqslant A$, (5) if $A \cap B \leqslant A \cap C$, then $A - C \leqslant A - B$.

The guiding hope was that these would provide necessary and sufficient conditions for this primitive relation to be represented through some probability measure P on E: $A \leqslant B$ if and only if $P(A) \leqslant P(B)$. Later investigations have revealed that further, less intuitive combinatorial postulates are needed for this purpose.

There is an intimate connection between de Finetti's notion and the above inductive *not least* conditional *if*. Some calculation will show that

$$\begin{aligned} &if AB \quad \text{iff} \quad A - B \leqslant A \cap B, \\ &A \leqslant B \quad \text{iff} \quad if(A \,\Delta\, B)B. \end{aligned}$$

Here, 'Δ' denotes symmetric difference. Thus, the De Finetti axioms generate a conditional logic. Amongst others, (5) becomes universally valid, (4) is Conditional Excluded Middle, while (3) becomes the 'Δ-principle' (cf. van Benthem, 1984b):

$$if(A \,\Delta\, B)B, \quad if(B \,\Delta\, C)C \quad \text{imply} \quad if(A \,\Delta\, C)C.$$

The latter holds for *not least* — and indeed, we have a conjecture: does the De Finetti logic coincide with the latter one?

4.7. INTENSION

The usual approach in conditional semantics has been a hierarchical one. Possible worlds can be more or less close to some vantage world, and the conditional is only concerned with 'closest' antecedent worlds. In our perspective, where the universe itself may already be derived from some vantage point, a hierarchy is just some binary relation. Thus, the generalized quantifier *if* will now assign binary set relations to (*finite*) structured universes $\langle E, R\rangle$, say. The earlier numerical perspective was a democratic one, with an empty (or universal) relation R — but now, certain occasions may wield greater influence than others.

One typical hierarchical conditional considers top-ranking occasions only:

all R-maximal occasions in A are in B.

We shall review the earlier intuitions (Section 4.2.) for this example. First, of course, *Quantity* has been abandoned — but there remains a principle of *Quality* (cf. Section 1.7.):

> conditional relations are invariant under the action of R-isomorphisms between universes $\langle E, R\rangle$.

For, such isomorphisms preserve the relevant hierarchical structure. Within a single universe, the force of this postulate depends on the R-pattern. If R is empty (or universal), every permutation is an R-automorphism, and Quality becomes Quantity. Thus, the present approach subsumes the earlier numerical one. If, on the other hand, every individual is uniquely distinguished by its position in the hierarchy, then identity is the only R-automorphism, and the postulate becomes empty.

Continuing with the other constraints, *Conservativity* remains equally plausible. The case of *Confirmation* is slightly more interesting. Evidently, this principle should remain valid — but, this may impose some (mild) conditions on the hierarchy. For instance, the top-ranking conditional satisfies clauses (1), (2) and (4) without further ado — but for clause (3), *transitivity* and *irreflexivity* are needed for R.

Proof. Ad (2): Suppose that *if* AB. Consider any top-ranking w in $A \cup C$. Either $w \in C$, and so $w \in B \cup C$; or $w \in A$ — whence it is R-maximal in A, and therefore, by the assumption, $w \in B$, $w \in B \cup C$.

Ad (3): Suppose again that *if* AB. Now remove a counter-example from $A - B$. Let w be R-maximal in the resulting set A^-. We are done if it also was R-maximal in A. Suppose otherwise. Then some $w_1 \in A$ lies R-above w. But then, by the above assumptions on R (and the finiteness of our universe), there must be some R-maximal v in A which is R-above w. This v belonged to B, and so it has not been removed. Being in A^-, then, it refutes the R-maximality of w in the latter set: a contradiction. □

This example shows three 'degrees of freedom' for a particular conditional: general constraints, a choice of truth definition, and special conditions on auxiliary semantic structure. This theme will be studied in greater depth in Chapter 9.

Variety plays the same role as before. In our paradigm example, its validity depends on the existence of R-maximal worlds in non-empty sets: something which again depends on the above assumptions.

On infinite universes, the top-ranking conditional need not satisfy either Variety or Confirmation. This reflects a well-known problem with Lewis' possible worlds semantics. On finite models, his truth condition has the intent of our top-ranking — but the infinite case leads to the less intuitive clause that 'some $A \cap B$-world is R-closer than every $(A - B)$-world'. Still, there is a close connection between the above example and the Lewis tradition.

THEOREM. The top-ranking conditional validates precisely the basic subjunctive logic S.

Proof. The main idea is that all axioms from Burgess (1981) (presented in Section 4.9. below) are valid for our conditional — while conversely, Burgess-counter-examples, which can always be taken to be finite, are hierarchies in the above sense. □

After this digression, we consider the remaining intuitions of Section 4.2. *Extension* remains appropriate, be it with respect to 'extensions' of universes $\langle E, R \rangle$ in the ordinary model-theoretic sense. And finally, *Uniformity* acquires a new flavour in the present perspective. The characteristic thought-experiment now consists in adding new worlds to a hierarchy. This will influence our description of the possible outcomes, in terms of preserving or destroying certain R-patterns. One particular implementation will be presented here, for the purpose of illustration.

By Confirmation, confirming instances can always be added, at each position in the hierarchy. For counter-examples, let us distinguish three possible actions: (1) insert in top position (without $A \cap B$-superiors), (2) insert below some $A \cap B$-world (not necessarily *immediately* below), (3) insert below some $(A - B)$-world (likewise). Uniformity now says that allowing such an action once means allowing it always. The force of this principle will appear below.

The richer hierarchical perspective also suggests new intuitions of its own. For instance, as seen from each individual occasion, the hierarchy is only relevant in as far as it is 'accessible'. Formally, call a substructure of $\langle E, R\rangle$ a *subhierarchy* if each of its occasions retains all its R-superiors and R-inferiors from E. Then the principle of *Relevance* states that conditional assertions are preserved in passing from a hierarchy to its subhierarchies. That is,

if $\langle E', R'\rangle$ is a subhierarchy of $\langle E, R\rangle$, then
$if_{\langle E, R\rangle} AB$ implies $if_{\langle E', R'\rangle} (A \cap E')(B \cap E')$.

As in Section 4.3., the effects of our combined intuitions may be investigated, for finite transitive irreflexive hierarchies.

THEOREM. The only two conditionals satisfying Quality, Conservativity, Confirmation, Variety, Extension as well as Uniformity and Relevance are

all X are Y, all top-ranking X are Y.

Proof. By a simple calculation, these two conditionals satisfy all principles mentioned.

Conversely, consider any conditional *if* subject to these constraints.

Claim 1: Relevance rules out action (1) occurring in the statement of Uniformity. For, by Confirmation, one single $A \cap B$-occasion verifies *if AB*. Action (1) would allow the addition of a single R-isolated $(A - B)$ occasion, while *if AB* remains true. But then, by Relevance, the latter alone would verify the conditional: which contradicts Variety.

Claim 2: Either *if* is inclusion, or it allows both action (2) and (3) occurring in the statement of Uniformity.

For, if there exists any situation $if_{\langle E, R\rangle} AB$ with A not included in B, then that hierarchy contains some $(A - B)$-occasion. Now, this occasion w_1 cannot occur in top position. For, otherwise, removing it leaves the conditional true (by Confirmation); and hence, in retrospect, action (1) was allowed after all. So, by an earlier argument, there must be some other R-maximal occasion w_2 above w_1 — and evidently, the latter must be in $A \cap B$. But then, the same reasoning of removal/reversal (applied to w_1, w_2) shows that action (2) is admissible. Thus, $(A-B)$-occasions may be inserted below w_2, in particular also below w_1. And that again means that action (3) is admissible too.

Claim 3: When actions (2), (3) of Uniformity are admitted, the conditional is that of top-ranking.

Here, in one direction, each situation where all top-ranking A are B can be created from single $A \cap B$-occasions (where the conditional holds, by Variety), through judicious insertion of confirming instances intermingled with (2), (3)-insertions. Conversely, suppose that at least one situation is admitted with some top-ranking $(A - B)$-occasion. Omit this occasion (by Confirmation): in reverse, action (1) has been allowed, in contradiction to the first claim. □

Thus, on the above analysis of broad intuitions, the hierarchical perspective allows just ordinary modal entailment as in the purely numerical case, while adding one new possibility (top-ranking), which generates the basic subjunctive logic in the Lewis—Stalnaker tradition.

4.8. THE RANGE OF CONDITIONAL TRUTH DEFINITIONS

Even though conditionals were treated as abstract generalized quantifiers in the above, the presentation of specific examples, or the statement of classification theorems usually proceeds by definition in some standard logical language. Indeed, such descriptions may be viewed as possible *truth definitions* for abstract conditionals, satisfying certain intuitive constraints, with respect to some background class of universes of occasions.

In the logically simplest case, call a conditional *if first-order definable* if there exists some formula $\varphi = \varphi(X, Y)$ in the monadic first-order language with identity and unary X, Y such that for all E and A, $B \subseteq E$,

$$if_E AB \quad \text{iff} \quad \langle E, A, B\rangle \models \varphi.$$

For instance, two of the conditionals in Section 4.3. were first-order definable: $\forall x(Xx \rightarrow Yx)$ (*all*), $\exists x(Xx \wedge Yx) \vee \forall x(Xx \rightarrow Yx)$ (*some or all*). Because of the preservation of first-order statements under isomorphism, all such conditionals satisfy Quantity.

Definitely outside of this class is the third conditional in the earlier Trilemma. *not least* is not even first-order definable on the finite sets alone. This follows from the characterization of first-order definability in the tree of numbers, presented in Section 2.5.

The limitations of a first-order language also become apparent in the following illustration, drawn from the Introduction to Suppe (1974).

Intermezzo: The Logic of Dispositions

The only well-known alternative to material implication known in the thirties was modal entailment: *all X are Y* (*X* is a 'sufficient condition' for *Y*). Now, when philosophers of science started considering conditional assertions in scientific contexts, they ran into the problem that entailment does not work for dispositional statements. E.g., the sentence *This lump of sugar is soluble*, which presumably means *If this lump of sugar is put into water, it dissolves*, cannot be transcribed as 'all watering occasions for this lump are dissolution occasions'. One problem is that continuously dry objects would have to be called soluble then, for trivial reasons. This could be remedied by enlarging the setting to all possible occasions — whether actual in this world or not. But even so, the conditional remains too strong in another sense. One is only referring to all watering occasions 'under normal circumstances'. Typically, this means that dispositional conditionals will not admit of strengthening their antecedents, as this may bring in non-normal circumstances. (*If this lump of sugar is put into water and withdrawn at once, then . . .* ?)

These problems led Carnap to formulate an amendment to the 'Received View' of scientific theories. In addition to ordinary first-order predicate logic, one would have to allow intensional (notably, counterfactual) logic, even in the observational base of the theory (cf. Carnap, 1956).

There is a curious weakness to the argument for such a move. One considers a certain kind of natural language statement, one tries a simple-minded predicate-logical transcription: this then turns out to fail — and one concludes that *no* predicate-logical transcription whatsoever will be adequate. This pattern of reasoning is also quite current in defense of the thesis that 'predicate logic is insufficient for semantics'.

To settle the problem in a more definite manner, it should be clear which logic of dispositionals is to be explicated. One obvious candidate here is the earlier basic subjunctive logic. Modulo some reasonable background assumptions, there is now an answer.

THEOREM. No first-order definable conditional $\varphi(=, X, Y)$ generates precisely the basic subjunctive logic.

Proof. See van Benthem (1984b). □

This result also provides a justification for the break with monadic first-order definability in current counterfactual semantics.

But perhaps, dispositional statements have a logic different from the above subjunctive one? Still, whatever its precise form, that logic will lack Persistence, whether downward or upward. But then, the definability theorem of Section 2.5. already implies a refutation. For, it may be shown that any first-order definable conditional satisfying both Conservativity and Confirmation will eventually settle down to either a ↓MON or a ↑MON type.

One curious feeling remains. As with other instances of definability questions, laborious formal analysis has eventually confirmed the earlier heuristics: 'it cannot be done *simply*. So, it cannot be done *at all*'. Could it be that a profound Principle of Perfection governs our world:

The truth is always simple?

End of intermezzo

From the austere monadic first-order language, one may ascend in at least two directions. One is to increase logical power, passing on to *higher-order* notions; the other is to increase power of perception, *enriching the vocabulary*. There is an argument for a preference here. A truth definition ought to be as simple as possible, not presupposing any higher-order entities in the models. If the latter are thought important, they should be incorporated into these models explicitly, together with the crucial axioms governing them. (Cf. also Chapter 9 on this issue.) E.g., the probabilistic approach of Section 4.6. shifted the higher-order complexity of *not least* to explicit probability measures P, using which the truth definition could become first-order again.

For first-order definable conditionals, it may be ascertained which syntactic constraints (in the presence of arbitrary 'hidden variables') are induced by the intuitive postulates of earlier sections. Essentially, the latter amount to ordinary model-theoretic 'preservation properties'. For instance, *Conservativity* corresponds to syntactic *restriction* of all quantifiers to the antecedent predicate X, as has been noted in Section 2.5. *Confirmation* requires at least Monotonicity, which induces syntactically *positive* occurrences for the consequent predicate Y. The further parts of Confirmation impose additional restrictions — which are as yet undetermined.

On the basis of these constraints, special purpose conditions may decrease the range of possible truth definitions even further. For instance, 'strengthening the antecedent' essentially removes occurrences of the existential quantifier, leaving syntactically 'universal' schemata only.

4.9. CONDITIONAL LOGICS

Besides global intuitions, there exist also convictions concerning the validity, or desirability of particular inference patterns for conditionals. Now, proposals for 'conditional logics' have varied widely. Moreover, their motivation is sometimes unclear — especially in those cases where merely some suspect 'classical' laws are removed from the usual corpus. Therefore, let us take stock of the privileged conditional logics that have emerged naturally in this chapter. Our medium will be the simple language of Section 4.1., with atomic $if\,XY$, for Boolean terms X, Y.

First and foremost, Section 4.2. inspires a *minimal conditional logic* M, with principles

(1) $\dfrac{if\,XY}{if\,X(Y \wedge X)}$

(2) $\dfrac{if\,X(Y \wedge X)}{if\,XY}$ (Conservativity)

(3) $\dfrac{if\,XY}{if\,X(Y \vee Z)}$

(4) $\dfrac{if\,XY}{if\,(X \vee Z)\,(Y \vee Z)}$

(5) $\dfrac{if\,X(Y \wedge Z)}{if\,(X \wedge Y)Z}$

(6) $if\,XX$ (Confirmation)

Actually, axiom (3) is redundant here.

The next important logics arise from the trilemma of Section 4.3. First, there was (S5-)modal entailment, axiomatizable as the *classical conditional logic C*, consisting of M with the following additions:

(7) $$\frac{\mathit{if}\, XY \quad \mathit{if}\, YZ}{\mathit{if}\, XZ}$$ (Transitivity)

(8) $$\frac{\mathit{if}\, XY}{\mathit{if}\, (X \wedge Z) Y}$$ (Left-monotonicity)

(9) $$\frac{\mathit{if}\, XY \quad \mathit{if}\, XZ}{\mathit{if}\, X(Y \wedge Z)}$$ (Conjunction)

(10) $$\frac{\mathit{if}\, XY \quad \mathit{if}\, ZY}{\mathit{if}\, (X \vee Z) Y}$$ (Disjunction)

There is a good deal of redundancy here, for clarity of presentation.

Then, the *not least* conditional generates the *preferential conditional logic P*. The latter lacks classical laws such as transitivity or downward persistence. Our conjecture is that P can be recursively axiomatized. But in practice, its principles are difficult to locate (cf. also Section 4.6.).

What we do know is this.

THEOREM. $M \subseteq P \subseteq C$ — and all inclusions are proper.

Proof. Here are the only two non-immediate assertions.

$M \neq P$: the Δ-principle of Section 4.6. was valid for P; but it fails for the conditional *all but at most one* which does satisfy the minimal conditional logic.

$P \subseteq C$: Let the inference from $\mathit{if}\, X_1 Y_1, \ldots, \mathit{if}\, X_n Y_n$ be refuted by inclusion in some model. Already, *not least* X_i are Y_i $(1 \leqslant i \leqslant n)$, because *all* are. On the other hand, at least one object inhabits $X - Y$. Now, add a number of copies of this object, behaving exactly the same as to (non-)membership of the relevant sets, such that the cardinality of $X - Y$ starts exceeding that of $X \cap Y$. This procedure disturbs none of the previous relations $\mathit{if}\, X_i Y_i$ — and thus, we have a *not least*-counterexample. □

Finally, *some or all* validates an *exemplary conditional logic E*, with axioms of reflexivity and symmetry. Evidently, this is not a serious competitor.

Of greater interest are some intermediate logics. For instance, the classical logic adds *two* kinds of principle to the minimal one. There is 'transmission' (transitivity, left-monotonicity), but also 'combination', as in Conjunction and Disjunction. The latter phenomenon is of interest by itself, and we may define the *subjunctive conditional logic* S as the result of adding the latter two principles to M. The motivation for the name lies in the following result.

THEOREM. S is precisely the basic subjunctive logic.

Proof. The principles of Burgess' presentation are Reflexivity, Conjunction, Right-Monotonicity, Disjunction as well as the inference from *if* XY, *if* XZ to *if* $(X \wedge Y)Z$. All derivations involved in this equivalence are straightforward, both ways. □

As a byproduct, the final Burgess axiom, often thought rather ad-hoc, receives a natural motivation through clause (3) of Confirmation.

It remains to establish the place of S in the above scheme. Obviously $M \subseteq S \subseteq C$, where all inclusions are proper. But, there is a deeper connection between M and S.

THEOREM. M coincides with the one-premise fragment of S.

Proof. Consider any invalid inference from *if* XY to *if* ZU in M. We shall find a Lewis-model which is an S-counterexample. First, some transformations are useful, into M-equivalent assertions:

if XY to *if* $X(Y \wedge X)$ to *if* $(X_1 \vee X_2)X_2$, where X_1, X_2 are disjoint disjunctions of complete state descriptions (composed out of the proposition letters occurring in X, Y, Z, U), such that $X \leftrightarrow X_1 \vee X_2$, $Y \wedge X \leftrightarrow X_2$;

if ZU likewise to *if* $(Z_1 \vee Z_2)Z_2$.

Now, as *if* ZU is non-derivable, Z_1 cannot be empty. By itself, one single world verifying some state description from Z_1 would already falsify the conclusion — being a closest $Z_1 \vee Z_2$-world to itself which is not Z_2. But, the premise imposes the condition that closest $X_1 \vee X_2$-worlds must be X_2-worlds. I.e., for every X_1-world, there must be some closer X_2-world. The latter condition is only operative when $Z_1 \subseteq X_1$. And even then, the obvious dodge is to pick a Z_1-world, with respect to some vantage world in $X_2 - Z_2$. This again will fail to falsify the conclusion only if $X_2 \subseteq Z_2$. But then, with these two inclusions, in M we have, successively,

$if(X_1 \vee X_2)X_2$, $if(Z_1 \vee X_2)X_2$ (by axiom (5)),

$if(Z_1 \vee (X_2 \vee (Z_2 \wedge \neg X_2)))\,(X_2 \vee (Z_2 \wedge \neg X_2)))$ (by axiom (4)).

Applying some Boolean identities then, the conclusion $if(Z_1 \vee Z_2)Z_2$ emerges after all: a contradiction. □

The counterexamples obtained in the above are even 'connected' in the sense of Lewis (1973): whence M is also the one-premise fragment of the full original Lewis logic. This observation raises a similar question for S itself.

THEOREM. S is the many-premise fragment, in our restricted formalism, of the full Lewis logic.

Proof. In one direction, every S-principle is Lewis-valid. Conversely, let the inference from $if\, A_1B_1, \ldots, if\, A_kB_k$ to $if\, AB$ be refuted on some finite model with irreflexive transitive R, in some world w. I.e., there is an R-closest A-world where B fails, say x. Now, consider the 'generated submodel' consisting of w, x and all worlds on R-paths in between these two. *If* AB still fails at w, while all $if\, A_iB_i$ are still true ($1 \leqslant i \leqslant k$). (Note that x is the only A-world in this new model.) Next, this model in rearranged in concentric spheres, according to the distance function given by

> distance$(w, y) =_{\text{def}}$ the maximum length of an R-path going from w to y.

This procedure may distort truth values; but, we still have

(i) $if\, AB$ is false at w (recall that there was only one A-world),

(ii) all $if\, CD$ formerly true at w remain true (for, R-closest C-worlds under the new arrangement must have been R-maximal in C before).

A somewhat different proof may be extracted from Lewis (1981). □

Finally, as for the connection between the subjunctive and the preferential logic, the answer is as follows:

THEOREM. P is properly contained in S.

Proof. Conjunction is an S-principle which is P-invalid. But, any inference which is S-invalid is outside of P too. For, let the inference be refuted in some finite S-model, and hence also, in some finite Lewis-model (cf. the preceding argument). A suitable finite version of the infinite 'blow-up' argument in Section 4.5. will produce a *not least* counter-example. (This time, one multiplies copies of worlds, proceeding from outer to inner circles, according to the following instruction: in circle $i + 1$, take as many copies for each world as the sum total of all copies created in circles $1, \ldots, i$.)

Thus, the major conditional logics from a linear sequence: $M \subseteq P \subseteq S \subseteq C$.

4.10. GENERAL PATTERNS OF CONDITIONAL INFERENCE

The preceding section ended on an all too familiar track in intensional semantics: a proliferation of logics. But, our perspective not only generates specific privileged logics — it also provides a means for investigating possible conditional inference patterns, without being tied to exclusive clusters. Thus, we can look at arbitrary inferential theories, asking for the conditional relations validating at least, or just these. This is not just the common question of 'modelling' some given logic: we are after the entire range of modellings, so to speak.

First, consider pure *if*-patterns without connectives, expressing ordinary relational conditions. Of these, only reflexivity and transitivity are serious candidates. Then, earlier results from Sections 1.5., 2.6. express that

(1) every reflexive transitive conditional is *transmitting*, allowing both strengthening of antecedents and weakening of consequents.

(2) Modulo Variety, inclusion is even the *only* reflexive transitive conditional. In fact, the relevant argument would go through with even less:

(3) Inclusion is the only conditional relation allowing strengthening of antecedents.

This explains why 'non-monotonicity' is the hallmark of all current non-classical conditional logics.

The above cluster of requirements is reminiscent of Scott (1971), who demands — in our terminology — the following fundamental properties for a 'conditional': reflexivity, transitivity, left- and right-monotonicity. Assuming Conservativity, (1) tells us that these postulates are not independent. Moreover, assuming Variety, (2) adds the insight that the only Scott conditional must be modal entailment.

In addition to pure patterns, there are also mixed conditional inferences, involving connectives — such as many of the axioms in Section 4.9. By way of illustration, we look at *Conjunction* in the present prespective.

First, this principle follows already from the above two, on the basis of M:

THEOREM. Every reflexive transitive conditional satisfies Conjunction.

Proof. Assume *if* AB, *if* AC. From the first, *if* $A(B \cap A)$ (by CONS). From the second, *if*$(B \cap A)C$ (as reflexive transitive conditionals are downward persistent), and hence *if*$(B \cap A)(B \cap A \cap C)$ (by CONS), *if*$(B \cap A)(B \cap C)$ (as reflexive transitive conditionals are upward monotone). By transitivity then, *if* $A(B \cap C)$. □

To some extent, a converse is true as well:

THEOREM (Quantity, Variety). On finite universes, *all* is the only conditional satisfying Conjunction.

Proof. If *if* has this property, and *if* AB holds without $A \subseteq B$, then $B \cap A \subsetneq A$ and also *if* $A(B \cap A)$. But then, by intersecting $B \cap A$ with equally large distinct sets $C \cap A$ (for which *if* AC, by Quantity), one obtains *if* AD for ever smaller sets $D \subseteq A$: and in the end, *if* $A\emptyset$. Therefore, by Variety, $A = \emptyset$; and so $A \subseteq B$ after all: a contradiction. □

On infinite universes, matters change. E.g., the infinitary conditional *all but finitely many* (as defined in Section 4.5.) satisfies VAR, CONJ without being transitive.

This observation raises an interesting general topic, viz. the non-obvious connection between inferences among *if*-schemata which are valid on the finite sets, and those which are valid everywhere.

Note also that the argument in the preceding proof can be generalized to infinite cardinalities. For infinite b, when $a_1, b \in$ *if* and $a_2, b \in$ *if*, with $a_1, a_2 < b$, then also $a_1 + a_2, b \in$ *if*. Moreover, assuming VAR, the middle position $b, b \notin$ *if*: otherwise, one suitable conjunction would yield $b, 0 \in$ *if*.

Next, consider the sister principle of *Disjunction*. First, on the basis of M, this follows already from CONJ:

$$\frac{\dfrac{if\,AC}{if(A \vee B)(C \vee B)} \qquad \dfrac{\dfrac{if\,BC}{if(B \vee (A \wedge {}^{-}B))(C \vee (A \wedge {}^{-}B))}}{if(A \vee B)(C \vee (A \wedge {}^{-}B))}}{\dfrac{if(A \vee B)((C \vee B) \wedge (C \vee (A \wedge {}^{-}B)))}{\dfrac{if(A \vee B)(C \vee (B \wedge (A \wedge {}^{-}B))}{if(A \vee B)C.}}}$$

The converse is not true: the disjunctive conditional *some or all* lacks CONJ.

So, the cumulative effect of all S-axioms on quantitative conditionals is just that of M + CONJ. Essentially, this leaves us with the couples $0, n$ on the finite sets, and then, in the infinite realm, intervals $a, b \ldots 0, b$ $(a < b)$; subject to the Confirmation constraint across the rows (cf. Section 4.3.). Intuitively then, S-conditionals are those which are 'classical' on the finite sets, and then tolerate a limited number of exceptions on the infinite ones (cf. Section 4.5.).

A less standard example beyond S is the following 'dual' of Conservativity:

$$if\,AB \quad \text{if and only if} \quad if(A \vee B)B. \qquad (*)$$

From left to right, this follows from M. The converse contains some strengthening of the antecedent. How much? As always, the quantitative approach gives a useful impression. The above equivalence holds for *if* if and only if it satisfies the following numerical condition:

$$a, b \in if \quad \text{if and only if} \quad a, b' \in if, \quad \text{for all cardinalities } a, b, b'.$$

So, *if* fills the tree of numbers in entire \-diagonals, as its behaviour depends on the number a only. (Thus, actually, it is more of a dual to the *symmetric* quantifiers of Section 1.5., whose behaviour depends on the number b only.)

As a consequence, on the basis of M, any *if*-pattern satisfying (*) will have the monotonicity type $\downarrow$MON$\uparrow$ (cf. Section 2.2.). Indeed, the only M-principle needed here is that form of Confirmation which allows upward $\nearrow$-travel in the tree: i.e., clause (3) in Section 4.2. Question: is there a direct deduction from principle (*) together with Confirmation (3) to the above two monotonicity properties?

There is an interesting general issue behind this question, concerning 'completeness' of quantitative reasoning. Is there a general system of conditional logic μ with the following property: whenever a set L of inference schemata for *if* implies another schema γ for all *quantitative* conditionals *if*, then γ is derivable from L via μ? Obviously, μ will include Boolean equivalences, as well as arbitrary substitutions for L-principles. But this seems insufficient by itself. The question makes sense for arbitrary quantifiers as well (cf. Section 1.7.).

Other mixed patterns arise when *negation* is introduced. The ubiquity of modal entailment shows again with the best-known inference of this kind.

THEOREM (Variety, Extension). The only reflexive conditional satisfying *Contraposition* is inclusion.

Proof. As always, $A \subseteq B$ implies *if* AB. Conversely, suppose that *if* AB. Thanks to Extension, Strong Conservativity holds — and so we may argue: *if* AB, *if*$(A-B)(A-A)$ (by Contraposition), *if*$(A-B)\emptyset$, $A-B=\emptyset$ (by Variety): i.e., $A \subseteq B$. $\square$

One important variation on the above theme is the completeness issue of, given a set of inference patterns, which conditionals validate *precisely* these. For instance, no specific conditional relation has been found yet validating the minimal conditional logic, and no more.

These samples conclude our first foray from the extensional into the intensional realm.

CHAPTER 5

TENSE AND MODALITY

Conditionals may be viewed as relations between propositions, as we have seen, in striking analogy with the extensional treatment of determiners and quantifiers. But the more traditional intensional constructions are rather *operators* on propositions, such as modality or tense. Still, given the generalization made in Chapter 3 to arbitrary extensional categories, a similar move is possible for intensional notions. Instead of pursuing this topic in its full generality, we present two concrete cases.

5.1. TENSES IN REAL TIME

Let us fix one temporal structure, the real number line IR. Tenses may be regarded as operations on propositions — which latter, in this model, correspond to sets of real numbers: the times when they are true. In the area of 'tense logic', a linguistic approach has been predominant, tenses being thought of as all operations definable in some operator language with PAST, FUTURE, etc., or perhaps in some first-order predicate logic with temporal parameters (cf. van Benthem, 1982b). For instance, the basic tenses in the Prior tradition are

$F\varphi$: it will be the case (at least once from now) that φ,
$P\varphi$: it has been the case (at least once before now) that φ.

Combinations of these operations will then account for compound tenses (*will have been, had been, would be*). Modulo logical equivalence, there are fifteen tenses of this kind on the real numbers (Hamblin's 'Fifteen Tenses Theorem'). Even so, there are tenses beyond this language; notably, the *progressive* (*Lucas is crying*), for which one may add a further operator, say Π. In the limit, such additions tend to converge to a more liberal medium of expression, viz. a first-order language on IR having special predicates $=$, $<$ for *identity* and *precedence* of moments in time, as well as unary predicates, representing slots for time-dependent component propositions.

Example: For a moment of evaluation t_0, Fq may be written as

$\exists t > t_0 Qt$, Pq as $\exists t < t_0 Qt$, and, e.g., Πq as $\exists t_1 < t_0 \exists t_2 > t_0 \forall t(t_1 < t < t_2 \rightarrow Qt)$.

Many other potential 'tenses' can be expressed in this way. One famous result in tense logic is Kamp's Theorem stating that all such first-order definable tenses on IR can already be defined in an operator language with two binary notions

$$\text{SINCE}\,(p, q)\colon \exists t_1 < t_0\,(Pt_1 \wedge \forall t(t_1 < t < t_0 \rightarrow Qt))$$
$$\text{UNTIL}\,(p, q)\colon \exists t_1 > t_0\,(Pt_1 \wedge \forall t(t_0 < t < t_1 \rightarrow Qt)).$$

These expressions correspond to temporal adverbs rather than actual tenses.

In the present perspective, we are interested in an alternative, more structural approach. With all possible tense denotations viewed as functions from sets of reals to sets of reals, can we find suitable semantic constraints characterizing the ones actually realized in language?

A first, obvious condition is that tenses respect only the temporal order. That is, in earlier terms, there is a maxim of *Quality* for such functions f:

for all order-preserving automorphisms π of IR,
$\pi[f(A)] = f(\pi[A])$, for all $A \subseteq R$.

Equivalently,

$$y \in f(A) \quad \text{iff} \quad \pi(y) \in f(\pi[A]), \quad \text{for all } y \in R, A \subseteq R.$$

All first-order definable tenses obey Quality; but the converse is not true.

Against this background, a common mathematical condition (already prominent in Chapter 3) turns out to characterize the Priorean basic tenses, viz. *Continuity*:

$$f\Big(\bigcup_i A_i\Big) = \bigcup_i f(A_i), \quad \text{for all families } \{A_i \mid i \in I\}, A_i \subseteq R.$$

THEOREM. The only IR-tenses satisfying Quality and Continuity are precisely those defined by the schema 'f is some union of *pa, pr, fu*'; where

$$pa(A) = \{y \in R \mid \text{for some } x \in A, x < y\}$$
$$pr(A) = A$$
$$fu(A) = \{y \in R \mid \text{for some } x \in A, y < x\}.$$

Proof. By Quality, a tense f can only have a limited range of choices for each singleton set $\{x\}$ $(x \in R)$. $f(\{x\})$ either contains or is disjoint from each of the three regions $\{y \in R \mid y < x\}, \{x\}, \{y \in R \mid x < y\}$ (see Figure 23).

Fig. 23.

For instance, if $f(\{x\})$ includes any $y_1 < x$, then it will also include all other $y_2 < x$. (To see this, choose any IR-automorphism sending y_1 to y_2, by suitable '*stretching*', leaving x and all $y > x$ fixed.) Moreover, f makes the same choice for all $x \in R$, again by Quality. (This time, one uses the fact that *translations* from, say, x_1 to x_2 are IR-automorphisms.) Thus, Quality already enforces strong uniformity of behaviour, in terms of the above three notions. Finally, Continuity implies that these pointwise choices determine the values of f completely at all larger arguments A, as $f(A) = \bigcup_{x \in A} f(\{x\})$. □

Notably, the progressive tense fails the Continuity test: it assigns the empty set to each singleton, and yet whole intervals are assigned their topological interior. Thus, to obtain a more liberal *hierarchy* of possible tenses, Continuity is to be relaxed progressively. Still, this much of its motivation remains clearly valid, also for the progressive: membership of $f(A)$ must only depend, 'locally', on some 'episode' in A. This is expressed in the following condition of *Bounded Continuity*:

$y \in f(A)$ iff $y \in f(A_0)$,
for some *bounded convex* subset A_0 of A.

A generalization of the above argument gives us all qualitative bounded continuous tenses — being all finite unions of

pa, pr, fu (as above)
int, with $int(A) = \{y \in R \mid y$ lies in some open A-interval$\}$
le, with $le(A) = \{y \in R \mid y$ is a left-most boundary of some open A-interval$\}$
ri, with $ri(A)$ the set of similar right-most boundaries;

as well as some variants:

fu^*, pa^*, with $fu^*(A) = \{y \in R \mid$ open A-interval lies to the right of $y\}$ and $pa^*(A)$ its past dual,
le^+, ri^+, with $le^+(A) = le(A) \cap A$, $ri^+(A) = ri(A) \cap A$.

This classification is proved in van Benthem (1983c). The reader may form a fair impression of its proof by observing that, this time, the options are to be charted for all 'pilot cases' of bounded convex intervals: i.e., singletons as well as open, half-open and closed intervals. Again, Quality makes f accept or reject obvious regions in their entirety.

There is still a linguistic aspect to this outcome, as progressives and 'boundary tenses' (*just arrived, about to go*) are indeed about the next level of common occurrence.

Relaxing Bounded Continuity even further, a more liberal hierarchy arises, whose natural endpoint seems to be the notion of mere *Monotonicity* for functions f:

$$\text{if} \quad A \subseteq B, \quad \text{then} \quad f(A) \subseteq f(B).$$

Note that this is a consequence of both Continuity and Bounded Continuity. By then, infinitely many tenses will qualify: for instance, all first-order definable ones having only positive syntactic occurrences of their unary predicate parameter.

Of course, the preceding analysis is still tied up with (dense, continuous) real time. It therefore becomes of interest to extend the analysis to other major temporal structures. For instance, on discrete *integer* time, the given characterizations are no longer valid. In particular, many other 'tenses' besides the three basic ones satisfy both AUT and CONT, such as *to-morrow, yesterday*. (Technically, the reason is that the integers ZZ have far fewer automorphisms than the reals IR. In a sense, the topological order on ZZ also encodes metric structure.) So, one can search for additional constraints characterizing the basic tenses in discrete time, perhaps even: on all linear orders. Another line worth investigating is the behaviour of tenses on temporal *interval* structures, rather than point-based ones (cf. van Benthem, 1982b).

Once we have this restricted view of tense, as a highly constrained class of operations on denotations of propositions, it seems reasonable to extend it to these propositions themselves. After all, although all propositions denote sets of points in time (in the present limited setting), there is no need to endorse the *converse*: that all such sets correspond to propositions. And indeed, the more esoteric mathematical subsets of IR (say, all irrational points in time) hardly seem to qualify as the life times of natural language propositions. Thus, we can profitably study certain restricted ranges of propositional denotations. Prime examples are the *convex* ('uninterrupted') sets of reals, or at most, the finite unions of the latter (when allowing 'repetitive' events). In such a setting, the earlier questions of definability, but also traditional concerns about axiomatizing complete temporal logics, assume new forms. Thus, much of traditional tense logic will have to be rethought.

5.2. A STRUCTURAL VIEW OF MODALITY

There is a certain formal similarity between the cases of tense and modality. So, we can repeat the above analysis of temporal precedence orders with *accessibility patterns* of possible worlds. Nevertheless, the situation is not wholly analogous, because accessibility seems to be more of a postulated construct than an independently given notion. Hence, one should proceed more cautiously, and more abstractly — as in the earlier chapter on conditionals.

Consider modalities as unary operations m on sets of possible worlds (denotations of propositions). As a point of departure, attention will be restricted to *quantitative* operations, insensitive to world patterning. Of the basic postulates presented for conditionals, none seem directly applicable, except for *Variety* — in the sense that, on non-empty universes E, modalities should not be constant.

As for more special features of modality, one obvious source is the so-called 'minimal modal logic' (cf. Chellas, 1980), whose central axiom is *distribution* of possibility over disjunction: $\Diamond(\varphi \vee \psi) \leftrightarrow \Diamond\varphi \vee \Diamond\psi$. For set functions m, this becomes the following principle:

$$m(A \cup B) = m(A) \cup m(B), \quad \text{for all } A, B \subseteq E.$$

This is a finite form of the earlier *Continuity* — and in fact, the motivation of the distribution axiom would seem to support the latter in its full strength.

In addition, there is the non-controversial T-axiom $\varphi \rightarrow \Diamond\varphi$ ('actual truth implies possible truth'), whose semantic effect is to make modalities *extravert* (cf. again Chapter 3):

$$m(A) \supseteq A, \quad \text{for all } A \subseteq E.$$

Now, which set operations m pass all the above tests? Consider any non-empty *finite* universe E of possible worlds. As in Chapter 3, m must assign, for each argument A, some union of A, $E - A$. Moreover, as $A \subseteq m(A)$, this can only be either A itself, or all of E. Also, $m(\emptyset)$ must be $\emptyset$: otherwise $m(\emptyset) = E$, and hence $m(A) = E$ for all A (by Monotonicity, a consequence of Continuity), which would make m a constant function, contradicting Variety. Then, as before, consider m's behaviour on singletons. Either, it is the identity, and m becomes the identity throughout, or it assigns E, and m assigns E everywhere, except for its empty argument. Thus, as in Chapter 3, the quantitative setting produces only the most obvious 'classical' possibilities:

THEOREM. The only quantitative T-modalities are the trivial one ('plain truth') together with S5-modality.

These arguments become slightly more interesting for the minimal logic without the T-axiom. (One curious alternative appears, viz. having $m(\emptyset) = \emptyset$, $m(\{a\}) = E - \{a\}$, and $m(A) = E$ for all sets $A \subseteq E$ having two or more elements.) Even so, the need for a more generous policy is obvious, if one is to obtain modalities weaker than S5. As in Chapter 4, here is where an accessibility pattern R is to be introduced, relaxing Quantity to *Quality*, i.e., invariance for R-automorphisms of the universe. Then, in line with the previous section, special accessibility structures may be investigated — say, the infinite binary tree of branching possibilities — pursuing the same questions as for tenses in Section 5.1. But actually, with the above abstract manner of introducing the relation R, it may be more appropriate to proceed differently.

For instance, one might assume that the suitable R-patterns are merely defined by some class of first-order conditions on R; a not unrealistic assumption in modal logic. Or, there might even be a first-order theory μ in R, X, Y, representing our semantic analysis of R and m, such that $\langle E, R, A, B\rangle \vDash \mu$ if and only if $B = m_{\langle E, R\rangle}(A)$. In the latter case, μ *defines* Y *implicitly*, given R, X, in the usual semantic sense. Hence, by Beth's Definability Theorem, there must be some *explicit* first-order *definition* of the form

$$\forall z(Yz \leftrightarrow \tau(z, X, R)).$$

In other words, our modal operator m has a first-order *truth definition* in the theory μ. And then, we have returned to an earlier theme, being the range of possible truth definitions — and questions arise similar to those in Section 4.8. Notably, which syntactic restrictions on τ are imposed by the laws of the minimal modal logic? This particular question will be answered in Chapter 9, be it with a somewhat different motivation.

Here is a different perspective to conclude with. There is also a more direct way of construing the alternative relation R on our universes, as arising out of the minimal axioms. The following result may be found in Segerberg (1971):

THEOREM. A modal function m satisfies the minimal modal logic on a universe E iff there exists some binary relation R on E satisfying

$$m(A) = \{y \in E \mid \text{for some } x \in A, Ryx\}, \quad \text{for all } A \subseteq E.$$

Proof. From right to left, this is standard. Conversely, set

$$Ruv \Leftrightarrow_{\text{def}} u \in m(\{v\}).$$

By Continuity then, $m(A) = \bigcup_{x \in A} m(\{x\})$, and the above identity follows. □

Incidentally, this point of view is not restricted to modality. Returning to the previous case of conditionals, the same kind of analysis yields the following result. A natural hierarchy between occasions may be defined by the stipulation

$$x \leqslant y \quad \text{iff} \quad if\{x, y\}\{y\}.$$

THEOREM. A conditional generalized quantifier relation *if* satisfies the basic subjunctive conditional logic iff

— its induced relation $\leqslant$ is a partial order,

— $if AB$ is equivalent to $\forall x \in A \exists y \in A \cap B\, x \leqslant y$.

A proof may be found in van Benthem (1985a).

Actually, on infinite universes, this representation becomes more complicated, involving an equivalence between $if AB$ and

$$\forall x \in A \exists y \in A(x \leqslant y \wedge \forall z \in A(y \leqslant z \rightarrow z \in B)).$$

Quite similar considerations may also be used to obtain a *temporal* precedence relation from a basic conditional *if X, then* (*next*) *Y*. (In this case, the pervasive condition of Conservativity will not be met: but, this creates no insuperable problems.)

Once a pattern of possible worlds has been introduced in accordance with the basic axioms, the effects of additional modal axioms will show up in further restrictions on the alternative relation. For instance, the characteristic axiom of T will now make R *reflexive*, and, e.g., that of S4 ($\Diamond\Diamond\varphi \rightarrow \Diamond\varphi$) enforces precisely *transitivity*. For a systematic study of such correspondences, see van Benthem (1984a).

These examples will suffice to show how we propose to enter the world of intensionality with the techniques of this book.

CHAPTER 6

NATURAL LOGIC

One recurrent 'underground' ideal in semantics is the construction of a *natural logic*, being a system of reasoning based directly on linguistic form, rather than logical artefacts. Thus, the ideal division of labour would come about, with the logician borrowing the linguist's grammatical analysis in his account of inference, without having to set up his own shop for producing 'logical forms'. Some strands in this book tend in such a direction — and we shall examine the prospects for a natural logic in the light of the preceding chapters.

The search for a natural logic is often viewed as a reactionary move, betraying a misplaced nostalgia for the age of *traditional logic*, which was in closer contact with natural language surface forms than its modern Fregean successor. But then, many notions of the generalized quantifier framework are indeed closely related to central concepts in pre-Fregean syllogistic. And so, to put the above enterprise in perspective, we shall consider some connections with traditional logic first.

6.1. ANALYZING TRADITIONAL LOGIC

The main framework of Chapters 1, 2, with its relational format QXY for basic quantified statements, resembles what is called the *Two Term Theory* of predication, which dominated logic until the nineteenth century. Given this general analogy, various more specific notions and themes turn out to be related. For instance, some of the basic inferences in traditional logic, such as principles of *Conversion* ('symmetry') or many *syllogisms*, have played an essential role in Sections 1.5., 2.6. Also very conspicuously, the classical *Square of Opposition* appeared in various definability results.

Of course, thanks to modern Fregean standards, such notions can often be treated somewhat more systematically in the present setting. For instance, the Square can be viewed as an instance of the general scheme of Figure 24 (with $Q = all$):

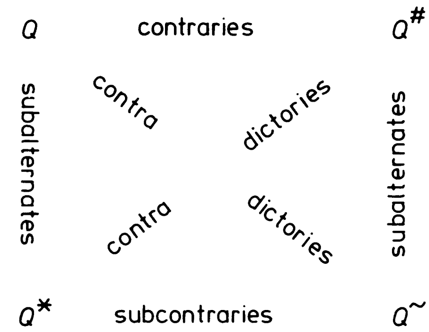

Fig. 24.

Here the following formal operations on quantifiers occur:

$Q^{\sim}AB$ iff *not* QAB	(*outer negation*)
$Q^{\#}AB$ iff $QA(A-B)$	(*inner negation*)
$Q^{*}AB$ iff *not* $Q^{\#}AB$	(*dual*)

For Q = *all*, these become the other three corners of the Square: *not all, no* and *some*. This general pattern is complete, in the following sense. Repeating these operations yields no new quantifiers: indeed, under composition, they form the 'Klein 4-Group' of Figure 25.

	id	~	#	*
id	id	~	#	*
~	~	id	*	#
#	#	*	id	~
*	*	#	~	id

Fig. 25.

For a very general linguistic use of this schema, cf. Löbner (1984), who presents squares of determiners, connectives, temporal adverbs, intensional verbs, etc. — and compare also the 'propositional' Square in Section 3.1.

Perhaps the most striking analogy has to do with *Monotonicity*.

Often considered a latter-day logical discovery, this notion lay at the heart of the syllogistic. Notably, downward monotonicity of an expression is what traditional logicians would call *distributed* occurrence of predicates: occurrences where the containing statement is true about 'all of the predicate' (i.e., about all parts of its denotation). Likewise, upward monotonicity reflects a central classical type of argument called the *Dictum De Omni*:

'whatever is true of every X is true of what is X'.

What this terse formulation comes to, in our terminology, is this. If every X is Y, and 'X' occurs in upward monotone position in some statement ... X ... , then that same statement holds for Y: ... Y .... The Dictum De Omni has been regarded as the principle par excellence governing syllogistic reasoning. And, its influence even extends beyond the latter into the logic of relations (cf. Sommers, 1982). For instance, De Morgan's famous non-syllogistic relational example *All horses are animals. Therefore, all horse tails are animal tails* can be subsumed under it (by considering the true statement 'all horse tails are *horse* tails'). Thus, monotonicity seems a promising key notion for a natural logic reviving classical ideals (cf. Section 6.2. below).

Thanks to these connections, we can also analyze some of the basic principles of traditional logic in our generalized quantifier theory — an idea inspired by Van Eyck (1985).

One pervasive presupposition of classical logic is *Existential Import*: statements QAB are only considered for *non-empty* predicates A, B (I). Once this is assumed, the Square of Opposition yields one obvious inferential connection: the basic quantifier Q (= *all*) implies its subalternate or *dual* (i.e., *some*) (II). (The other 'subalternacy' relation, that from $Q^{\#}$ to $Q^{\sim}$, is already a consequence of II — by the definition of $^{\#}$, $^{\sim}$.) As a third basic classical principle, we select one form of *Conversion*, viz. that for the opposite quantifier: $Q^{\#}AB$ iff $Q^{\#}BA$, for all A, B (III). (Again, conversion for the dual quantifier Q^{*} is already a consequence of III.) Note that these conditions represent special features of the quantifiers in the classical Square, which need not hold for other quartets of expressions related by $^{\#}$, $^{\sim}$, *. Indeed, within the field of logical quantifiers (as defined in Chapter 2), these three syllogistic principles determine precisely the quantifier Q = *all* and its descendants.

THEOREM. The syllogistic principles I, II, III determine exactly the intended interpretation of the Square of Opposition.

Proof. First, in the Tree of Numbers (Section 2.2.), the four intended quantifiers have aceptance patterns satisfying these three conditions. Conversely, consider any non-empty quantifier pattern Q. The geometric constraints corresponding to the above three principles are as follows. I merely amounts to disregarding the top node (0, 0) — leaving a starting row (1, 0)—(0, 1). As for II, note that $Q^{\sim}$ denotes the complement of Q, $Q^{\#}$ its mirror image along the axis (n, n) $(n \geqslant 1)$, and Q^{*} the latter's complement again. In particular, then, no central position (n, n) itself $(n \geqslant 1)$ can belong to Q. (Otherwise, $(n, n) \in Q$, $(n, n) \notin Q^{*}$: contradicting II.) Finally, III expresses symmetry of $Q^{\#}$. By a standard argument (cf. Section 1.5.), this amounts to dependence on the right-hand number only: for any b, $(a, b) \in Q^{\#}$ iff $(a', b) \in Q^{\#}$, for all a, a'. Geometrically, then, Q itself must consist of entire north-west\south-east diagonals. The final argument becomes this. If Q is non-empty, it must also contain a diagonal like above, of the form $\{(a, b) \mid b \geqslant 0\}$. Now, if $a \neq 0$, this set will contain the pair (a, a) with $a \geqslant 1$: quod non. Therefore, the only possibility is the right-most edge of the tree — which is the pattern of the quantifier *all*. □

This result expresses, in a sense, that the Syllogistic is complete with respect to its subject matter.

The above is a 'denotational' analysis, in terms of structural constraints. There remains the more traditional viewpoint of the description of valid inference, without immediate semantic reflection. The perspective of this book throws some new light on this old task too.

6.2. LOGIC BASED ON GRAMMATICAL FORM

What is needed to set up a natural logic with ordinary grammatical structures as its vehicle? An instructive example is provided by monotonicity reasoning, a central feature in traditional logic, as we have seen.

One important idea, implicit in the formulation of the Dictum De Omni, is that inference rules can be *global*, operating at statement level, without presupposing any proof-theoretic fine structure analysis. At least in an intuitive psychological sense, this seems realistic. For the present case, this means the following. There is a general *inclusion* or *implication order* $\sqsubseteq$ on arbitrary denotations (cf. Chapter 3), generalizing specific notions of consequence, such as set inclusion for predicates (*all XY*) or ordinary $\leqslant$ on truth values (material *if, then*). Other examples are implications between adverbs (*very loudly* $\sqsubseteq$ *loudly*), adjectives (*red and angry* $\sqsubseteq$ *red, Dutch* $\sqsubseteq$ *European*), or determiners (*at least five* $\sqsubseteq$ *at least three*). Now, occurrences of expressions in other expressions can be *inferentially sensitive* to this general implication relation. More precisely, let us call an occurrence of an expression X in an expression $\varphi = \ldots X \ldots$ *positive* if $[\![X]\!] \sqsubseteq [\![Y]\!]$ implies $[\![\ldots X \ldots]\!] \sqsubseteq [\![\ldots Y \ldots]\!]$ (notation: $\ldots \overset{+}{X} \ldots$). An occurrence is *negative* if the inverse correspondence holds (notation: $\ldots \overset{-}{X} \ldots$).

When read in dynamic terms, positive occurrence of X in φ means both: increasing the denotation of X will at most increase that of φ/decreasing the denotation of X will at most decrease that of φ. So, the main point of the distinction +/− is not any privileged direction, but that of direct versus inverse correlation.

Now, in order to exploit positive and negative occurrences in inference, they must be syntactically available. Thus, this information must have been built in during the very process of sentence formation. (Actually, there are various more theoretical questions concerning the connection between the above semantic and other, more syntactic ways of defining 'positive'/'negative' occurrences. Compare also Section 2.5.)

Inference marking of inferentially sensitive positions requires several phases. First, the 'logic' of certain lexical items consists (partly) in their +/− effects on their linguistic environment. Then, the +/− effects of various grammatical rules have to be taken into account. And finally, the +/− effects of nesting of such markers are to be established by some rule of calculation. In this way, sentences are constructed (or understood) with marked inferentially sensitive positions of key items. This process may be spelt out as an algorithm on phrase structure trees. Here, we shall only outline the construction.

First of all, a specific grammar is needed — here: a simple set of phrase structure rewriting rules. These are then to be marked for their +/− behaviour in a suitable manner. For the basic rules of Chapter 1, this works out as follows:

$$\mathrm{S} \Rightarrow \overset{+}{\mathrm{NP}}\ \mathrm{VP}.$$

The reason is this. In 'subordinating' phrase structure rules, the *functor* gets a + marking. Similarly,

$$\mathrm{NP} \Rightarrow \overset{+}{\mathrm{Det}}\ \mathrm{N}.$$

To be more precise, this rule says that, within any subexpression NP, its Det occurs in positive position. This is 'local' information, as the NP itself may be deeply embedded — which may affect the final position of Det in the whole expression: itself +, − or neutral.

Another, entirely trivial case are the rules

$$\mathrm{NP} \Rightarrow \overset{+}{\mathrm{PN}} \text{ (proper names)}$$
$$\mathrm{VP} \Rightarrow \overset{+}{\mathrm{V}} \text{ (intransitive verbs).}$$

Transitive verbs work along the earlier line:

$$\mathrm{VP} \Rightarrow \mathrm{V}\ \overset{+}{\mathrm{NP}}$$

Note the marking here. The V-position is (usually) governed by the direct object NP, rather than vice versa: witness the three cases $\overset{+}{\textit{love}}$ *a dog*, $\overset{-}{\textit{love}}$ *no dog* and *love one dog.*

Another general pattern is found with more 'coordinating' phrase structure rules. These are treated *conjunctively*:

$$\mathrm{N} \Rightarrow \overset{+}{\mathrm{Adj}}\ \overset{+}{\mathrm{N}} \text{ (intersective adjectives)}$$
$$\mathrm{N} \Rightarrow \overset{+}{\mathrm{N}}\ \overset{+}{\mathrm{R}} \text{ (relative clauses)}$$

Finally, here is one syncategorematic rule:

$$\mathrm{R} \Rightarrow \textit{who}\ \overset{+}{\mathrm{VP}}.$$

Moreover, monotonicity effects of specific lexical items are to be acknowledged. For instance, some determiners have the double monotonicity behaviour described in Section 1.4. To some extent, this may even exhaust their 'logic', witness the definability theorem for the Square of Opposition. Connectives also have their expected behaviour — here displayed in a convenient syncategorematic form:

$$\mathrm{V} \Rightarrow \overset{+}{\mathrm{V}}\ \textit{and}\ \overset{+}{\mathrm{V}} \quad \text{(or even } \overset{+}{\mathrm{V}}\ \overset{+}{\textit{and}}\ \overset{+}{\mathrm{V}}\text{).}$$

But also, *or* gets the same marking: monotonicity does not exhaust the 'logic' here. Then, of course, negation reverses direction. Here is a sample rule:

$$\mathrm{V} \Rightarrow \textit{not}\ \overset{-}{\mathrm{V}}$$

(Or even: $\overset{+}{\textit{not}}\ \overset{-}{\mathrm{V}}$. A bit clumsily, '(does) *not bow*' implies '(does) *not or* *hardly bow with a smile*'.) One last example shows one subordinating connective of Chapter 4:

$$S \Rightarrow \textit{if}\ \overset{-}{S}, \overset{+}{S}.$$

These lexical indications are not precise as they stand — because one item's sphere of action may extend beyond its immediate phrase structure rule (witness the determiners). This aspect works out somewhat more smoothly in a *categorial grammar*, where lexical items carry type information encoding several arguments, which can be marked already. E.g., *all* could be construed as having type $((\overset{-}{e}, t), ((\overset{+}{e}, t), t))$.

Finally, the rule of *over-all calculation* for nesting is simply algebraic multiplication:

$$++=+,\ +-=-,\ -+=-,\ --=+.$$

Putting all this together, concrete sentences can be inferentially marked. Some examples are shown in Figure 26.

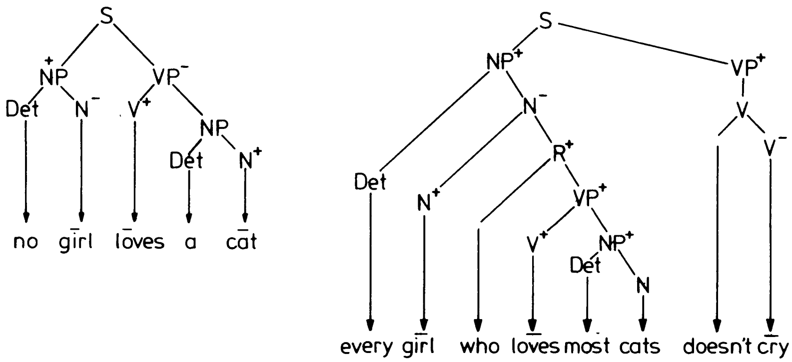

Fig. 26.

The first example is straightforward — and its predictions are borne out in practice. The sentence indeed implies: that no *pretty* girl loves a cat, no girl *desperately* loves a cat, no girl loves a *curly* cat. These inferences can become blocked in quite similar cases. For instance, the determiner *one* has no monotonicity effects in either argument, and hence it stops the + markers on *$\overset{+}{\textit{loves}}$ a $\overset{+}{\textit{cat}}$* from surfacing: *one girl loves a cat.* Another kind of barrier arises in cases such as *a $\overset{+}{\textit{cat}}$ and* *no* $\overset{-}{\textit{dog}}$ *appeared*, where the two determiners disagree on the marking of the verb phrase. As long as they work in tandem, however, markings do percolate to the surface: *a* $\overset{+}{\textit{cat}}$ *and all* $\overset{-}{\textit{dogs}}$ $\overset{+}{\textit{appeared}}$. The second example introduces no new points of principle.

Another way of computing the first example uses categorial grammar again. *No* will have type $(\overset{-}{p}, (\overset{-}{p}, t))$ (with $p = (e, t)$), *girl* has p, *loves* (e, p), *a* $(\overset{+}{p}, (\overset{+}{p}, t))$, *cat* p. Then, function/argument combinations work out as follows: *no* $\overset{-}{\textit{girl}}$ $(\overset{-}{p}, t)$, *a* $\overset{+}{\textit{cat}}$ $(\overset{+}{p}, t)$, $\overset{+}{\textit{loves}}$ *a* $\overset{+}{\textit{cat}}$ p. In the latter case, the type for *a cat* has been 'lifted' to obtain $((e, p), (e, t))$: a procedure described in full detail in Chapter 7. Then finally, *no girl loves a cat* receives type t. A further analysis of inference marking in a flexible categorial grammar is given in van Benthem (1986a).

Now, how much inference is accounted for by the above mechanism? We consider a few cases.

Example (syllogistic reasoning). Of the four valid syllogisms in the 'First Figure', three are immediate consequences of the above. These are *Barbara* (*all MP, all SM/ all SP*), *Celarent* (*no MP, all SM/ no SP*) and *Darii* (*all MP, Some SM/ some SP*). The fourth, *Ferio* (*no MP, some SM/ not all SP*) requires some additional reasoning, however.

More specifically, Barbara goes like this. The first premise *all MP* serves as a '*cue*' for the presence of an $\sqsubseteq$-relation. Then the second premise is marked thus: *all* $S\overset{+}{M}$ — and the conclusion follows. With Ferio, one can argue: *no MP, some* $\overset{+}{S}M$, *all SP* — *no MP, some PM* — *no PM, some PM* — contradiction. Thus, in addition to Monotonicity, Conversion is needed, as well as propositional Reductio ad Absurdum. An alternative would be to read *no MP* as a cue for $[\![M]\!] \sqsubseteq [\![\textit{non-}P]\!]$ or $[\![P]\!] \sqsubseteq [\![\textit{non-}M]\!]$, turning the second premise into *some S non-P*, which reduces to *not all SP*.

Example (propositional reasoning). Many well-known propositional inferences revolve around monotonicity. Examples are Modus Tollens: $P \rightarrow Q, \neg \overset{-}{Q}/\neg P$, and Constructive Dilemma: $P \rightarrow R$, $Q \rightarrow R$, $\overset{+}{P} \vee \overset{+}{Q}$ $/R$. Actually, in the latter case, an additional principle is used, viz. the Boolean identity $R \vee R = R$. Further examples require more formal tricks, such as $\neg(P \wedge \overset{-}{Q})$, Q — $\neg(P \wedge \overset{-}{Q})$, $\textit{true} \rightarrow Q$ — $\neg(P \wedge \textit{true})$ — $\neg P$. And, many elementary steps, such as $P \wedge Q/P$, do not fit this mould at all.

Still, monotonicity is involved in many propositional inferences, especially less elementary, and more interesting ones. (Compare also the earlier point about 'global' inference.) Here is one more general example. Let there be two premises, of any complexity, which are 'linked' in the following sense: one has a positive occurrence of P, the other a negative one. Then, intuitively, there ought to be some typical combination of the two, via the 'bridge' P — as in $P \rightarrow \overset{+}{Q}$, $\overset{-}{Q} \rightarrow R/P \rightarrow R$. And indeed, there is. Observe that $\alpha(\overset{+}{P})$ is equivalent to $\alpha(\textit{true}) \wedge (P \vee \alpha(\textit{false}))$, $\beta(\overset{-}{P})$ to $\beta(\textit{false}) \wedge (P \rightarrow \beta(\textit{true}))$. Therefore, we may draw the following inferences from our premises singly: $\alpha(\textit{true})$, $\beta(\textit{false})$; as well as $\alpha(\textit{false}) \vee \beta(\textit{true})$ from the two combined. For instance, with some obvious calculation on *true, false*, one has

$$\neg(P \wedge \neg(\overset{+}{Q} \wedge R)), (P \wedge \overset{-}{Q}) \rightarrow R/\neg(P \wedge \neg R), \textit{true},$$
$$\neg P \vee (P \rightarrow R); \quad \text{i.e.,}\ P \rightarrow R.$$

These are still schematic applications, to formal languages. Natural language examples are to be obtained using the above phrase structure grammar. There are some interesting discrepancies between the two cases. Note, for instance, that the syllogism Barbara cannot even be expressed directly in our fragment. Circumlocution is required, using participles or relative clauses: 'All *girls wept*. All *weepers/those who wept, were afraid*. Therefore, all *girls were afraid*'. Likewise, the earlier fundamental principle of Conservativity requires the use of adjectives or relative clauses: 'all *girls wept*' if and only if 'all *girls were weeping girls/were girls who wept*'. So, applying monotonicity patterns in natural language takes more than mere substitution: viz. a judicious look through linguistic wrappings. The moral is that inference is an art of *omission* as well as *combination*.

As was clear in the above examples, monotonicity reasoning is to be augmented with various other types to describe our natural logic. The preceding chapters suggest a systematic policy here. First, there are general principles deriving from the implication relation $\sqsubseteq$ as such. It is *transitive* and *reflexive*, and hence so are its cue words (*all*, material *if, then*). Moreover, this order forms a *lattice* with greatest lower bounds (*and*) and lowest upper bounds (*or*), with their usual properties. (This explains why these connectives are so ubiquitous.) Perhaps also, negation should be included at this level — making all axioms of *Boolean Algebra* available (i.e., the full propositional calculus). Next, various categories of expression may contribute their own broad inferences. One notable example is *Conservativity* for determiners. With each such addition, inferential power may increase in unexpected ways.

Example. Conjunction follows from Monotonicity and Conservativity:

$$\dfrac{\dfrac{all\,XY}{all\,X(Y\,and\,\overset{+}{X})} \qquad all\,XZ}{all\,X(Y\,and\,Z)}$$

Then, specific lexical items may have their particular base logic — such as Conversion (symmetry) for *some, no.* And also, we need 'definitional' connections, such as the equivalence between *not every* and *some . . . not*, or *not a* and *no.*

Example. Some quantifier interchanges:

Some girl loves no cat, some girl loves not a cat, not every girl loves $\bar{a}$ *cat.* Here, we observe a marking, and use the implication $[\![every]\!] \sqsubseteq [\![a]\!]$ (assuming Existential Import) to obtain *not every girl loves every cat.*

Even in this sketchy form, the mechanism of a modest natural logic will have become clear.

Of course, natural language fragments like the above also have a 'Fregean logic', through the usual transcription into predicate-logical form. Thus, the vexed controversy between classical and modern logic can be given an exact content. How do the two logics compare for various fragments of natural language? In particular, when is the natural logic of some fragment *complete* with respect to its Fregean rival? In this way, a philosophical quarrel becomes a matter of research.

It remains to be noted that there is quite a tradition of attempts at rehabilitating pre-Fregean logic. One very interesting school devoted to this program is that of Sommers (1982) and his followers. (See especially Englebretsen (1981) for an attempt at formulating a logic on some sort of linguistic form.)

Roughly speaking (it is rather hard to fathom), Sommers' system consists of the following ingredients: a monotonicity calculus, algebraic working rules for connectives, and a logic for the operators of Quine (1966) (cf. Section 3.1.) handling permutation, identification and projection on argument places of predicates. Thus, quite different mechanisms are to be added to the simple monotonicity calculus if the full power of Fregean logic is to be approximated. This points at an interesting moral. From the present point of view, 'predicate logic' is a conglomerate of quite diverse mechanisms of inference: some relatively elementary, taking a free ride on the grammatical construction of the sentence (as in the above), others requiring inspection of more complex structures at a late stage of utterance. Notably, manipulation of inferences in the presence of anaphoric connections may be such a 'higher' facility. The latter comes in which we have to recognize, say, Modus Ponens in such varying patterns as the following:

$$\frac{\text{if Julia tries, she will succeed} \qquad \text{Julia tries}}{\text{she will succeed}}$$

$$\frac{\text{if Julia tries, she will succeed} \qquad \text{she tries}}{\text{Julia will succeed}}$$

$$\frac{\text{if she tries, Julia will succeed} \qquad \text{Julia tries}}{\text{she will succeed.}}$$

Clearly, such inferences are pieces of *discourse*, and our processing of them requires inferential mechanisms at text level.

Thus, 'logic' is becoming an ever more elusive (though also omnipresent) notion in this book: occurring across all linguistic categories (see Chapter 3), and at various levels of linguistic processing, as we have seen here.

PART II

DYNAMICS OF INTERPRETATION

CHAPTER 7

CATEGORIAL GRAMMAR

Sentences of natural language may be analyzed as having a function/argument structure, as was originally observed by Frege. For instance, in *Julia weeps*, the verb acts as a functor assigning a truth value to entities mentioned in the proper name position. But also, e.g., in *Julia weeps bitterly*, the adverb may be regarded as denoting a function from verbs to (complex) verbs. Thus, the interpretation of natural language expressions involves a hierarchy of functions — and the task of a categorial grammar is to assign suitable types of function ('categories') to linguistic expressions, so as to make the puzzle 'fit'. This chapter is concerned with some logical aspects of the categorial mode of description.

7.1. THE ORIGINAL VERSION

As in Chapter 3, there are two *basic types* t ('truth value') and e ('entity'), with complex types generated recursively by the rule 'if a, b are types, then so is (a, b)'. Expressions of type (a, b) denote functions mapping type a-denotations to type b-denotations. Simple expressions of natural language will now be assigned some type, while sequences of expressions may or may not receive a type through 'functional applications' of their component types. This is the basic idea of Ajdukiewicz' 'categorial grammar'. (In general, other basic types could occur too; and also, in case of ambiguity, simple expressions could be assigned multiple types.)

Here are some samples of type assignment. Negation (*not*) takes sentences to their negations: type (t, t). Conjunction (*and*) conjoins two sentences to form a new one: type $(t, (t, t))$. The above intransitive verb *weep* had type (e, t), while a transitive verb (*love*) takes two entities to a truth value: $(e, (e, t))$. The above adverb (*bitterly*) would have type $((e, t), (e, t))$. Finally, in general, expressions in subject position may be more complex than just proper names (witness *every lady weeps*), in which case they receive type $((e, t), t)$. I.e., in general, subjects operate on predicates, rather than vice versa. Other important types are found in the table of Section 3.3.

Now, evaluation takes place as in the following example.

Not (*every lady*) *loves Romeo*:
$(t, t) \quad ((e, t), t) \quad (e, (e, t)) \quad e.$

The following sequence of applications shows this to be a sentence:

$$
\begin{array}{lll}
(t, t) & ((e, t), t) & (e, t) \\
(t, t) & t & \\
t. & &
\end{array}
$$

Evidently, not every sequence evaluates to a sentence, witness *every lady Romeo*. On the other hand, sequences which do have an outcome t (or any other single type) may arrive there in different ways; sometimes (not always) corresponding to non-equivalent 'readings'. For instance, consider

Not (*a lady wept*) *all that year*:
$(t, t) \quad t \quad (t, t).$

One analysis combines the left-most types first (*all that year* (*not* . . .)), the other starts with the right-most types, producing the different scope reading *not* (*all that year* . . .).

As it stands, this is still a very crude way of linguistic description. For instance, functional application is allowed both ways, whereas natural language has constraints of 'directionality'. Standard categorial grammars account for this by asymmetric encoding of arguments, with the familiar slashes $a\backslash b$, b/a introduced by Bar-Hillel to indicate left- or right-searching operators. Important as it is, directionality — and indeed descriptive 'fit' — will not be a major concern in this chapter. Throughout, a simple non-directed version will be used as a vehicle for exposition. The insights obtained in this way will probably generalize to the more sensitive categorial grammars employed in actual linguistic description.

Even so, the above simple calculus already has its problems; of which we mention one. Notably, there remain certain 'missing readings', as with (*Every crook*) *fears* (*some detective*): $((e, t), t) \quad (e, (e, t)) \quad ((e, t), t)$. This sequence cannot be evaluated to the sentence type t as it stands. Perhaps the best-known remedy is that of Montague Grammar: re-categorize transitive verbs as $(((e, t), t), (e, t))$. (There still remains the question of obtaining enough different readings for this sentence: cf. Section 7.3.) The price to be paid for this solution is the failure of the basic sentence *Julia loves Romeo*: $e \quad (((e, t), t), (e, t)) \quad e$. To make the latter come out right again, proper names are to be 'lifted' to type $((e, t), t)$ as well. A more attractive general solution to the problem of transitive verbs will be presented below.

By and large, categorial grammar has remained an undercurrent in linguistics, much as the related type-theoretical approach has been in logic. Nevertheless, its more flexible version to be introduced now has additional linguistical and logical interest.

7.2. MORE FLEXIBLE VERSIONS

Expressions of natural language need not stay in their basic category, but can assume higher types when desired for the purpose of interpretation. For instance, the negation particle *not* can also behave as predicate negation (*does not cry*) or subject negation (*not every lady*). Still, there is a system to such type changes: not anything goes. Thus, Geach (1972) has proposed the following rule for deriving higher occurrences from the basic one:

G if an expression occurs in type (a, b), then it may also occur in any type $((c, a), (c, b))$ (for arbitrary c).

This proposal has some pleasant effects. First, it accounts for the various uses of negation, starting from one basic assignment (t, t) (sentential negation), such as intransitive verb negation: $(e, t), (e, t)$, and then, iterating, also transitive verb negation: $(e, (e, t)), (e, (e, t))$, etc. In the original grammar, a host of initial types would have to be postulated for the lexical item *not*, without capturing its essential unity of meaning.

Next, even with their original intended type, transitive verbs now allow for the earlier complex direct objects. For instance, *Every crook fears some detective* may be evaluated as follows:

$$
\begin{array}{lll}
((e, t), t) & (e, (e, t)) & ((e, t), t) \\
((e, t), t) & (e, (e, t)) & (((e, (e, t)), (e, t))\ (!) \\
((e, t), t) & (e, t) & \\
t. & &
\end{array}
$$

Also, additional readings appear. Until now, (*no lady*) *wept* (*all that year*) could only be read as follows: $((e, t), t) \quad (e, t) \quad (t, t)$, and then $t \quad (t, t)$, and finally t: i.e., 'all that year (no lady)'. But now, the inverse scope reading 'no lady (all that year)' may also be obtained:

$$\begin{array}{lll}
((e, t), t) & (e, t) & (t, t) \\
((e, t), t) & (e, t) & (e, t), (e, t)\ (!) \\
((e, t), t) & (e, t) & \\
t. & &
\end{array}$$

In addition to G, other rules may be considered, such as, after all, the lifting rule from Montague (1974):

M If an expression occurs in type a, then it may also occur in any type $((a, b), b)$ (for arbitrary b).

This rule lifts proper names from type e to $((e, t), t)$, so as to obtain one homogeneous class of subject type. (This again can be motivated by pointing at phenomena of 'coordination', such as *Potgieter and a few Boers crossed the Vaal.*) It also allows for general changes in the order of application. Instead of evaluating $(a, b) + a$ to b, one may now also evaluate as $(a, b) + ((a, b), b)$: reversing function and argument roles.

At this point, it should be recalled how G and M arose very naturally already in the categorial model theory of Section 3.3. (Indeed, the latter appeared as a special case of the former.)

This flexible approach has been used recently in diverse areas of semantics (cf. Bach, 1984). One example is argument inheritance in morphology (cf. Moortgat, 1984). Deverbal nominalizations such as 'build-*ing*' inherit argument positions from the underlying verb, as shown in 'building Xanadu'. Now, prima facie, -*ing* takes an activity (type (e, t)) to an object (type e): and so it has type $((e, t), e)$ itself, leaving no room for further arguments. But, e.g., by the rule G, the latter type can also occur as $((e, (e, t)), (e, e))$, allowing *building* to be of type (e, e), having a slot for a 'postponed' argument *Xanadu*.

Another example concerns NP-structure (cf. Hoeksema, 1984). In many languages, such as Iraqi Arabic, there is strong syntactic evidence against the Montagovian analysis of NPs with relative clauses R, as being of the form Det $(N\ R)$. They should rather be (Det $N)R$. Using the above mode of combination, the correct semantic reading can be obtained, saving syntactic appearances:

Let $p = (e, t)$. There are even two ways of proceeding:

(i)
$$\begin{array}{lll}
\text{Det} & N & R \\
(p, (p, t)) & p & (p, p) \\
\quad (p, t) & & (p, p) \\
(p, p), (p, t)\ (G) & & (p, p) \\
\quad\quad (p, t) & &
\end{array}$$

(ii)

Det	N	R
$(p, (p, t))$	p	(p, p)
$((p, p), p), ((p, p), (p, t))$ (G)	$((p, p), p)$ (M)	(p, p)
$((p, p), (p, t))$		(p, p)
	(p, t).	

Corresponding readings will be given in Section 7.5. (As it turns out, (ii) is the correct analysis.)

Finally, in this comparative linguistic mode, the present approach also enables us to lift the apparent initial restriction in Chapter 1 to *SVO-languages*. Other major sentence structures in quite diverse natural languages can now be handled too (cf. Zwarts, 1985, on the categorial analysis of SOV and VSO).

There is also a different, more general kind of motivation for a flexible approach to categories and their combinations. In Chapter 10, the present categorial grammar is used, not to match a pre-established norm of 'grammatical correctness', but rather to provide an independent notion of 'semantical interpretability' with which the former may be compared.

7.3. THE LAMBEK CALCULUS

As it happens, perhaps the most elegant version of a flexible categorial grammar had been proposed already in the little-known paper Lambek (1958). A 'non-directional' version of Lambek's system will be the paradigm in what follows.

The mechanism of type change shows striking resemblances with logical calculi of natural deduction for *conditionals* (in the 'vertical' tradition; cf. Chapter 4 for the 'horizontal' mode).

Using this insight, Lambek constructed a calculus of sequents

$$A \Rightarrow b, \quad \text{or} \quad A \Rightarrow B;$$

meaning that the sequence of types $A = a_1, \ldots, a_n$ reduces to the single type b, or to the sequence of types $B = b_1, \ldots, b_m$. In our presentation, the axioms and rules of the Lambek calculus L will be the following:

(1) $a \Rightarrow a$

(2) $a\ (a, b) \Rightarrow b \qquad (a, b)\quad a \Rightarrow b \qquad$ (*function-elimination*)

(3) $\dfrac{A\ a \Rightarrow b}{A \Rightarrow (a, b)} \qquad \dfrac{a\ A \Rightarrow b}{A \Rightarrow (a, b)}$, for non-empty A (*function-introduction*)

(4) $$\frac{A \Rightarrow b}{B\,A\,C \Rightarrow B\,b\,C} \qquad (\textit{replacement})$$

(5) $$\frac{A \Rightarrow B \qquad B \Rightarrow C}{A \Rightarrow C} \qquad (\textit{transitivity})$$

These rules capture the earlier idea of evaluation. A sequence of expressions evaluates to a type a, if, starting from the corresponding sequence of original types, some succession of admissible type combinations and replacements yields the single type a.

Example (derivation of M and G):

M: $$\frac{a\ (a, b) \Rightarrow b}{a \Rightarrow ((a, b), b)}$$

G: $$\cfrac{\cfrac{\cfrac{\cfrac{(c, a)\ c \Rightarrow a}{(a, b)\ (c, a)\ c \Rightarrow (a, b)\ a} \qquad (a, b)\ a \Rightarrow b}{(a, b)\ (c, a)\ c \Rightarrow b}}{(a, b)\ (c, a) \Rightarrow (c, b)}}{(a, b) \Rightarrow ((c, a), (c, b)).}$$

The analogy with logical deduction shows already in the rules. Function elimination is like Modus Ponens, while function introduction corresponds to Conditionalization.

Example (a logical law).

Here is a derivation of the logical law $(a \rightarrow (b \rightarrow c)) \rightarrow (b \rightarrow (a \rightarrow c))$.

$$\cfrac{\cfrac{\cfrac{\cfrac{a\ (a, (b, c)) \Rightarrow (b, c)}{a\ (a, (b, c))\ b \Rightarrow (b, c)\ b} \qquad (b, c)\quad b \Rightarrow c}{a\ (a, (b, c))\ b \Rightarrow c}}{(a, (b, c))\ b \Rightarrow (a, c)}}{(a, (b, c)) \Rightarrow (b, (a, c)).}$$

Nevertheless, other logical laws may fail. For instance, $a \Rightarrow (b, a)$ is underivable, as it requires a 'vacuous' conditionalization. Also, $(a, (b, c)) \Rightarrow ((a, b), (a, c))$ is underivable, as it requires using an assumption twice — a practice not allowed in our type evaluation (as it stands).

Thus, in a sense, we are now studying the logic of *uses of* premises — in itself, also a task of logical interest.

The calculus L has several useful meta-properties. Notably, strengthening rule (4) to replacement by arbitrary *sequences* adds no new derivable sequents to the system. One proves this by observing that, in the extended system, every derivable sequent $a_1, \ldots, a_n \Rightarrow b_l, \ldots, b_m$ gives rise to a decomposition of $a_1, \ldots, a_n$ into m successive subsequences $A_1', \ldots, A_m'$ such that $A_1' \Rightarrow b_1, \ldots, A_m' \Rightarrow b_m$ are all derivable in the original calculus.

The great strength of non-directional rules also shows in the following observation.

THEOREM. If $A \Rightarrow b$ is derivable, then so is $A' \Rightarrow b$, for any permutation A' of A.

Proof. As every permutation is a composition of interchanges between neighbours, it suffices to show that, if $Ab_1b_2C \Rightarrow d$ is derivable, then so is $Ab_2b_1C \Rightarrow d$. The following example shows the general principle at work. $a_1a_2b_1b_2c_1c_2 \Rightarrow d$, $a_1a_2b_1b_2c_1 \Rightarrow (c_2, d)$, $a_1a_2b_1b_2 \Rightarrow (c_1, (c_2, d))$, $a_2b_1b_2 \Rightarrow (a_1, (c_1, (c_2, d)))$, $b_1b_2 \Rightarrow (a_2, (a_1, (c_1, (c_2, d))))$, $b_1 \Rightarrow (b_2, (a_2, (a_1, (c_1, (c_2, d)))))$, $b_2b_1 \Rightarrow b_2\,(b_2, (a_2, (a_1, (c_1, (c_2, d)))))$ $\Rightarrow (a_2, (a_1, (c_1, (c_2, d))))$, $a_2b_2b_1 \Rightarrow a_2\,(a_2, (a_1, (c_1, (c_2, d))))$ $\Rightarrow (a_1, (c_1, (c_2, d)))$, etc. to $a_1a_2b_2b_1c_1c_2 \Rightarrow d$. □

Using this insight, one may as well re-axiomatize the calculus L by having just one version of each earlier rule, adding a principle of *Permutation*:

(6) $$\frac{A \Rightarrow b}{\pi[A] \Rightarrow b}\ \text{, for all permutations } \pi \text{ of the sequence } A.$$

It should be realized, however, that this feature makes some of the earlier applications somewhat problematic. For instance, NP structure Det $N\,R$ can now be read in *any* order: which may be more than we bargained for. If so, judicious constraints on (6) will have to be imposed eventually. On the other hand, the frequent occurrence of permutations of phrases (e.g., in relative clauses or questions) is an undeniable fact of linguistic life. (See also Section 7.8. for a point of descriptive strategy here.)

Natural Deduction

In actual practice, a tree-format with natural deduction rules makes the calculus L easier to handle. Again, there is an elimination rule (Modus Ponens) and an introduction rule (Conditionalization), just as in ordinary implicational logic. The difference with the latter is slight, but subtle. It resides in the so-called 'structural rules' for inferences: each conditionalization is to withdraw one occurrence of a premise, no more, no less. (So, the ordinary rule of 'Thinning' for premises is not allowed. On the other hand, 'Permutation' of premises is still admissible, at least in the present version of L.) Incidentally, the 'ordinary' variant here yields the so-called *constructive* or *intuitionistic implicational logic I*. Amongst others, I still lacks classical tautologies such as Peirce's Law $((q \to p) \to q) \to q$.

Example. Lifting transitive verbs, and lowering them.

$$\dfrac{\dfrac{\dfrac{\dfrac{\overset{1}{e}\quad e,(e,t)}{e,t}\qquad (e,\overset{2}{t}),t}{t}}{e,t}\;-1}{((e,t),t),(e,t)}\;-2 \qquad\qquad \dfrac{\dfrac{\dfrac{\dfrac{\overset{2}{e}\quad e,\overset{1}{t}}{t}}{(e,t),t}\;-1 \qquad ((e,t),t),(e,t)}{e,t}}{e,(e,t)}\;-2$$

This *two-way* reduction is not available, e.g., for e and $(e, t), t$.

Example. Det N R structure.

Here are two different proof trees for the sequent

$$p,(p,t) \quad p \quad p,p \Rightarrow p,t:$$

(i) $$\dfrac{\dfrac{p\quad p,p}{p}\quad p,(p,t)}{p,t}$$

(ii) $$\dfrac{\dfrac{\dfrac{\overset{1}{p}\quad p,p}{p}\qquad \dfrac{p\quad p,(p,t)}{p,t}}{t}}{p,t}\;-1$$

Other examples from preceding chapters with easy proofs are
— 'homomorphic inflation': $(b, a) \Rightarrow (a, t), (b, t)$;
— determiners in direct object position:

$$(e, t), ((e, t), t) \Rightarrow (e, t), ((e, (e, t)), (e, t)).$$

By varying proof trees, the power of this mechanism emerges. For instance, argument order can be reversed as follows:

$$\dfrac{\dfrac{\dfrac{\overset{2}{e} \qquad \dfrac{\overset{1}{e} \quad e, (e, t)}{(e, t)}}{t}\; -1}{(e, t)}\; -2}{e, (e, t)}$$

As a result, a standard example such as *Every boy loves a girl* gets four proof trees, corresponding to four different readings — again, an abundance which may have to be limited eventually.

Finally, another point of application should be noted. The motivation in the original examples was this. Traditional categorial grammar assigns initial types, after which the only mode of combination is 'function application'. The corresponding natural deduction trees have only Modus Ponens, so to speak. We then allowed 'type shifts' for single expressions, so as to get greater flexibility; after which the old process of application leads to a single final type. But now, a more radical view is possible too. Just call those sequences A recognizable to type b for which there exists a calculus proof going from A to b — regardless of how this is effected (with a final Modus Ponens phase or not). We shall adopt the latter point of view henceforth. Nevertheless, one could search for some kind of 'normal form' for proofs, consisting of a top part with 'inflation rules' for single types, followed by a bottom part with Modus Ponens only. The inflation rules needed for this purpose turn out to be generalized forms of G and M (cf. van Benthem, 1985e).

7.4. RECOGNIZING POWER

The first meta-theoretical question about the Lambek calculus concerned its *decidability*. As we have seen, L is formally similar to the intuitionistic calculus I, whose decidability had already been established by a 'Cut Elimination' method from proof theory. Extending this method to L, Lambek showed that his original calculus (which has directed slashes \, /) is decidable: there exists an algorithm for determining whether any given sequent $A \Rightarrow b$ is derivable. For the present non-directed version, an analogous proof applies (W. Buszkowski, personal communication).

One way of viewing the Lambek proof is as showing that L can be presented equivalently as a 'cut-free' calculus, whose rules at most increase complexity:

(1) $a \Rightarrow a$

(2) $$\frac{A \Rightarrow b \quad B, c \Rightarrow d}{A, B, (b, c) \Rightarrow d}$$

(3) $$\frac{A, b \Rightarrow c}{A \Rightarrow (b, c)}$$

(4) 'Permutation'

For meta-theoretic purposes, the latter version often has its advantages.

The next obvious question concerned the *weak recognizing capacity* of L. More precisely, starting from a finite alphabet, one assigns types (one or several) to each symbol, and then recognizes a language consisting of all those strings $s_1, \ldots, s_k$ for which there exists a corresponding sequence of initial types $a_1, \ldots, a_k$ such that $a_1 \ldots a_k \Rightarrow t$ is derivable in L. (Here, t is any distinguished type for sentences.) Mathematical linguists have been interested in the position of L (as a language-accepting device) in the Chomsky hierarchy of formal languages. An early proof from the sixties purporting to show that directional L-grammars accept precisely all *context-free* languages was recently shown to be defective (cf. Buszkowski, 1982). But, the latter author has a revised version for the uni-directional variant of L.

Still, even if the Lambek grammars were to recognize no new languages beyond the context-free realm (and hence, beyond what was already recognized by traditional categorial grammars), their *strong* recognizing power (providing additional structures for sentences) would remain a virtue.

For the present non-directed calculus, the problem of recognizing power is open. Of course, there are context-free, and even regular languages beyond its grasp — viz. those where the order of symbols is essential. For, the Permutation property of L has the effect that all permutations of strings recognized will also be accepted. On the other hand, some languages of this 'permutation-closed' kind will be recognized which are not context-free. An example is the set of all strings with equal numbers of the symbols a, b, c (cf. also Section 8.3.). It can be recognized by taking a traditional categorial grammar for the related regular language $(abc)^*$ (i.e., all strings consisting of a number of copies of abc) with the full L-rules on top.

How does one prove that a given L-grammar recognizes precisely a particular set of strings? Usually, things are set up so as to get *at least* that set; the *at most* forms the difficulty. For this purpose, one needs to know what *cannot* be derived in L. Now, one very useful property of L-derivations involves counting occurrences of basic types. We define the *e-count* (and likewise, *t-count*, etc.) of any type a by the following recursion:

$$e\text{-count}\,(e) = 1, \quad e\text{-count}\,(t) = 0; \text{etc.}$$
$$e\text{-count}\,((a, b)) = e\text{-count}\,(b) - e\text{-count}\,(a).$$

(So, one counts the number of positive occurrences of e minus that of the negative ones.) For instance, the determiner type $((e, t), ((e, t), t))$ has e-count 2, t-count -1. For sequences of types A, the counts are obtained by adding those for their members. Now, L-derivations have the following invariant:

THEOREM. If $A \Rightarrow B$ is derivable in the calculus L, then the e-count (t-count, . . .) of A equals that of B.

The proof is by induction on the rules used in L-derivations. By way of contrast, note that the ordinary implication calculus I lacks this property. For instance, it can prove $t \Rightarrow (e, t)$, as we saw, which is not valid in L. Note that equality of count is a necessary, not a sufficient condition for derivability: otherwise, all L-reduction arrows could be reversed.

As a first application of this result, e.g., there is no L-justification for deriving adjectives (type $(e, t), (e, t)$) from determiners: the former have

e-count 0, instead of 2. Next, here is the promised application to questions of recognition.

Example. Polish propositional logic.

Consider the alphabet $\{p, \neg, \wedge\}$, and assign types as usual:

$$p \mapsto t \qquad \neg \mapsto (t, t) \qquad \wedge \mapsto (t, (t, t)).$$

Evidently, all well-formed formulas will be recognized as t, already on the Ajdukiewicz schema alone. What the L-apparatus adds are at least all permutations of such formulas. But in this case, it recognizes no more. For, suppose that $a_1 \ldots a_k \Rightarrow t$ is derivable in L, with all a_i $(1 \leqslant i \leqslant k)$ from among the above three types. As the conclusion t has t-count 1, so should the initial sequence. Now, t-count $((t, t)) = 0$, t-count $((t, (t, t))) = -1$. Therefore, the only admissible sequences to the left are those with, say,

n occurrences of $\wedge$, $n + 1$ occurrences of p,
and an arbitrary number of occurrences of $\neg$.

But then, as is easily seen (cf. Section 10.4.), any such sequence is a permutation of some propositional formula.

This example might suggest that using L-rules on top of a traditional categorial grammar will always recognize the permutation closure of the original language. For a counter-example, see van Benthem (1985e), which also presents many other results in this area. We conclude with the current state of affairs:

THEOREM. All permutation closures of context-free languages have a recognizing L-grammar.

CONJECTURE. The converse of the theorem holds as well.

As a final curiosity, one can also determine the class of languages that would be recognized by using the full intuitionistic implication calculus as a system of combination. These turn out to form a small fragment of the *regular* languages: increasing power of type change does not necessarily improve recognizing capacity!

In the remainder of this chapter, however, we are interested in new kinds of logical question generated by this type of categorial grammar, rather than studying 'old about new'. Evidently, the fundamental theme in the above is *variation of types*. Now that a given sequence of expressions may be evaluated to different types in different ways, it becomes of interest to describe the resulting variety. Two main results will be obtained here: (1) for any sequence, the set of its possible types has at most three basic types, 'generating' the others; (2) for any sequence, and any one of its types, there exist at most finitely many non-logically equivalent readings in that type. Thus, the variety induced by L has definite bounds — as it should. Perhaps more fundamental, however, is the search for some independent semantic notion of 'admissible type change', which can serve as a touchstone for the given syntactic rules of type manipulation. We shall present one relevant proposal here, and a completeness theorem for the calculus L with respect to this semantics.

7.5. SEMANTICS FOR TYPE CHANGE

In L-derivations for sequents $A \Rightarrow b$, there is an obvious semantic motivation for the individual steps. A function-elimination amounts to an instruction to *apply* a function to an argument, a function-introduction requires creating a new function by *abstraction*. More formally, the 'meanings' of derivations of $A \Rightarrow b$ may be given using terms in a logical type-theoretical language, having an infinite supply of individual variables $x_a, y_a, \ldots$ for each type a, and allowing the following kinds of term formation:

— if t_1 is of type (a, b), and t_2 of type a, then $t_1(t_2)$ is a term of type b ('application'),

— if t is of type b, and x is a variable of type a, then $(\lambda x \cdot t)$ is a term of type (a, b) ('lambda-abstraction').

This language has its obvious interpretation in hierarchies of functional domains, starting from domains D_e ('entities'), D_t ('truth values'; usually $\{0, 1\}$), and ascending via the stipulation that $D_{(a,b)} = (D_b)^{D_a}$.

To assign a lambda term to an L-derivation of $A \Rightarrow b$, one takes *distinct* variables $x_{a_1}, \ldots, x_{a_n}$ for each type occurrence in A, and proceeds inductively as in the following examples.

Example (axioms). Axiom $a \Rightarrow a$ has a term assigned x_a. (Given any a-value, it yields the same a-value, of type a.) Axiom $(a, b)\ a \Rightarrow b$ gets $x_{(a,b)}(x_a)$, of course.

Example (longer derivations). The earlier derivation of rule G is treated as follows:

$$\cfrac{\cfrac{\cfrac{\cfrac{x_{(c,a)}(x_c)}{x_{(a,b)},\ x_{(c,a)}(x_c)} \qquad x_{(a,b)}(x_a)}{x_{(a,b)}(x_{(c,a)}(x_c))}}{(\lambda y_c \cdot x_{(a,b)}(x_{(c,a)}(y_c)))}}{(\lambda y_{(c,a)} \cdot (\lambda y_c \cdot x_{(c,a)}(y_c)))).}$$

(*substitute* the relevant expressions for $x_{(a,b)}$, x_a)

This example follows the structure of proofs in the original formulation of the calculus L in Section 8.3. It is also possible to read off lambda-terms from natural deduction trees, or other related formalisms. For instance, the two trees given for the analysis of Det NR structure yield, respectively,

(i) $x_{(p,(p,t))}(x_{(p,p)}(x_p))$, and the non-equivalent
(ii) $\lambda y_p \cdot x_{(p,(p,t))}(x_p)(x_{(p,p)}(y_p))$.

The precise procedure to be followed here will appear below. It will be clear already how these lambda terms act as *type changers*, expressing constructive instructions for obtaining a denotation of type b from given denotations of types $a_1, \ldots, a_n$.

As it stands, this type-theoretical language is too rich. Many of its terms do not correspond to admissible type changes in the calculus L. Thus, the question is to find plausible restrictions — on the way to a perfect match. First, lambda terms with 'vacuous abstraction' are to be excluded. E.g., $\lambda y_e \cdot x_t$ corresponds to the non-derivable sequent $t \Rightarrow (e, t)$. Then lambda terms with subterms ('subroutines') without free variables are to be excluded. E.g., $x_{((e,e),t)}(\lambda y_e \cdot y_e)$ corresponds to the non-derivable sequent $((e, e), t) \Rightarrow t$. Finally, repetitions of the same variable occurring freely in some subterm are forbidden. E.g., it was observed before that $(a, (b, c))$ does not evaluate in L to $((a, b), (a, c))$ — even though there exists a lambda term $(\lambda y_{(a,b)} \cdot (\lambda y_a \cdot x_{(a,(b,c))}(y_a)(y_{(a,b)}(y_a))))$. The resulting restricted class of lambda terms will be called Λ.

Although these particular restrictions all have a certain rationale, it is evident that there exists a whole hierarchy of various kinds of 'lambda recipes', rather than one unique preferred one. And indeed, once our completeness result has been obtained, it may be modified at once to deal with other calculi.

It now remains to connect up the L-calculus and its Λ-semantics:

THEOREM ('$L = \Lambda$'): A sequent $a_1, \ldots, a_n \Rightarrow b$ is derivable in the calculus L if and only if there exists a term t in Λ of type b with exactly the free variables $x_{a_1}, \ldots, x_{a_n}$.

Proof. The theorem follows from the slightly stronger claim that a sequent $a_1, \ldots, a_n \Rightarrow b_1, \ldots, b_m$ is L-derivable iff there exist Λ-terms $t_1, \ldots, t_m$, of types $b_1, \ldots, b_m$, whose free variables use up $x_{a_1}, \ldots, x_{a_n}$ in m disjoint segments.

From L to Λ. Rules (1), (2) have been treated already in the examples. Rule (3): suppose $t(x_{a_1}, \ldots, x_{a_n}, x_a)$ is given, of type b. Then the Λ-term $(\lambda x_a \cdot t)$ matches the conclusion. Rule (4): The terms for the conclusion are all single variables (for B, C), together with the already given term for $A \Rightarrow b$. Rule (5): Let $A = a_1, \ldots, a_n$, $B = b_1, \ldots, b_m$, $C = c_1, \ldots, c_k$. We have Λ-terms $t_1, \ldots, t_k$ using up $x_{b_1}, \ldots, x_{b_m}$, as well as Λ-terms $s_1, \ldots, s_m$ using up $x_{a_1}, \ldots, x_{a_n}$. But then, the Λ-terms corresponding to the conclusion of the Transitivity rule are obtained by *substituting* the matching terms s_i for the variables x_{b_i} occurring in the terms t_j. It is easily checked that no Λ-restrictions are violated in this process. Rule (6): there is nothing to prove.

From Λ to L. The case of a single term illustrates the general procedure. (The following induction uses the fact that, if an application or abstraction term is in Λ, then so are its immediate components.) (i) t is a variable of type a: $a \Rightarrow a$ is L-derivable. (ii) t is an application $t_1(t_2)$: if L derives $a_1, \ldots, a_i \Rightarrow (b, c)$ as well as $a_{i+1}, \ldots, a_n \Rightarrow b$, then it also derives $a_1, \ldots, a_n \Rightarrow c$, by its rules (4), (2), (5). And (iii), t is an abstraction $(\lambda x_b \cdot t_c)$: if L derives $a_1, \ldots, b, \ldots, a_n \Rightarrow c$, then it also derives $a_1, \ldots, , \ldots, a_n \Rightarrow (b, c)$, by rules (3) and (6). □

Notice how the above proof *effectively* provides each eligible lambda term with a derivation whose meaning it is. In other words, viewing lambda terms as possible 'readings', constructions are provided exemplifying each reading. The procedure in the other direction, providing L-derivations with readings, is effective too.

The following general picture now emerges. At one end of the spectrum lies a *pure application language* without lambdas, whose terms correspond with derivations in the original Ajdukiewicz categorial grammar. At the other end, there is the *full abstraction/application language*, whose terms may be seen to correspond precisely with derivations in the intuitionistic implication calculus I. (This observation is known from general logic.) The Lambek calculus, and possible variants, lie somewhere in between.

Still, the lambda-connection does not provide a semantics in all senses of the word, such as providing appealing 'pictures' for refuting derivability. And indeed, there are other useful formats of semantic description for the Lambek calculus. For instance, in another long-standing tradition, Buszkowski (1982) provides an 'algebraic semantics'. Here, we conclude with a somewhat different perspective on this issue.

Constraints on Admissible Transitions

The lambda language is a medium for recording, rather than constraining type transitions. If we are to find additional conditions, one source for these is Chapter 3, with its emphasis on general categorial mechanisms.

An important general structure permeating all categories was the implication order $\sqsubseteq$ defined in section 3.2. Now, it seems attractive to demand *preservation* of such *implications*, when changing denotations from one category to another. Thus, an admissible lambda term $\tau(x_a)$ would have to satisfy, for all u, v in D_a,

$$u \sqsubseteq v \textit{ only if } \text{(perhaps even: } \textit{if and only if}\text{)}$$
$$[\![\tau]\!]_u^{x_a} \sqsubseteq [\![\tau]\!]_v^{x_a}.$$

(There is some self-explanatory notation here.)

Some of the earlier transitions have this property. An example is the G-form $(a, b) \Rightarrow ((c, a), (c, b))$, with lambda term

$$\lambda y_{(c,a)} \cdot \lambda y_c \cdot x_{(a,b)}(y_{(c,a)}(y_c)).$$

Others do not. For instance, the M-form $(e, t) \Rightarrow (((e, t), t), t)$ has an associated term $\lambda y_{((e,t),t)} y(x_{(e,t)})$, sending sets $X \subseteq E$ to the family of sets $X^+ \subseteq P(E)$ containing X as an element. But there is no guarantee that, if $X_1 \subseteq X_2$, then also $X_1^+ \subseteq X_2^+$.

The general principle behind these examples is that of Chapter 6. Only those lambda terms will guarantee preservation of $\sqsubseteq$ whose occurrences of x_a are syntactically *positive*, in the following sense:

— x_a occurs positively in x_a (but in no other variable).

— if x_a occurs positively in A, then also in $A(B)$ (for arbitrary terms B),

— if x_a occurs positively in A, then also in $\lambda y \cdot A$ (for any variable y distinct from x_a).

And this fits in with the above formulas. $x_{(a,b)}$ occurs positively in the first lambda term, whereas $x_{(e,t)}$ does not in the second. (Still, for trivial reasons — $\sqsubseteq$ being the identity on D_e — the original Montague rule $e \Rightarrow ((e, t), t)$ remains admissible in this sense.) Thus, preservation of implication is a genuine additional constraint on the calculus of type change.

Remark: Through the above correspondence between lambda terms and derivations, this condition may also be translated into one on L-proofs:

whenever Modus Ponens is applied, say with 'premises' $a_1 \ldots a_i \Rightarrow b$ and $a_{i+1} \ldots a_n \Rightarrow (b, c)$ to obtain $a_1 \ldots a_n \Rightarrow c$, all type occurrences $a_1, \ldots, a_i$ should be transferred to the right-hand side eventually (by Conditionalization).

Whether the above constraint is reasonable remains a matter of debate. E.g., in Groenendijk and Stokhof (1984), it is claimed that failure of preservation with M, a rule independently motivated by the needs of 'coordination', rather shows that one *needs* an L-connected *family* of types for single expressions, with some of their behaviour accounted for at one level (for instance, logical relations) and other aspects (such as combinatorial 'affinities') elsewhere. Our flexible categorial grammar then provides a 'family tree' for these various types.

Another possible constraint suggested by Section 3.3., might be to focus on those transitions whose associated recipe makes the lifted denotation a *homomorphism* in its new category. In general, this is a very restrictive condition, only fulfilled when the main argument of the term 'really' occurs as the head functor in its matrix. This is what happens in Montague's M rule M: $\lambda y_{(e,t)} \cdot y_{(e,t)}(x_e)$. It can also be observed with the transition underlying the Keenan and Faltz treatment of transitive verbs:

$$\lambda y_{((e,t),t)} \cdot \lambda y_e \cdot y_{((e,t),t)}(\lambda z_e \cdot x_{(e,(e,t))}(z_e)(y_e)).$$

(The 'λz' is only here for a minor combinatorial purpose.)

Finally, there is a constraint which turns out to be satisfied by all lambda terms, and yet has a certain intuitive content. In Chapter 3, a special role was played by the *logical* items, invariant for automorphisms π induced by underlying permutations of the individual domain D_e. One obvious question is: will such a logical item in type a

remain logical, when lifted to a higher type by our rules? The answer is affirmative, and it follows from this result:

THEOREM. For every lambda term τ_b with the free variables $x_1, \ldots,$ x_n, every permutation π of the above kind, and every suitable interpretation function $[\![\ ]\!]$,

$$\pi([\![\tau_b]\!]^{x_1 \ldots x_n}_{d_1 \ldots d_n}) = [\![\tau_b]\!]^{x_1 \ldots x_n}_{\pi(d_1) \ldots \pi(d_n)}$$

The proof is by induction on the complexity of the term τ. So, if an item x_a is invariant for such permutations π already, its lifted forms $\tau(x_a)$ will be invariant too.

7.6. DIGRESSION: POSSIBLE WORLDS SEMANTICS

The analogy between the Lambek calculus and the intuitionistic logic I also suggests a more traditional semantics for L; not for its derivations, but for its sequents as such. An outline of this approach follows here. (For details, cf. van Benthem, 1985e.)

The usual semantics for I has models

$$M = \langle \mathfrak{I}, \sqsubseteq, V \rangle,$$

with $\mathfrak{I}$ a set of 'forcing conditions' or 'information pieces', $\sqsubseteq$ a partial order ('possible growth of information'), and V a valuation giving a truth value to each proposition letter at every $i \in \mathfrak{I}$ (with truth persisting along $\sqsubseteq$-successors).

Evaluation is as follows:

$$\begin{array}{lll} M \vDash p[i] & \text{iff} & V_i(p) = \textit{true} \\ M \vDash \alpha \rightarrow \beta[i] & \text{iff} & \textit{for all } j \sqsupseteq i, M \vDash \alpha[j] \textit{ only if } M \vDash \beta[j]. \end{array}$$

Finally, validity for a sequent $\alpha_1, \ldots, \alpha_n/\beta$ ('$\alpha_1, \ldots, \alpha_n \vDash \beta$') means that, in all M, at all $i \in \mathfrak{I}$, if $\alpha_1, \ldots, \alpha_n$ hold at i, then so does β. A very straightforward Henkin completeness argument then establishes that always

$$\alpha_1, \ldots, \alpha_n \vDash \beta \quad \text{iff} \quad \alpha_1 \ldots \alpha_n \Rightarrow \beta \quad \text{is derivable in } I.$$

The difference with L lies in one's view of the 'information pieces'. These are now, as it were, *checks* for verification, that can be cashed only once. This idea leads to the following new models

$$M = \langle \mathfrak{I}, \oplus, V \rangle;$$

where $\oplus$ is now a binary operation of *addition* of information pieces. Here, $\oplus$ will be required to be

associative $(i \oplus (j \oplus k) = (i \oplus j) \oplus k)$ and *commutative* $(i \oplus j = j \oplus i)$; though not necessarily *idempotent* $(i \oplus i = i)$.

This new structure is reflected in the key clause for evaluating complex types:

$$M \models (a, b)[i] \text{ iff } \textit{for all } j \in \mathfrak{I} \textit{ such that } M \models a[j], M \models b[i \oplus j].$$

This time, no persistence clause is added for the valuation: L-types do not enjoy the *heredity* possessed by all I-formulas. Finally, semantic consequence becomes

$a_1, \ldots, a_n \models b$ if, for all $M = \langle \mathfrak{I}, \oplus, V \rangle$ and all $i_1, \ldots, i_n \in \mathfrak{I}$,
if $M \models a_1[i_1]$ and ... and $M \models a_n[i_n]$,
then $M \models b[i_1 \oplus \cdots \oplus i_n]$.

For instance, it may be checked that all inference rules of L preserve this notion of consequence.

Now, the usual logical questions can be raised with respect to this semantics. For a start, a judicious formulation of a *semantic tableau* method may be used to establish *decidability* for the above notion of consequence. Then, to see that the latter indeed coincides with L-derivability, it is to be shown how non-L-derivability leads to counter-examples in the above models. (This is 'completeness': 'soundness' having been observed already.) For the latter purpose, a simple Henkin model suffices, with

$\mathfrak{I}$: all finite sets of formulas (viewed typographically as occurrences),

$\oplus$: union (in the sense in which, e.g., $\{p, q\} \oplus \{p, r\}$ would become $\{p, q, p, r\}$, rather than $\{p, q, r\}$)

$V_i(p) = \textit{true}$ iff $i \Rightarrow p$ is L-derivable.

By induction on the complexity of types a, a Truth Lemma can then be proved:

$$M \models a[i] \text{ iff } i \Rightarrow a \text{ is } L\text{-derivable.}$$

As a final topic of interest, we mention the possible semantic study of logics 'intermediate' between L and I. For instance, various additional axioms turn out to correspond to possible extra conditions on the operation $\oplus$.

Example: The 'double use' principle $a\ a, (a, b) \Rightarrow b$ will be valid if and only if the following form of idempotence holds:

$$i \oplus i \oplus j = i \oplus j, \quad \text{for all } i, j.$$

Remark: As for the general comparison of L and I, van Benthem (1985e) shows that L cannot be faithfully embedded in I; whereas the converse question is open.

All these topics can be extended to the case where the type change calculus also has an operation $\cdot$ of *concatenation* (as was the case in Lambek's original paper). The calculus rules here will be the usual ones for *conjunction*, even though not all usual $\wedge$-laws can be derived, owing to the general restrictions in L. The matching semantic clause becomes

$M \vDash a \cdot b\ [i]$ iff *there exist* j, k *with* $i = j \oplus k$ *such that* $M \vDash a\ [j]$ and $M \vDash b\ [k]$.

Despite all these matchings, it should be admitted that the present semantics stays very close to syntax — and indeed, it may be closely related to the algebraic semantics of Buszkowski (1982). Still, the present models, with their 'dynamic' view of evaluation as 'consuming' an index i to establish the truth of a formula may have some suggestive value beyond the present setting.

7.7. VARIATION IN OUTCOMES

The present grammar is extremely generous in what it counts as 'interpretable'. For, every sequence of types evaluates to a single combined type, even a whole family of them:

$$a_1 \ldots a_n \Rightarrow ((a_1, (\ldots (a_n, b) \ldots)), b), \quad \text{for all types } b.$$

Nevertheless, the *range of outcomes* for a sequence is subject to certain restrictions.

One obvious conjecture would be that the set of 'outcomes' for a sequence A, i.e., those types b for which $A \Rightarrow b$ is provable, has a 'generator': one single type b such that (i) $A \Rightarrow b$, and (ii) if $A \Rightarrow d$, then $b \Rightarrow d$, for all types d. In propositional logic, this property is immediate, once the language has *conjunction*. For the pure implication calculus I, however, the answer seems to be unknown. The conjecture fails for the Lambek calculus.

Example: The sequence $(e, t)\ (t, e)$ evaluates to both (e, e) and (t, t):

$$\dfrac{\dfrac{\dfrac{\overset{1}{e} \quad e, t}{t} \quad t, e}{e}}{e, e}\;-1 \qquad\qquad \dfrac{\dfrac{\dfrac{\overset{1}{t} \quad t, e}{e} \quad e, t}{t}}{t, t}\;-1$$

These two outcomes are incomparable in L (even though their e- and t-counts are the same). This may be shown by noting the impossibility of producing either implication in the special cut-free version of L presented at the beginning of Section 7.4. (A natural language example of this ambiguity might be the phrase *the woman who fears that # runs*, with a gap that is indifferently e or t.) More generally, n-cycles of this kind occur for all finite n.

To return to the matter of the range of outcomes, here is the principle of generation.

THEOREM. For any sequence of types $a_1, \ldots, a_n$, its set of outcomes is L-generated by the two types $((a_1, (\ldots (a_n, e) \ldots)), e)$ and $((a_1, (\ldots (a_n, t) \ldots)), t)$, together with at most one of the types e, t.

Proof. The first-mentioned two types are obviously derivable outcomes. Moreover, they ensure that all types of the form $((a_1, (\ldots (a_n, b) \ldots)), b)$ are available — through the following observation: $((a_1, (\ldots (a_n, b_2) \ldots)), b_2) \Rightarrow ((a_1, (\ldots (a_n, (b_1, b_2)) \ldots)), (b_1, b_2))$ is L-derivable. Then, if $a_1, \ldots, a_n \Rightarrow (b_1, b_2)$ is derivable in L, we are done, since the following sequent will be L-derivable: $((a_1, (\ldots (a_n, b_2) \ldots)), b_2) \Rightarrow (b_1, b_2)$. (The argument is this: $a_1, \ldots, a_n \Rightarrow (b_1, b_2)/ a_1, \ldots, a_n, b_1 \Rightarrow b_2/b_1 \Rightarrow (a_1, (\ldots (a_n, b_2) \ldots))/((a_1, (\ldots (a_n, b_2) \ldots)), b_2)\ b_1 \Rightarrow b_2$.) Finally, it remains to add e or t, if these types actually occur among the outcomes. By the invariance of e-count, the two cannot occur together for the same sequence $a_1, \ldots, a_n$. □

This result does not imply that the set of *all* types is generated by some 'finite basis' of types. For instance, all types in the following sequence are mutually non-derivable in L: $t, (e, t), (e, (e, t)), \ldots$. The e-count is different in all cases.

In a sense, the preceding theorem merely shows that all complexity of outcomes is already contained in the relation $\Rightarrow$ between single types. Moreover, it is the 'inflated' types which do all the work.

Next, we turn to a second kind of variety, concerning the *range of readings* for an outcome.

One and the same sentence may have categorial analyses with different meanings, as was observed before. In other words, given a sequence $a_1, \ldots, a_n$ and one of its outcomes b, there is still variety in derivations of the sequent $a_1, \ldots, a_n \Rightarrow b$. It is usually assumed that only *finitely* many different readings can exist. For instance, the sentence *every crook does not tell lies*, when read as the type sequence $((e, t), t)/(t, t)/(e, t)$, has precisely the two readings: 'not (every)' and 'every (not)'. We shall prove this expectation of finiteness, by an appeal to the completeness theorem of Section 7.5. In that earlier perspective, different derivations for a sequent $a_1, \ldots, a_n \Rightarrow b$ correspond to different Λ-terms of type b, having the free variables $x_{a_1}, \ldots, x_{a_n}$. The question is, even if there are infinitely many of these, is their number *finite modulo logical equivalence*? For instance, there remain only two candidates for the above sequence; essentially:

$$x_{(t,t)}(x_{((e,t),t)}(x_{(e,t)})) \quad \text{and} \quad x_{((e,t),t)}(\lambda y_e \cdot x_{(t,t)}(x_{(e,t)}(y_e)))$$

THEOREM. A sequence of types can have only finitely many logically distinct readings in any given type of outcome.

Proof. The assertion follows from the following claim about the corresponding logical type theory:

Up to logical equivalence, the set of Λ-terms of a fixed type b, with a fixed finite number of free variable occurrences $x_{a_1}, \ldots, x_{a_n}$, is finite.

To see this, consider such terms in *lambda normal form*; i.e., with as many 'lambda-conversions' performed as possible:

$$(\lambda x \cdot t_1)(t_2) \rightarrow [t_2/x]t_1.$$

This process is admissible, as lambda conversions do not lead outside of the class Λ. This procedure restricts the occurrences of the remaining λ-operators: they can no longer occur in contexts of the form $(\lambda x \cdot t_1)(t_2)$. Moreover, by renaming bound variables where necessary, each occurrence of λ may be marked by a unique variable y_a.

LEMMA. Each occurrence λy_a in a normal form corresponds to a unique occurrence of a as a subtype of b or a_1 or ... or a_n.

Then, as there are only finitely many such subtypes, our normal forms will be composed of just a finite stock of symbols: occurrences of λy, x_{a_i} and applications linking these — and hence the assertion of the theorem follows.

The proof of the lemma is by induction on the complexity of terms, noting that each subterm of a term in normal form is itself in normal form. Case 1: the term is a variable. There is nothing to be proved. Case 2: the term is a lambda-abstract $\lambda y_{b_1} \cdot t'$, with $b = (b_1, b_2)$, t' of type b_2. Apply the inductive hypothesis to t', with respect to the types $a_1, \ldots, a_n, b_2$. Case 3: the term is an application $t_1(t_2)$. This case will be illustrated by an example. First, notice that t_1 cannot begin with a λ. Hence, it is either a variable or an application. In the latter case, consider its left-most constituent, and so on. Eventually, one finds a 'leading variable' x_{a_i} on the left. For instance, suppose that the shape of the term is as follows: $((x(s_1))(s_2))(t_2)$, where t_2 has type c, s_2 has type c_2, s_1 type c_1 and x type $a_i = (c_2, (c_1, (c, b)))$. Now, s_1, s_2, t_2 divide up the remainder of the variable occurrences $x_{a_1}, \ldots, x_{a_n}$, and the inductive hypothesis may now be applied to (i) t_2 with respect to its 'x-types' plus c, (ii) s_2 with respect to its 'x-types' plus c_1, and (iii) s_1 with respect to its 'x-types' plus c_2. As all these sets of types are disjoint, the required conclusion (for this case) follows. □

7.8. DISCUSSION

The above theory revolved around one specific calculus of type change. But there is a need for a certain latitude in the enterprise: L certainly is not the final word. And indeed, most results and techniques developed would also work for a wider range of such systems. Here are some possible directions.

To begin with, L is perhaps overly generous in its strong capacity for recognizing phrase structures. Therefore, various constraints may be needed to decrease its tolerance of permutations. And of course, *directional* variants ought to be studied in any case. Nevertheless, it should be noted that *two* strategies may be discerned among linguists working in this area. One is to keep the calculus as weak as possible, adding rules only when forced by the facts of syntax. Another is to keep the calculus strong but simple, adding filters afterwards to account for syntactic peculiarities of specific natural languages. (cf. Flynn, 1983).

On the other hand, there are also type transitions which L does not capture yet. One urgent example is that of *conjunction*, with basic sentential type $(t, (t, t))$, which ought to be raised to predicate conjunction (e, t), $((e, t), (e, t))$. But, no such L-derivation is possible: the e-counts do not match. What is needed here is a certain amount of 'recycling' of types, using the same type e twice. But perhaps, another line of approach is preferable here. The conjunction *and* is one of a class of lexical items which do not fit the uniform functional mould (a, b) too well: they are *coordinators*, not subordinators. So, it might be better to admit *binary* types as well, such as, say, $(t, t; t)$ for conjunction, and extend the calculus L in some appropriate fashion. Alternatively, a binary 'type-coordinator' might be added to our calculus, forming sequences of 'parallel' arguments.

Various other arguments have been advanced for considering strengthenings of the calculus L, relaxing its structural rules (see Bach *et al.* (eds.), 1986). Further interesting questions arise from the interplay of slashes $\backslash$, $/$, neglected in this chapter. For instance, which theoretically possible 'dualities' between these two directions are observable in natural language?

Then also, more substantial 'mathematical' type transitions were studied in Section 3.3. As was observed there already, these seem to involve a lambda/application language having *identity* as well. Is there still an L-like system in this area? Even so, there are many semantic family resemblances across types which are not captured by the above approach at all, such as that between verbs and their participles or nominalized forms. But that is another story.

Finally, type change is just one instance of what might be called a broad 'semantic mechanism', operative in natural language. Other examples are inference, manipulating temporal or modal perspective, etc. Such mechanisms can *interact*, witness the earlier points of contact between *type change* and *inference* (Sections 6.2., 7.5.). These interactions can themselves be studied systematically. For instance, van Benthem (1986a) presents a Lambek calculus with systematic monotonicity marking, which describes the changing inferential behaviour of expressions, when given different categorial construals. Another general interaction of this kind is the interplay of *type change* and *intensionalization*, studied in Section 7.9. below. Thus, natural language is more than a sum of expressions, or production rules. There is higher-order structure too, which deserves semantic attention — both within and across human languages.

7.9. APPENDIX: INTENSIONAL TYPE CHANGE

One attraction of a flexible categorial grammar is its perspicuity. The interplay of simple basic type assignments with simple rules for type change produces complex higher type behaviour which other theories, such as Montague's, have to spell out in formidable detail. For instance, it has been pointed out in Partee and Rooth (1983) how Montague's 'inflexible' approach forces him to assign words to 'worst case categories', representing the most complex environments they might find themselves in (where we have given 'best case' analyses, plus a set of rules for coping with adversity).

There remains another, independent source of complexity in Montagovian type assignments, being an 'intensionalization', providing expressions with complex intensional types, which may be needed when they find themselves in opaque contexts. The most convenient categorial setting here is just a type theory with *three* basic types: e, t as well as s ('possible worlds', or 'indices'). For instance, intransitive verbs already get type $((s, e), t)$ with Montague: they denote properties of 'individual concepts' (type (s, e)) rather than of bare individuals. Again, Montague's general mechanism injects an over-dose of types s, 'just in case'. Can we devise a simpler scheme, inspired by the above?

First, there is an entirely reasonable point of view that might justify, at least, the rule

change t into (s, t).

Consider the earlier extensional e, t-grammar. Objects of type e were to be *entities*. What about those of type t? Intuitively, one thinks here of some set of *propositions* (rather than the bare truth value domain $\{0, 1\}$). Moreover, the latter come as an ordered structure, with a relation of implication, perhaps even a full *Boolean Algebra*. (As a matter of natural language, this seems a very mild assumption.) But then, it is well known that every such algebra can be represented as an algebra of subsets of some carrier set S of ultrafilters (or 'maximal state descriptions'). This then will be our base domain D_s; with $D_t = \{0, 1\}$ arising as a means of classifying subsets by their characteristic functions. (Alternatively, thoroughly Boolean-minded semanticists, such as Keenan and Faltz (1985), might think of this move as making the whole semantics 'Boolean-valued' rather than 'two-valued'.) Propositions will now correspond to sets of possible worlds, in the Leibnizian sense, and hence we baptize the above principle the *Leibniz Rule*.

There is more to the above construction of possible worlds than has been stated. For instance, once upon this road, we will also have to reconsider some other basic types, notably, that of properties $((e, t))$. Recent philosophical research has tended to take individuals, propositions *and properties* (unary, binary and higher) as primitive entities (cf. Bealer, 1983). So, perhaps, properties should become an intensional base type too. (The earlier notation p $(=(e, t))$ in fact suggested this.) Note, however, that the above move already assigns properties type $(e, (s, t))$ — or equivalently, $(s, (e, t))$: i.e., extensions of properties-old-style across all possible worlds. In that, there is already a lot of intensionality.

Should *individuals* be raised in type too? No very convincing arguments have been advanced for the transition from e to (s, e). For instance, the common claim that this is needed with terms such as *the president* lacks force. The latter have NP-type $((e, t), t)$ already on the old approach — and hence can be given new types $((e, t), (s, t))$ $(=(s, ((e, t), t)))$, or even $((e, (s, t)), (s, t))$ $(=(s, ((s, (e, t)), t)))$. There is enough intensional padding here to go a long way.

In fact, so much intensionality is generated by our simple rule that the opposite worry might arise. Certain expressions, such as the Boolean connectives, do not really seem to be affected by these intensionalizations at all. For instance, predicate negation will still be defined 'pointwise', descending to specific indices and using the old extensional version: $not_{\text{int}} = \lambda P \cdot \lambda s\ not_{\text{ext}}(P(s))$. This is another ubiquitous feature in the Montagovian tradition. In many semantic value clauses, several layers of intensionality have to be unpacked, to operate at lower, essentially extensional levels. One interesting question then is the following. Exactly when can a function A from $(s, (e, t))$ to $(s, (e, t))$ be *represented* by a function A^* from (e, t) to (e, t), in the sense that $A(P)(s) = A^* \circ P(s)$ (where '$\circ$' denotes *composition*)? One direct answer is this:

> 'if and only if, for all P, Q, and all $x, y \in S$, if $P(x) = Q(y)$, then $A(P)(x) = A(Q)(y)$'.

Strong indifference principles such as this are likely to hold for many expressions — notably determiners — which pay only nominal allegiance to intensionality.

On the other hand, the Leibniz Rule alone does not suffice for handling all of Montague's legitimate intensional concerns. In particular, the familiar case of intensional *seek* remains problematic. *Seek* (*a unicorn*) has extensional component surface types $(e, (e, t))$ $((e, t), t)$; which may be intensionalized to, say, $(s, (e, (e, t)))$ $(s, ((e, t), t))$. We want to combine these two to get type (presumably) $(s, (e, t))$. We have before us the behaviour of *seek* and *a unicorn* across all possible worlds. Now, if the basic idea of possible worlds semantics is to have any illuminating power at all, one should try to explain the meaning of the combined intensional expression by making genuine use of these 'multiple extensions'. Say, someone seeks a unicorn intensionally if she seeks one extensionally in all of her belief worlds. By contrast, the Montagovian intensional types assigned in this situation seem to be motivated only by the mechanics of blocking undesired 'extensional' inferences.

Still, the problem is that $(s, (e, (e, t)))$ and $(s, ((e, t), t))$ do not combine naturally to $(s, (e, t))$. One possible solution here might seem to follow from an earlier-mentioned liberalization of the Lambek calculus, so as to allow multiple uses of premises. But, this only returns the usual extensional reading 'locally' at the world of evaluation. What is needed instead, in line with the above intuitive proposal, is to combine $(s, (e, (e, t)))$ and $(s, ((e, t), t))$ *directly* into $(s, (e, t))$, giving the intension of extensionally *seeking a unicorn*. The general categorial principle here seems to be the following transition:

$$\textit{from} \quad a_1, \ldots, a_n \Rightarrow b \quad \textit{to} \quad (s, a_1), \ldots, (s, a_n) \Rightarrow (s, b).$$

One might call this *Frege's Rule*, honoring this well-known policy of employing full intensional reference in opaque contexts. Its *meaning* cannot be a simple lambda-recipe, however — given the preceding discussion — and perhaps, no general explication is possible here, covering all types of intensionalization.

A more systematic approach to these matters is proposed in van Benthem (1986b). Here are a few relevant points.

(1) The above Leibniz Rule may be compared directly with Montagovian modes of intensionalization. For instance, transform any exten-

sional type a into an intensional version $I(a)$ by making the following replacement throughout:

$$t \mapsto (s, t), \qquad e \mapsto (s, e).$$

Then $I(a)$ is equivalent to a type $(s, M(a))$, where M is defined by the well-known PTQ-recursion:

$$M(e) = e, \quad M(t) = t, \quad M(a, b) = ((s, M(a)), M(b)).$$

(2) Such global intensionalizations do not affect earlier possibilities of categorial combination: intensionality produces no syntactic constraints. For, the following implication is provable:

$$\textit{if} \quad a_1, \ldots, a_n \Rightarrow_L b, \quad \textit{then} \quad I(a_1), \ldots, I(a_n) \Rightarrow_L I(b).$$

(3) Nevertheless, there may be good descriptive reasons for a 'mixed strategy', as some items remain extensional in meaning (such as the earlier 'not'), whereas others really exploit their intensional freedom (such as the conditional 'if'; cf. Chapter 4). This still leaves room for categorial combination between the two kinds of item, as extensional types can intensionalize to some extent already via the *Lambek* rules. Notably, 'not' can go from type (t, t) to type $(s, t), (s, t)$, with a corresponding lambda-recipe which is precisely the explication presented earlier! One general issue then becomes to *define* generally when an expression of some intensional type is 'really' *extensional*, generalizing the earlier criterion for extensional operations on unary predicates.

(4) Finally, the interaction between *intensionalization* and changing potential for *inference* remains to be explored in more detail — reconsidering some of Montague's decisions. It certainly seems possible to retain *monotonicity* inference in this area, provided that the data of 'inclusion' are now given in a suitably strong intensional sense, across worlds.

CHAPTER 8

SEMANTIC AUTOMATA

An attractive, but never very central idea in modern semantics has been to regard linguistic expressions as denoting certain 'procedures' performed within models for the language. For instance, truth tables for propositional connectives may be viewed as computational instructions for finding truth values. Another example is the proposal in Suppes (1982) to correlate certain adjectives with procedures for locating an individual in some underlying comparative order. And finally, the frequent proposals in a more computer-oriented setting for translating from natural language into programming languages are congenial too. In this chapter, this perspective will be applied to determiners, or more in particular, quantifier expressions.

In order to correlate quantifiers with procedures, the earlier generalized quantifier perspective will be used. As in Chapter 2, a quantifier denotes a functor $Q_E AB$ assigning, to each universe E, a binary relation among its subsets. Viewed procedurally, the quantifier has to decide which truth value to give, when presented with an enumeration of the individuals in E marked for their (non-)membership of A and B. Equivalently, the quantifier has to recognize a 'language' of admissible sequences in a 4-symbol alphabet (as there are four distinct types of A, B-behaviour). But then, we have arrived at the familiar perspective of mathematical linguistics and automata theory.

This observation turns out to be more than a formal perspectival trick. We shall find surprising connections. In particular, the *Chomsky Hierarchy* turns out to make eminent semantic sense, both in its coarse and its fine structure. For instance, the borderline regular/context-free is connected with that between first-order and higher-order definability. But also, within these broad classes, machine fine structure is correlated with a significant semantic hierarchy. Thus, what is often regarded as the main formal stronghold of pure syntax, can also be enlisted in the service of semantics.

Another motive for this study comes from within the preceding theory. Especially in Section 2.4., the idea was mentioned to view quantifiers as procedures, to obtain a better insight into their hierarchy of complexity. Nevertheless, there, another road was pursued, in terms of 'uniformity' constraints on truth value patterns for quantifiers. The present perspective provides the latter with a perhaps more solid and convincing background.

Evidently, quantifiers form one very special type of linguistic expression. But, as in Chapter 3, extension of notions and results is often possible to other, sometimes even all categories. Usually, such extensions require sensitivity to further model-theoretic structure than mere feature lists (as in the above). We shall introduce 'graph automata' for this purpose later on, when looking at conditionals and adjectives. This further step is similar to the move in contemporary mathematical linguistics toward 'tree automata', recognizing more structured syntactic objects than flat linear sequences.

There are some interesting general aspects to procedural semantics. Notably, the denotations studied here are 'intensional', in the sense that one and the same input/output behaviour for a quantifier may be produced by widely different automata. But also, this functional view of denotations (in the pre-Cantorian sense) still has obvious links with our earlier type-theoretical system (Chapters 3, 7). Such general issues will be discussed in the final section below.

Finally, the procedural perspective may also be viewed as a way of extending contemporary concerns in 'computational linguistics' to the area of semantics as well. Complexity and computability, with their background questions of *recognition* and *learning*, seem just as relevant to semantic understanding as they are to syntactic parsing.

8.1. QUANTIFIERS AND AUTOMATA

The action of a quantifier Q may be viewed as follows. For any pair of arguments A, B, it is fed a list of members of A, which can be tested one by one for membership of B. At each stage, Q is to be ready to state whether it accepts or rejects the sequence just read. In its simplest mathematical formulation, then, Q is just presented with finite sequences of zeroes and ones (standing for cases in $A - B$ and $A \cap B$, respectively), of which it has to recognize those for which the couple 'number of zeroes, number of ones' belongs to Q in the sense of Chapter 2. In other words, Q corresponds to a language on the alphabet $\{0, 1\}$. Moreover, its language is a rather special one, in that different enumerations of the argument set would not have affected acceptance: quantifier languages are *permutation-closed*. (Compare the similar phenomenon with the Lambek languages of Chapter 7.) Eventually, of course, the potential dependence on order of presentation may be put to good use in describing certain more contextual, or pragmatic expressions. For instance, the quantifier *every other* may be presentation-dependent in this sense (cf. Löbner, 1984).

Actually, the following discussion could also be couched in terms of a four-element alphabet, enumerating universes E with the four labels $A - B$, $A \cap B$, $B - A$ and $E - (A \cup B)$. But here, the two additional symbols would remain 'inert', only encumbering notation. Moreover, there are mathematical differences between languages with a two-element alphabet and more complex ones, that will be exploited to good effect below. For the relevant mathematical theory, here and elsewhere, the reader is referred to Ginsburg (1966), Hopcroft and Ullman (1979).

Now, one of the main themes in mathematical linguistics is the interplay between the notion of a language and its description by means of accepting (or generating) automata. And indeed, the familiar quantifiers turn out to be computable by means of well-known automata.

EXAMPLE. The quantifier *all* is recognized by the finite state machine of Figure 27.

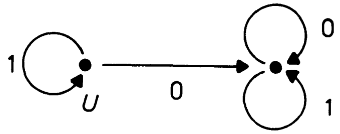

Fig. 27.

Here and henceforth, the starting state is the left-most one. Also, U is an accepting state. The automaton accepts only strings of 1's; i.e., those cases where all A-objects are concentrated in $A \cap B$. (Alternatively, it suffices to present the *all* language in Kleene regular set notation: 1*.) Similar two-state finite state machines will compute the remaining quantifiers in the Square of Opposition.

Other quantifiers may lead beyond regular languages and finite state machines:

EXAMPLE. The quantifier *most* is not recognized by any finite state machine. (This follows from the Pumping Lemma for regular languages. Cf. Hopcroft and Ullman, 1979.) Its language is context-free, however

— and so it can be recognized by a push-down automaton. The idea of the latter is simply to store values read, crossing out complementary pairs 0, 1 or 1, 0 of top stack symbol and symbol read, as they occur. When the string has been read, the stack should contain symbols 1 only. (Actually, the recipe does not quite fit the usual format of a push-down automaton. But such subtleties are postponed until Section 8.3.)

Thus, higher-order quantifiers may induce context-free languages. Interestingly, it is not all that easy to advance beyond this stage, finding non-contrived natural language quantifiers whose associated procedure would be essentially of the complexity of some Turing machine. One (not uncontroversial) example is (*relatively*) *many*, in the sense of Section 1.2. This case will be considered in somewhat more detail in Section 8.3. But on the whole, one finds natural language quantifiers, even the higher-order ones, within the context-free realm. Thus, they are essentially 'additive', in a logical sense to be explained below. There is some foundational significance to this observation, as *additive arithmetic* is still an axiomatizable (indeed, decidable) fragment of mathematics. The 'Gödel Border' of non-axiomatizability only arises in the next step, when adding *multiplication* to the system (cf. Mendelson, 1964). So, natural language shows a wise restraint in these matters.

8.2. FIRST-ORDER QUANTIFIERS AND FINITE STATE MACHINES

Definability

The examples in the preceding section motivate an obvious conjecture; which turns out to be justified:

THEOREM. All first-order definable quantifiers are computable by means of finite state machines.

Proof. Recall the tree pattern for first-order quantifiers described in Section 2.5. It consists of some arbitrary finite top triangle, followed by a 'Fraïssé threshold' at level $2N$. Now, the theorem is proved if the associated language can be shown to be regular.

The first top triangle declares some fixed finite set of sequences to be in Q; and all the corresponding singleton sets are of course regular. Then, the Fraïssé row adds a finite number of accepted patterns, of (at most) the following three types:

— exactly i occurrences of 1, at least j occurrences of 0,
— at least i occurrences of 1, at least j occurrences of 0,
— at least i occurrences of 1, exactly j occurrences of 0.

It remains to be checked that the latter types of language are indeed regular.

An example will make this clear. 'Exactly two 1-s, at least five 0-s' may be described as follows in Kleene notation. Take all (finitely many) possible distributions of five 0's over three slots formed by two 1 boundaries, and then 'fill up' with suitable iterations 0* wherever possible. □

Through this connection, existing results about finite state machines become available for first-order quantifiers, given by their definitions. For instance, it follows that, for any two such quantifiers, the question is decidable if they are equivalent.

Given the simple nature of finite state automata, a *converse* of the above theorem seems likely. But, there are obstacles.

EXAMPLE. The automaton of Figure 28 recognizes the (non-first-order) quantifier *an even number of*.

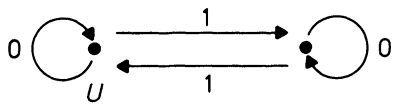

Fig. 28.

Again, U is an accepting state.

Thus, an additional restriction is needed to filter out the procedures corresponding to logically elementary quantifiers among all finite state machines. And in fact, the earlier examples suggest the following distinction. The machine graph for *an even number of* has a non-trivial *loop* between two states: something which did not occur with *all, some,* etcetera. Let us call a finite state machine *acyclic* if it contains no loops connecting two or more states. This notation has an independent motivation as well. The central concept in the monograph McNaughton and Papert (1971) is that of a *testable* regular language, being a simple construct out of languages recognized by 'counter-free' automata. The following result is in van Benthem (1985d):

> the testable regular languages are precisely those having an acyclic finite state recognizer.

Moreover, a condition is needed reflecting the earlier-mentioned permutation closure of our associated quantifier languages. A machine graph will be called *permutation-invariant* if the possibility of traveling from state U to state V by means of some sequence consisting of 0's and 1's implies that any permutation of that sequence will also force the passage from U to V. Note that the given automata for *all* and *an even number of* both had this property. Evidently, such automata recognize only permutation-closed languages. A converse will be proved later on in this section.

Evidently, permutation closure is a strong restriction on classes of languages. It should be kept in mind, however, that we can always decide to drop it, when studying more involved linguistic constructions. On the other hand, the restriction is a mathematically interesting one. Contrary to initial expectations, it does not seem to collapse the general theory of automata and languages; leading rather to new versions of old questions. Some examples will appear in what follows.

As an illustration of the above two conditions, it may be observed that permutation-invariant and acyclic finite state machines are 'convergent': all non-trivial continuing paths end eventually in one single absorbing state. This fact will be significant (although it is not actually used) in the proof of the following result.

THEOREM. The first-order definable quantifiers are precisely those which can be recognized by permutation-invariant acyclic finite state machines.

Proof. First, a closer look is required now into the automata associated with first-order quantifiers. Interestingly, the Tree of Numbers (Section 2.2.) now proves to be computationally significant too.

Let Q be a first-order quantifier in the tree, with its Fraïssé pattern. Now, interpret this tree structure *itself* as a graph for a machine M_Q with infinitely many states corresponding to the tree nodes. Its starting state is $(0, 0)$. The transition arrows for reading a zero go from states (a, b) to $(a + 1, b)$, those for reading a one from (a, b) to $(a, b + 1)$. Finally, the accepting states will be those in Q.

CLAIM. M_Q accepts a finite sequence s if and only if the couple $0(s)$, $1(s)$ = number of zeroes in s, number of ones in s, belongs to Q.

Proof. By induction on the length of sequences s, it may be shown that reading s will take M_Q from its starting state to the state $0(s)$, $1(s)$. The assertion then follows. □

Now, with a first-order pattern for Q, this infinite machine accepting device may be reduced to an equivalent *finite* state machine. Recall the earlier Fraïssé threshold, at level $2N$. Above it, the state graph remains as before; on it, one concludes with Figure 29.

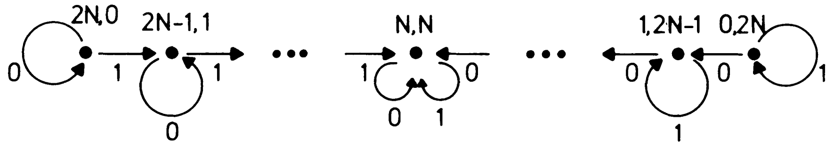

Fig. 29.

It is easy to check that this modified machine still accepts the same language.

Finally, inspection shows that machine graphs of the latter kind are both acyclic and permutation-invariant. This concludes the first half of the argument.

The converse direction of the theorem follows from the inductive character of acyclic finite state machines. In such automata, for any accepting state, there is only a *finite* number of types of path leading to it (from the starting state), driven by successive transitions of the four types 0, 1, 0* or 1*. Thus, such states accept sequences of essentially the following forms:

exactly i 1/exactly j 0, exactly i 1/at least j 0,
at least i 1/at least j 0, at least i 1/exactly j 0.

Now, all these types by themselves are first-order definable — and hence so is their disjunction (which is the 'yield' of the accepting state considered). Finally, the whole machine itself again accepts the disjunction of everything admitted in its individual accepting states. □

Some further illustrations of this method may be obtained by comparing 'tree-based' automata for first-order quantifiers with those found ad hoc. For instance, the tree approach for *all* would give a 3-state automaton, which may be simplified to the earlier 2-state one. Conversely, given automata for first-order quantifiers may be 'normalized' into a tree-like shape.

The above theorem can be generalized to the more general class of merely quantitative quantifiers (not necessarily obeying CONS). The same notion of first-order definability may be employed then, but there is a four-letter alphabet, as explained above.

Finally, the question remains to describe the class of *all* quantifiers computed by finite state automata. Our conjecture is that these are all definable in a first-order language augmented with suitable 'periodicity quantifiers', such as *a k-multiple number of.* (For related definability results, and a first significant use of finite automata in logic, see Büchi, 1960.)

Fine-Structure

In addition to global definability questions, matters of finer detail are relevant too in the study of actual quantifiers. For instance, one particularly important case is that where the states correspond directly to semantic *truth values*. In the present setting, this leaves only finite state machines having two states, one accepting, the other rejecting.

THEOREM. The permutation-closed languages or quantifiers recognized by finite two-state automata are the following: *all, some, no, not all / an even number of, an odd number of, all but an even number of, all but an odd number of,* together with the extreme cases *empty sequence only, non-empty sequences only, sequences of odd length, sequences of even length.*

Proof. By brute force; enumerating all eligible automata. □

This enumeration result is reminiscent of earlier ones in Section 2.4. For instance, one *Uniformity* condition stated that a truth value change under the influence of a transition (a, b) to $(a + 1, b)$ or $(a, b + 1)$ will be the same throughout the tree of numbers. But, this amounts to sanctioning operating with one 'true' and one 'false' state, with fixed responses for their departing zero and one transition arrows.

Additional states may be needed for quantifiers of higher complexity.

EXAMPLE. (*Precisely*) *one* can be recognized by an automaton with three states; but no less.

Of course, all quantifiers mentioned can also be correlated with machines having large numbers of additional states. Thus, there is an issue of *minimal representation*; as in the well-known Nerode Theorem, representing input/output functions in terms of state transition machines. In fact, the Nerode representation method could be applied to first-order quantifier languages too.

EXAMPLE. Consider the class ONE of sequences with exactly one occurrence of 1. The crucial Nerode equivalence relation is the following:

> $s_1 E s_2$ if, for all sequences s, s_1-followed-by-s is in ONE iff s_2-followed-by-s is in ONE.

Its equivalence classes are the following three: sequences with *no, one* or *at least two* occurrences of 1. These will then be the states in the minimal representation, with the transition function defined by the stipulation that

> equivalence class of s, read symbol x $\Rightarrow$ equivalence class of s-followed-by-x.

The outcome is the automaton of the previous example.

As an application of this technique, the converse may be proved of an earlier observation. For any permutation-closed regular language, its Nerode recognizer will be 'permutation-invariant' in the sense of the characterization theorem for first-order quantifiers.

Infinite Models

When *infinite* models are considered, matters become less perspicuous. Of course, the earlier machines can operate on infinite sequences just as well as on finite ones — but, the problem is to find a well-motivated convention for acceptance, given a non-ending sequence of states traversed.

One possibility is to admit only those infinite sequences which cause the machine to remain in some accepting state after a finite number of steps. By inspection of the earlier Fraïssé recognizers, it may be seen that all first-order definable quantifiers are recognized in this way. But also, certain genuinely infinitary cases are admitted.

EXAMPLE. The following two-state machine recognizes the quantifier *almost all*, in the sense of 'with at most finitely many exceptions' (see Figure 30).

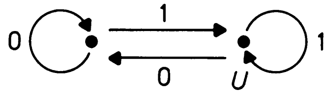

Fig. 30.

Here, U is an accepting state. Note that this machine is not permutation-invariant — although its infinite language is permutation-closed.

Another essentially infinite quantifier of this kind is *virtually no*, in the sense of 'only finitely many'. Indeed, it would be of interest to chart the effects of all our earlier two-state machines here.

What distinguishes the first-order quantifiers from the infinitary cases is that their rejection is also finitary: non-acceptable sequences lead the machine to some stable non-accepting state after a finite number of steps. We conjecture that, on the sets of countably infinite sequences, the first-order quantifiers are *precisely* those with a finitary acceptance and rejection behaviour.

Further relaxations of the acceptance convention are possible, such as allowing recurring *cycles* of accepting states — or even recognizing by certain infinite patterns. Clearly, the 'yield' of our machine classes will be rather sensitive to such stipulations. A deeper investigation of these phenomena must be left to another occasion.

8.3. HIGHER-ORDER QUANTIFIERS AND PUSH-DOWN AUTOMATA

With higher-order quantifiers, finite state machines usually become insufficient as computing devices — and the next level in the machine hierarchy is needed, that of push-down automata, handling a stack in addition to changing state when reading new input. The usual definition of these machines is rather restricted: the transition function takes only the top stack symbol into account, recognition is by empty stack only. However, by a suitable re-encoding of instructions, it may be assumed that the machine can keep track of a *fixed finite* top part of the stack, with a final read-off convention performing a finite state machine check on the stack contents. One way of simulating these effects in the orthodox format is by having new states $q, \langle q, s_1\rangle, \ldots, \langle q, s_1, \ldots, s_k\rangle$ encoding tuples of old states q with up to k top stack positions. Then, suitable ϵ-*moves* are to allow a trade-off between encoded sequence and top stack symbols displayed. Finally, another set of ϵ-moves will allow us to carry out the finite state machine instructions without consuming any more input.

For instance, the earlier-mentioned quantifier *most* would have an awkward corresponding automaton in the original sense, whereas it has an easy intuitive one in this more liberal version. Significantly, though, writing a context-free grammar for the *most*-language is not an entirely trivial exercise.

Before passing on to further examples, some general remarks should be made on push-down automata. First, the present discussion is couched in terms of *non-deterministic* machines of this kind. Deterministic push-down automata recognize a smaller class of languages, less natural than the context-free ones. Incidentally, the examples from the literature for this non-inclusion all lack permutation-closure. But, certain disjunctively defined quantifiers in our sense would seem to be examples also. Even so, most basic quantifiers considered here are 'deterministic' in an intuitive sense — and so, other types of automaton may be relevant too.

Recognition of sequences is made by inspection of stack content. One could also pursue the analogy with the finite state machine case, recognizing by designated state: an equivalent procedure, in general. Either way, the earlier parallel between machine states and semantic truth values becomes less direct.

Besides *most*, other prominent higher-order quantifiers in natural language are *almost all, many, few, hardly any*, etc. Evidently, the meaning of these expressions is underdetermined — as is the case, indeed, with *most* itself. To get an impression of their complexity in the present terms, then, reasonable formal 'explications' are to be considered. The following seem fair approximations of their *spirit*:

many: 'at least one third', *almost all*: 'at least nine-tenths'. Thus, on this type of reading, the above quantifiers all express proportionality.

Now, the complexity of the latter phenomenon is essentially context-free. A simple illustration is the following.

EXAMPLE. A push-down automaton recognizing *at least two-thirds* will operate as follows. It keeps track of two top stack positions, checking with the next symbol read. In case it encounters a symbol 1, with 1, 0 occurring on top (in any order), it erases the latter two, and continues. Likewise, with a symbol 0 read and two symbols 1 on top. In all other cases, the symbol read is simply stored on top of the stack. At the end of the process, the machine checks if the stack contains symbols 1 only: only then does it recognize the sequence just read.

This automaton recognizes the right strings — by a simple combinatorial argument. A similar procedure, somewhat more involved, will recognize *almost all*. The case of *many* is instructive too.

EXAMPLE. *At least one-third* may be recognized by a more curious automaton in the above spirit. The machine now keeps track of three top stack positions, having the following (non-deterministic) instructions. Symbols read may be pushed onto the stack. But also, the following ϵ-moves are allowed (without consuming input): combinations of two 0/one 1, one 0/one 1, or one 1 only, may be erased from the stack. This time, final recognition is simply by empty stack.

The arithmetical idea behind this procedure is that all couples (a, b) with $a \leqslant 2b$ are generated, starting from (0, 0), by means of the following operations: add (2, 1)/add (1, 1)/ add (0, 1). Obviously, if a string is accepted by the above machine, its 'occurrence couple' (number of zeroes, number of ones) has such an arithmetical form. Conversely, if a string has an occurrence couple a, b of the form $x(2, 1) + y(1, 1) + z(0, 1)$, then a judicious process of reading and erasing will produce an empty stack in the end. (A more general procedure in this vein will be stated in the proof of the main theorem below.)

Many other examples can be analyzed in the same way. We shall now investigate the situation from a higher logical standpoint.

Arithmetical Definability

When viewed arithmetically, the above quantifiers all express extremely simple conditions on the two variables a (number of zeroes) and b (number of ones). For instance, *all*: $a = 0$/*some*: $b \neq 0$/*one*: $a = 1$ /*most*: $a < b$/'*many*': $a \leqslant b + b$. Thus, simple first-order quantifiers employ just a, b, fixed natural numbers and identity, others add 'smaller than', and eventually, addition is needed. These were atomic cases; but, e.g., the earlier *an even number of* would require a quantified formula: $\exists x\, b = x + x$. In fact, once arbitrary first-order conditions on a, b involving $+$ are considered, the notion $<$ becomes definable, and hence so are all specific natural numbers n. Thus, we are dealing with first-order formulas $\varphi\,(=, +, a, b)$, interpreted as standard arithmetical statements. Now, we shall derive a general characterization of the present area of complexity.

But first, one natural preliminary notion to consider is first-order arithmetical definability in $=$, 0 and S (*successor*) only. It is easy to see that all first-order definable quantifiers (in the sense of Section 2.5.) are definable here already. (For instance, *at least two* becomes: $b \neq 0 \wedge b \neq So$.) The converse fails, however. E.g., the arithmetical formula $a = b$ defines *exactly one half*, which is not even recognizable by a finite state machine. But also, not every finite state quantifier is definable here: witness the earlier *an even number of*. In this light, the better behaviour of additive arithmetical definability, to be studied now, is the more surprising.

Here are some preliminary notions and results. First, there is the fundamental *Parikh Theorem*, stating that each context-free language L, in alphabet $a_1, \ldots, a_k$, say, induces a *semi-linear* set of k-tuples of natural numbers

$$\{(\text{number of symbols } a_1 \text{ in } s, \ldots, \text{number of symbols } a_k \text{ in } s) \mid \text{all sequences } s \text{ in } L\}.$$

Here, a 'semi-linear' set of k-tuples is a finite union of *linear* sets, consisting of all k-tuples produced by a schema of the following form:

$$(m_1, \ldots, m_k) + x_1(m_{11}, \ldots, m_{1k}) + \cdots + x_n(m_{n1}, \ldots, m_{nk});$$

with $x_1, \ldots, x_n$ non-negative integers.

All semi-linear sets correspond to first-order additively definable k-ary relations on the natural numbers — as the given arithmetical schema can be written out in first-order terms. E.g., $(1, 2) + x(0, 1) + y(2, 2)$ would become $\exists xy(a = 1 + y + y \wedge b = 2 + x + y + y)$. Hence, a first conclusion may be drawn:

THEOREM. Every quantifier computable by a push-down automaton is first-order additively definable.

Proof. The set of 'Parikh couples' for the (context-free) language of the quantifier is semi-linear, and hence definable in the required sense. Moreover, the language being permutation-closed, the latter definition fits it exactly. $\square$

In general, however, Parikh's Theorem cannot be converted, not even for permutation-closed languages. A standard counter-example on a

3-symbol alphabet is the following non-context-free set: 'all sequences with equal numbers of occurrences of the three symbols'. The corresponding arithmetical predicate is obviously semi-linear; but, the language is not context-free. (When intersected with the regular $(abc)^*$, it produces the well-known counterexample $a^nb^nc^n$.)

Fortunately, the present *binary* alphabet admits of a more interesting conclusion. (That such restrictions may produce strong effects is not unknown: compare the theorem that, on a 1-symbol alphabet, every context-free language is already regular.) First, observe that all previous examples were semi-linear. This was indicated already for the case *at least one third*, being $(0, 0) + x(0, 1) + y(1, 1) + z(2, 1)$. But also, e.g., *at least two-thirds* has such a form: $(0, 0) + x(0, 1) + y(1, 2)$. The latter is the simpler one; a difference which showed up already in their respective automata.

To proceed more generally, a result is now needed from Ginsburg and Spanier (1966), who proved that all first-order additively definable predicates are semi-linear. Their idea was to use *Presburger*'s early description of these predicates (cf. Mendelson, 1964, pp. 116/7). By the method of 'quantifier elimination', the additive predicates turn out to be equivalent to all Boolean combinations of 'atomic' formulas

$$t_1 = t_2,\ t_1 < t_2 \text{ and } t_1 = t_2 \ (\text{mod } k) \ (\text{i.e., } \exists x(t_1 = t_2 + x \vee t_2 = t_1 + x))$$

Here, t_1, t_2 are terms involving individual variables, 0, S and $+$. Now, all such Presburger normal forms can be shown to be semi-linear. (This is not trivial. For instance, it is not obvious that the intersection of two semi-linear sets is again semi-linear.)

Thus, first-order additive predicates correspond to semi-linear sets — and it remains to find suitable automata for the latter, generalizing the earlier examples. In view of earlier observations, the restriction to a 2-symbol alphabet must play some essential role here.

THEOREM. Every first-order additively definable quantifier is computable by a push-down automaton.

Proof. In view of the above, if suffices to find an accepting push-down automaton for every language having a semi-linear set of occurrence *couples*. The following construction is due to Stan Peters and Bill Marsh.

First, attention may be restricted to *linear* predicates, as finite unions of context-free languages are themselves context-free. Thus, suppose our predicate has the form

$$(m_1, m_2) + x_1(m_{11}, m_{12}) + \cdots + x_n(m_{n1}, m_{n2}).$$

Let N be the maximum of all natural numbers m_i, m_{ij} involved. Our automaton will have $2(N.N)$ states, being all couples of the forms

$$(i, j), (i, j)^* \quad (1 \leqslant i, j \leqslant N).$$

What these encode will become clear presently. The instructions are the following:

Reading: reading a symbol 0 in state (i, j) (with $i < N$), either go to $(i + 1, j)$, or remain in (i, j) adding a 0 to the stack;
reading 0 in state (N, j), remain in (N, j) and add 0 to the stack.

Likewise, for reading a symbol 1.

Similar moves are allowed for the $*$-states: here, and wherever appropriate.

Exchanging: the following ϵ-moves are possible:
from $(i + 1, j)$ to (i, j), pushing a 0 onto the stack;
likewise, with $(i, j + 1)$ and a symbol 1.

Conversely, if $i < N$, a 0 may be popped from the stack, going from state (i, j) to $(i + 1, j)$;
likewise, for $j < N$ and popping a symbol 1.

Lowering: when $i \geqslant m_{k1}$, $j \geqslant m_{k2}$, it is possible to jump from state (i, j) to $(i - m_{k1}, j - m_{k2})$.

Crossing: when $i \geqslant m_1$, $j \geqslant m_2$, it is possible to jump from state (i, j) to $(i - m_1, j - m_2)^*$.

Finally, the initial state is (0, 0) — and recognition is by empty stack and designated state $(0, 0)^*$.

The point of these instructions lies in the following

CLAIM: At each stage in the computation, with state (i, j), there exist numbers $x_1, \ldots, x_n$ such that (1) $i +$ 'the number of symbols 0 in the stack' equals 'the number of symbols 0 read' $-x_1 \cdot m_{11} - \cdots - x_n \cdot m_{n1}$, and (2) $j+$ 'the number of symbols 1 in the stack' equals 'the number of symbols 1 read' $-x_1 \cdot m_{12} - \cdots - x_n \cdot m_{n2}$. And, for $*$-states, a similar assertion holds, but with two more subtractions: viz. for m_1, m_2, respectively.

The *proof* of the claim is by induction on the number of admissible moves in a computation. Clearly, each of them preserves this invariant. Moreover, then, once in the final state, with the sum on the left-hand side equal to zero, the sequence processed must have had the above arithmetical character. Thus, only correct sequences are recognized.

Conversely, a judicious sequence of admissible moves will accept any string with the correct amounts. To see this, the following assertion may be proved, by induction on the sum $x_1 + \cdots + x_n$:

CLAIM: If the stack has either all zeroes or all ones, the state is (0, 0), and the occurrence totals of symbols 0, 1 in stack plus sequence still to be read are of the original linear form, then the machine will proceed to recognition. And likewise, in state (0, 0)* without the initial factor (m_1, m_2).

Proof. For $x_1 + \cdots + x_n = 0$, the only non-trivial task is to recognize a string of m_1 zeroes and m_2 ones. By obvious steps, state (m_1, m_2) can be reached here, popping and reading — and then, one crossing produces the desired final state.

For $x_1 + \cdots + x_n > 0$, the machine is to continue reading until the first couple of occurrence numbers (m_{i1}, m_{i2}) is exceeded (or perhaps (m_1, m_2) itself) — as is bound to happen. Say, we started reading symbols 1, stacking all of these, while using symbols 0 to raise the state. Then, transferring enough symbols 1 from stack to state, we arrive at the state (m_{i1}, m_{i2}) — from where we can drop to (0, 0), leaving a 'homogeneous' stack. □

From the claim, it follows at once that all correct strings will indeed be accepted. □

Comment: how to read and act in order to recognize a sequence of zeroes and ones with occurrence numbers given by

$$(m_1, m_2) + x_1(m_{11}, m_{12}) + \cdots + x_n(m_{n1}, m_{n2}).$$

Suppose that we have a homogeneous stack of symbols 1, are in state (0, 0), and the correct invariant holds (i.e., the totals of occurrences for 0, 1 in $\langle$state + stack + sequence to be read$\rangle$ still satisfy the above linear form). We want to read ahead to pick up enough 0's and 1's to reach the first m_{i1}, m_{i2} (or perhaps m_1, m_2 itself) in occurrence totals for $\langle$state + stack$\rangle$. (Then, we 'drop state', or 'cross over', and repeat the process.) Now, in order to get this m_{i1}, m_{i2}, we may have to pick up *too many* symbols 0 or 1 (though not both). E.g., suppose that it is too many 0 ('too many 1' will just remain on the stack). Then, we first transfer all symbols 1 from stack to state (this must be possible: if there were too many 1's already there, then there would have been no need to pick up more than the required number of 0's) — putting all symbols 1 read into the state, and *enough* symbols 0 (stacking the others).

Why does this work for two-symbol alphabets only? With, say, three symbols, one cannot maintain a 'homogeneous' stack — which is vital to the argument.

Our own original approach to this theorem proceeded by enumeration, looking for suitable push-down automata directly for each of the above Presburger forms. Still, this approach also provides some interesting concrete examples. Arithmetical predicates can be represented in the tree of numbers, and thus, (Boolean compounds of) Presburger formulas are often seen to reduce to manageable geometric patterns.

In all, the above theorems form an elegant characterization of all quantifiers computed by push-down automata.

Natural Language Quantifiers / Computational Concerns

As there is such a host of push-down automata, and so few linguistically realized quantifiers exploiting them, *additional* constraints seem of interest. For instance, earlier on, a possible restriction was mentioned to *deterministic* automata of some sort. One possibility here is to use deterministic push-down automata scanning some fixed finite top portion of the stack, provided with an additional facility for a final finite state machine check of stack contents. This type of automaton seems appropriate for all of the earlier concrete examples. (In particular, thus, *at least one-third* can be recognized deterministically after all.)

In this light also, another topic to be investigated are connections between the 'semantic' conditions on denotations from earlier chapters, and natural restrictions on machine instructions. In other words, what are the *computational effects* of our original generalized quantifier notions? One example is found in Chapter 10, where an equivalence is proven between quantifiers that are of *minimal count complexity* and those that are *continuous*, in a strong sense, related to that of Section 2.4. What the latter amounts to, in the tree of numbers, is this:

in each of the three main directions of variation in the Tree, i.e., increasing a (traveling along a /-line), increasing b (traveling along a \-line), playing off a and b (traveling along a horizontal line), the quantifier is to experience at most one truth value change. (Cf. also van Benthem, 1985a, for several uses of this notion.)

All previous examples of basic quantifiers satisfy this strong condition. Still, *uncountably* many others share it: being all patterns of similar rows $---\cdots+++$ (or $+++\cdots---$), whose true/false boundary shifts at most one position at a time, when descending the tree. But, upon combination with the above computational requirement of *linearity* — and modulo *Variety*, our ever-present simplifying assumption (cf. Sections 1.3., 2.1.), only a countable number of very regular patterns remains, viz. those dividing the Tree into a true and a false part with a 'periodic' boundary. But, there is a case for yet another restriction. The conditions of Continuity and Variety are *symmetric*: holding for Q if and only if they hold for its negation. Thus, it seems of interest to strengthen linearity to *bi-linearity*, in the appropriate sense. Then, the following classification results of 'strongly computable' quantifiers:

THEOREM. The bilinear continuous quantifiers are precisely those of the following forms ($n = 0, 1, \dots$):

at least 1/n + 1, at most n/n + 1,
less than 1/n + 1, more than n/n + 1.

For $n = 0$, this gives the four quantifiers in the Square of Opposition. For $n = 1$, one gets *most, not most, least* and *not least*; as in Section 2.4. Beyond, a computational hierarchy arises, as expected.

Proof. It suffices to consider one case only, the others being analogous. Suppose that $(0, 0) \in Q$, with $(1, 0) \notin Q$, $(0, 1) \in Q$. Because Q is linear, this must mean that $(0, 0)$ is its basic case, while it allows steps of $+x(0, 1)$. Moreover, non-Q has a basic case $(1, 0)$. Also, because of Continuity in /-lines, $(2, 0) \notin Q$, and so forth: that is, non-Q allows steps $+x(1, 0)$. Now Q, being linear, has only finitely many possibilities for occupying positions on the line adjoining the non-Q left-hand edge; say, at most down to $(n, 1)$. By Continuity, it will then also occupy $(0, 1), \dots, (n, 1)$ — but nothing beyond that. Thus, non-Q can make the jump $(n + 1, 1)$. By bi-linearity then, as well as the already described properties of Q and non-Q, their boundary is fixed, to form the pattern of *at least 1/n + 1* ($n = 0, 1, 2, \dots$). □

Again, questions of *fine-structure* arise here too. For instance, one natural class of push-down automata to consider are those with two states, whose stack alphabet is restricted to 0, 1. Even there, a hierarchy of further possibilities exists, depending on which actions are permitted in rewriting the stack, and which final reading-off convention is chosen. The *most* machine exemplified about the simplest case here, with additions or removals of at most one stack symbol at a time, and a final convention which essentially just inspects the top of the stack. Of the latter variety, there are only few representatives — but, an enumeration is omitted here.

Are there any natural language quantifiers beyond the context-free realm? Of course, it is easy to make up examples, such as *a square number of*. A more serious candidate is the earlier (*relatively*) *many*, in the sense that $b/(a + b) > (b + c)/(a + b + c + e)$, or equivalently, that $a \cdot c < b \cdot e$. (For these numbers, see Figure 6.) This notion is not computable by a push-down automaton. For, the above predicate is not semi-linear, as it essentially involves multiplication.

Still, this example is not a pure quantifier in the narrower sense; and we know of no natural non-additive cases in the latter area. (Another more complex candidate is the *ternary* determiner *more* A_1 *than* A_2 *are* B (cf. Section 1.8.), whose representation seems to require more than two symbols — in which case the preceding theorems would not apply. But, upon closer inspection of its Venn diagram, only *two* zones are involved after all. And the same is true, e.g., for *as many* A_1 *as* A_2 *are* B.) Thus, as was observed in Section 8.1., the quantifier system of natural language seems to stop just short of the arithmetical border-line where incompleteness and undecidability set in. Thus, while relatively strong, it still enjoys the theoretical properties of additive arithmetic. In particular, many questions about behaviour or comparison of even higher-order quantifiers must be *decidable*.

Actually, some caution is in order here. The above arguments only provide an effective road for obtaining push-down automata for arithmetically given quantifiers. The converse requires some supplementary reasoning. First, given a push-down automaton computing a quantifier, an equivalent context-free grammar is to be found. The standard proof of this correspondence in the literature is effective. Then, the proof of Parikh's Theorem provides an effective method for assigning a semi-linear set to a grammar. And finally, transcription of the latter into the format of additive arithmetic is effective too. Thus, for instance, *equivalence* of quantifiers as procedures (in the present sense) is indeed decidable. This is somewhat surprising, as equivalence of push-down automata in general is known to be undecidable. Evidently, the restriction to permutation-closed languages plays a major role in this respect.

8.4. OTHER LINGUISTIC EXPRESSIONS AND GRAPH AUTOMATA

Many details of the preceding sections are tied up with special features of determiners and quantifiers. Nevertheless, other types of expression are also amenable to procedural analysis.

An intensional parallel to the quantifier case arises with *conditionals*, when treated in the manner of Chapter 4. Then, as we have seen, universes will have to carry additional 'accessibility' or 'relative similarity' patterns, which requires a more sensitive kind of automaton to be used.

For instance, on the account of Stalnaker, essentially, the antecedent situations for a conditional form a *strict linear order*, the highest of which is to satisfy the consequent. A simple push-down automaton could do the required search here, keeping one occasion stored, and reading positions one by one. Its instruction would be to ignore lower objects in the order as they are encountered, while substituting higher ones. In the final position, a 'consequent check' is made. A more complex graph may be needed on Lewis' account of conditionals, viz. an *almost-connected* order of antecedent occasions, with a satisfaction clause as in Section 4.7. In the relevant computation, there may be larger 'indifference classes' now; but a similar idea works. Lower objects are disregarded, higher ones take precedence, while equally high ones falsifying the consequent take precedence over the others. Again, the same final check suffices.

These descriptions presuppose a somewhat enlarged concept of automaton already: one which can store objects surveyed, and test for their binary relations, as well as unary properties. Alternatively, with a finite alphabet of possible outcomes for our test or tests, the structure surveyed may be viewed as being a 'letter graph', rather than a linear word. That theoretical behaviour changes in this area follows by comparison with Section 8.2. The Lewis truth condition is *first-order* in A, B and the order R — but nevertheless, it cannot be computed by an ordinary finite state machine; as the latter lacks a facility for making R-comparisons. Still, the issue of a starting point of description is relatively arbitrary. More interesting here are *jumps* in complexity. For instance, when the occasion graph is no longer almost-connected, but, say, a mere strict partial order (as in Chapter 4), computation of the Lewis clause becomes more difficult; as an arbitrary store of individuals may have to be checked now and then. No push-down automaton in the above sense can do the latter job.

Another example is the proposal in Suppes (1982) to treat certain adjectives as procedures. For instance, the superlative *tallest* invites us to check if an individual x is in top position in the 'taller' order on the universe. (Compare the above Lewis case.) Here, a finite state machine suffices, as only unary checks are needed of individuals surveyed: '$\lambda x \cdot Rxy$', '$\lambda x \cdot Ryx$'. Interestingly, assigning the predicate *tall* itself involves a more complex procedure, on Suppes' view, viz. to check that fewer individuals precede our individual than succeed it in the 'taller' order. For this purpose, a push-down automaton will be needed.

While in this specific area, adjectival *modifiers* may be of interest. (Recall also the determiner modifier *almost*.) For instance, various 'numerical' proposals have been made for measuring the force of *very*, as it applies to adjectives — some of them *quadratic* (cf. Hoeksema, 1984). Here again, the earlier addition/multiplication boundary comes in sight.

With modifiers, one arrives at semantic *operations*, rather than properties or relations. To make the earlier automata compute operations as well, they will have to produce output — something which can be arranged in various ways (cf. Hopcroft and Ullman, 1979). Perhaps the simplest idea in the present setting would be to use the earlier machines as 'sluicing devices', deciding whether or not to pass certain items read. Used in this way, e.g., the automaton of Figure 28 passes 'every other B' from among the A's. Note that the outcome here depends on the order of presentation of input: 'permutation invariance' is again a non-trivial constraint on automata computing operations. In general, of course, more complex actions will have to be performed. This topic will not be pursued here.

Graph Automata

To obtain further insight into this general area, it seems of interest to study a reasonably simple, and yet widely applicable type of graph automaton. Yet, searching arbitrary graphs is a complex task. Fortunately, both the above cases, as well as parallel developments in mathematical linguistics, suggest a strong restriction, viz. to *finite acyclic graphs*, perhaps with a *root*. These are finite rooted graphs in which no loops occur. A pleasant property here is their 'inductive' behaviour: starting from bottom nodes, then passing on to nodes all of whose predecessors were surveyed previously, etc., the whole graph can be surveyed in a unique manner. This notion arises, for instance, with enriched syntactic trees in grammar: as long no loops have arisen in 'anaphoric annotation' and related processes, compositional bottom-up interpretation will still be feasible (cf. van Benthem, 1983b).

In line with the above, a *graph automaton* will be a device that can test nodes for features, moving up the graph (in parallel fashion) from daughters to mothers. More specifically, at each node, its feature combination is inspected, after which a check is made of the results already obtained on the daughters, the nature of which may depend on the feature values registered. As there is no limit to the number of daughters (we are working in semantic models, not on syntactic structure trees), an infinite enumeration of cases might be necessary here in stating the desired transition functions. In the spirit of this chapter, however, we want a uniform checking routine, finitely specifiable in advance: in the simplest case, a *finite state machine*. This machine will read 'final state markers' left on the daughters in the previous round, subject to an *entry convention* specifying, for each feature combination appearing on the original node, in which state to start up the machine. The latter's final state will again be marked on the original node itself, this being the result of the present round. Finally, the whole graph is accepted if our automaton stops on its root in an accepting state.

EXAMPLE. Computing first-order properties of graphs.

Consider the following condition on graphs with one feature A:

$$\forall x(Ax \rightarrow \forall y(R^+xy \rightarrow \neg Ay)) \wedge \forall x(\neg Ax \rightarrow \forall y(R^+xy \rightarrow Ay)) \quad \text{('alternation')}.$$

Here, R^+ expresses *immediate* dominance. A recognizing finite state automaton has three states:

> a_1 ('accept, with top position A'), a_2 ('accept, with top position $\neg A$'), b ('reject').

When reading feature A on a node, it enters state a_1, and starts scanning final states already reached on the daughters of that node. (Think of the former as displayed in some way. Some theorists like their machines to leave coloured pebbles on their trail, as in the story of Hans and Gretchen.) Now, when reading some a_1 or b, the machine enters state b, never to leave it. Only with encounters of a_2 only, will it produce state a_1 for our original node. The case of initial feature non-A is symmetrical.

Alternatively, one can give the finite state machine an *input/output* alphabet, so as to circumvent the subtlety of 'states displayed' versus 'states in action'.

Various further theoretical questions about this kind of automaton are investigated in van Benthem (1985g). Notably, several issues of *definability* arise. Any graph automaton M recognizes a certain class of finite 'featured graphs', or equivalently, it computes a property π_M of the latter. Given any automaton M, can its corresponding π_M be defined in some perspicuous fashion?

One general answer is this. For every M, π_M can be described by a *monadic existential second-order* sentence having the graph relation R and its feature predicates as first-order parameters. The idea is essentially to assert that there exist unary predicates on the graph (marking states assumed by M while processing the latter) such that 'suitable transcription into first-order terms of: entry convention, transition diagram for M, state marking convention, accepting state at the top node'.

But, the preceding definition of π_M merely mimicks the presentation of M itself in an unenlightening way. The point about this presentation is that it is *recursive* (and hence *implicit*): acceptance of nodes is made dependent on their features plus accompanying conditions on (non-)acceptance patterns among their daughters. What we would prefer is an *explicit* definition of π_M, if possible. To take a specific case, suppose that M's acceptance of a node depends on its features plus some *first-order* condition on (non-)acceptance of its daughters. (This case will occur when the finite state checking routine is *acyclic* in the sense of Section 8.2.) Then, can π_M be described as a first-order condition too, on the graph order plus the feature predicates? In general, it is hard to find such explicit descriptions.

EXAMPLE. A simple first-order recursion.

Machine M acts as follows: (i) if a node has feature A, then M ends up accepting it iff no predecessor was accepted, (ii) if the node lacks feature A, then M accepts it iff some predecessor was accepted.

Consideration of simple cases (close to the bottom leaves) will provide some impression of what is going on, but a general explicit description of the class of trees recognized remains elusive.

A great improvement occurs, however, when the machine is allowed to inspect *all* predecessors of any node, instead of only daughters. In that case, the implicit description of the graphs accepted may be formulated as a recursion over all predecessors in a class of well-founded structures. Then, techniques from 'modal provability logic' can be applied (cf. Smoryński, 1984), in particular the *De Jongh—Sambin Fixed Point Theorem*. The latter gives us an algorithm for converting implicit definitions into explicit ones, at least for simple (modally definable) cases. E.g., for the previous example, the condition obtained is the (correct)

$$(Ax \wedge \forall y(Rxy \rightarrow \neg Ay)) \vee (\neg Ax \wedge \exists y(Rxy \wedge Ay)).$$

In general, even with acyclic checking routines, our graph recursions will employ arbitrary numerical first-order quantifiers over predecessors (not just the 'modal' $\forall$, $\exists$). Moreover, there will often be a multiple recursion, defining a family of accepting state predicates at the same time. As De Jongh has shown in the meantime, the above algorithm can be extended to deal with all these cases. (In fact, it works even beyond the first-order realm; but, there are limitations too.) So, we can answer the question raised before:

if a graph automaton is acyclic, then its associated graph property is first-order.

One obvious *converse* question now becomes if all first-order properties of featured graphs are computable in our sense. For the case of inspecting all predecessors, this is refuted in van Benthem (1985g). For the case of inspecting daughters only, the answer may well be affirmative. In addition, there are analogous questions concerning the monadic second-order formalism introduced above.

These illustrations will have conveyed the flavour of this kind of automaton, and the agenda for research.

8.5. PROCEDURAL SEMANTICS

Taking procedures for actual denotations of certain linguistic expressions turns out to be a fruitful perspective. Indeed, Suppes has suggested that the idea extends to all of natural language — citing even a 'procedural' view of proper names, as 'criteria of identity'. When the enterprise is advocated in this generality, some caution is due.

First, the proper name case may point at a possible confusion. Of course, even on the orthodox view, in every interpretation of language, there is some functional connection between linguistic items and their denotations — and this function may come with a procedure. Thus, *Julia* denotes the girl Julia; but, a procedure may be necessary to recognize her if you saw her. Put differently, a procedural view might be appropriate to Fregean *senses*, while leaving traditional denotations undisturbed. But, the procedural perspective is not external in this sense. It is supposed to apply 'inside models', so to speak. But even there, a facile over-applicability threatens. For, in our *type-theoretic* framework (Chapters 3, 7), virtually every denotation is a function — and as such, may be thought of as a procedure. Thus, the present approach only acquires some bite by descending from this general level.

But then, semantic population problems only seem to get worse. One function, in the set-theoretic input/output sense, will correspond to many intensionally different procedures for computing it. Fortunately, however, this intensional move also suggests a more 'categorical' perspective upon type-theoretic models, with each functional domain containing enough, but not necessarily all set-theoretically possible arrows. (See Section 3.3. for a similar proposal.) Even so, this view does not yet give us *concrete* restrictions of the set of all possible functions. To obtain the latter, one might articulate general conditions of computability on semantic meanings. For instance, should one consider only *recursive* functions? Or perhaps *continuous* functions, in the sense of Scott's 'domain semantics'? Such global notions seem rather far removed from the examples which gave the idea of procedural semantics its initial flavour.

Therefore, pursuing further case studies, like the ones in this chapter, may be a more sensible way of finding substantial forms of procedural semantics at this stage.

Regardless of the eventual outcome of this program, it should be stressed that there need not be any incompatibility here with the usual denotational views. There is always a distinction available between 'truth conditions' and 'verification conditions' (suitably understood). The latter view may complement the former, by introducing concerns of computability into contemporary semantics.

Moreover, the type of research advocated here may help in providing a handle on questions of learnability, which are just as interesting in semantics as they have been in syntax. Still, there is a long way to go here. Notably, learning seems to involve an interplay of *computation* on with *representation* of semantic objects. Of course, judicious representation of the objects for computation has been used throughout this chapter, but it has not been treated as a topic in its own right. In a more elaborate theory, one would have to study the interaction of both aspects of semantic competence. (As computer scientists know, creating economical 'data structures' is just as important as developing elegant 'algorithms'.) Final judgments of *complexity* will then depend on a balance between these two components — witness also the discussion at the beginning of Section 8.4. Even so, the present approach may have served its purpose as a plea for a computational semantics.

PART III

METHODOLOGY OF SEMANTICS

CHAPTER 9

LOGICAL SEMANTICS AS AN EMPIRICAL SCIENCE

Despite a common interest in 'logical structure', often pursued with model-theoretic tools, logical semantics and philosophy of science are nowadays disparate subjects. There have been prominent philosophers whose work spans both fields: Carnap, Reichenbach, Quine and Putnam, to mention some resounding names. But even with these, combination does not necessarily mean integration. This lack of contact is somewhat surprising, as several natural bridge topics exist, such as the area of natural laws/counterfactuals/modality or modality/tense/time. Indeed, the difference between formal semantics of natural language (sentences or texts) and that of scientific theories seems only a gradual one, in the level of 'logical aggregation'.

In this chapter, we take a look at logical semantics using some notions and insights from the philosophy of science. Excessive methodological attention of this kind can easily smother an infant science in the cradle. Therefore, our aim is to comment, not to regiment.

9.1. LOGICAL SEMANTICS AND GENERAL PHILOSOPHY OF SCIENCE

Ever since the Vienna Circle enlisted logic for its theory of science, ideas from logical semantics have had some audience among philosophers of science. Two recent examples that come to mind are the role of the Kripke—Putnam theory of rigid reference in Kuhnian debates on theory change, or Mittelstaedt's use of the constructivist dialogue semantics of Lorenzen in an operationalist analysis of physical theories.

In contrast, philosophers of science have often shied away from applying their insights to deductive sciences, such as mathematics or logic. It took Lakatos' *Proofs and Refutations* to scrutinize the mathematical activity with a Popperian eye, and logic has not even been treated in this way yet. And yet one may study, at a general level, semantical *research programs*, just as has been done in other sciences. For instance, there is nothing immoral about asking whether the flood of completenes theorems in intensional logic shows that Kripke's research program has entered upon a Lakatosian 'degenerative phase'. Such questions may be vital for the understanding and development of a subject. Here is a more historical example. Why has the *formalist* program in logic been so much more successful than the *logicist* one; even though most of its tenets were refuted in due course? One reason must surely be Popper's point that stronger falsifiable claims are usually preferable over weaker, more tenable ones. (A logicist claim that everything may be formalized in this or that logical format, is not risky enough.) Such points are still topical. A modern semantic research program like Montague Grammar runs precisely the same risk of sterility, if its practitioners do not start adding falsifiable claims soon.

But also within a particular research program, philosophy of science may help the semanticist in recognizing problems for what they are. One prominent example is the vexed issue of the *empirical evidence* of semantics. A familiar procedure in the area is the gathering of a paradigm set of typical (non-)inferences, which a formal semantics is required to explain. These data are seldom discussed systematically; and yet they exhibit all the familiar problems of empricial science. First, the selection cannot be random (although a cheerfully naive inductivism is widespread in the area). One should try to be explicit about one's *search-light* theories, to quote Poper again — as semantic 'facts' do not present themselves: *we* do. Then there are more subtle aspects of the semantic data, calling for awareness of 'theory-loading'. To begin with, one does not work with raw instances of particular (in)valid inferences, but already with half-processed versions (in the form of schemata). One consequence is that we are explaining *classes* of inferences, rather than single cases — a phenomenon not unlike the situation in natural science, where one is explaining empirical generalizations, rather than isolated facts. Also, such schemata involve a fair amount of theory-loaded interpretation in their very structure. The uncertainties of mathematical *representation* of data in the empirical sciences are reflected in the vicissitudes of 'logical form' in semantics.

Not only the representation of semantic data is theory-loaded, the very judgments of (non-)validity are affected as well. Consider the simple sentence *Alexius will not have cried.* A Dutch linguist once insisted that it had only one reading: '$\neg WH$ cry (Alexius)'. A Montague grammarian found that two readings fitted into his system, adding '$W \neg H$ cry (Alexius)'. And the students in our tense logic class 'see' three readings, viz. also '$WH \neg$ cry (Alexius)'. Evidently, the semantic observer and the semantic theorist are not independent.

It has been objected that the above view equates semantics with the *social sciences*, with their perennial revision of the evidence. And, there is much to be said for this view (if only to refute the prejudice that social science could not be *exact*). For example, semantics and some parts of social science share the study of 'internal relations' (in Wittgenstein's sense) shaping our experience of life. Moreover, the well-known phenomenon of social laws changing their own domain of application by becoming known is reflected in the observable interplay between semantic judgments of validity and norms derived from semantic theories. In this light, some contemporary discussions of the status of *logical laws* appear antediluvian.

Admittedly, earlier discussions in this book of 'intuitions' concerning denotational constraints (cf. Chapters 1, 2, 4) also had an ancient 'immutable' ring. In fact, it was claimed in Section 4.2. that, whereas judgments of inference can be fluid, statements of more general semantic principle might be more permanent. Still, even there, changes are possible — and are even facilitated by our formal scrutiny.

At a more concrete level, philosophy of science provides new perspectives upon questions concerning specific semantic theories. For instance, the debate surrounding the status of *theoretical terms* is immediately relevant. Take the common complaint that the 'accessibility' or 'similarity' relations of intensional semantics have no 'real content'. The presupposition here seems to be a widespread *realist* conviction that semantics should provide 'the correct' account of meaning. Thus, sober *instrumentalist* views of such theoretical terms would not be acceptable. Now, this may or may not be a reasonable additional demand on semantic theories. But the burden of justification lies surely with those who want to be stricter about 'accessibility' in Kripke semantics than they would be about 'gravity' in Newtonian mechanics, or 'sublimation' in psychology.

On this topic of theories and theoretical terms, more detailed formal studies are available — an outline of which follows here.

9.2. THE FORMAL STRUCTURE OF SCIENTIFIC THEORIES

The simple picture of a theory as a single formal system is inadequate for empirical science (or mathematics, for that matter). A logically congenial richer format has evolved in the Ramsey/Przełęcki/Sneed tradition.

In Ramsey (1929), an empirical theory is presented as a two-stage affair. Individual *facts* are recorded in an observational 'primary language' L_0, which also allows for the formulation of experimental generalizations or *laws* T_0. Such laws are often taken to be purely universal statements, without nestings of quantifiers. The primary language may be interpreted in L_0-structures in some standard sense: these then serve as the appropriate (representations of) pieces of reality. Thus, the primary laws define a class MOD (T_0) of empirical situations where the primary system obtains.

Next, a convenient redescription occurs in some 'secondary language' L_t. There will now be a *dictionary* translating L_0-primitives into (possibly complex) L_t-predicates — as well as a set of *axioms*. The latter are the theoretical principles of the theory, the former represent what are sometimes called 'correspondence principles'. The whole theory $T = T(L_0, L_t)$ will then consist of these two components. What is the point of this manoeuvre? Well, T affords a simpler view of the reality described in the primary language, through the introduction of suitable abstract concepts.

The status of the L_t-terms differs from those in L_0. At the primary level, we are mostly interested in knowing if T's L_0-consequences are trustworthy. Now, as a simple point of logic: *if* $T(L_0, L_t) \vdash \varphi(L_0)$, *then* $\overline{\exists T}(L_0) \vdash \varphi(L_0)$; where $\overline{\exists T}(L_0)$ is the *existential second-order closure* of T with respect to its L_t-vocabulary. The latter so-called *Ramsey Sentence* of T contains, therefore, all its empirical consequences. Thus, theoretical terms need only be interpretable in some (not: one unique) way.

Incidentally, we can also see now why T_0 need only contain universal laws. More complex sentences, such as the universal-existential form $\forall x \exists x Rxy$, may be brought into their *Skolem normal forms* $\exists f \forall x Rxf(x)$; after which the existential prefix may be absorbed into that of the Ramsey sentence.

Several connections now seem relevant between T_0 and T. First, obviously, T should *explain* T_0 in the following sense ('experimental laws are derivable'):

(1) $\quad T_0 \subseteq T \upharpoonright L_0;$

or equivalently,

(2) $\quad \text{MOD}\,(T) \upharpoonright L_0 \subseteq \text{MOD}\,(T_0).$

A less immediate condition is that T be as strict as possible in fulfilling this task:

(3) $T \restriction L_0 \subseteq T_0$;

i.e., L_0-insights from T are 'T_0-safe'.

Together, (1) and (3) express that T should be a *conservative extension* of T_0. This is one well-known formulation of 'eliminability' of theoretical terms. But, it is tempting to equate (3) with the semantic requirement that

(4) $\mathrm{MOD}\,(T_0) \subseteq \mathrm{MOD}\,(T) \restriction L_0$.

I.e., each intended empirical situation can be expanded to a model for the whole theory T by the introduction of suitable theoretical term denotations. (Or, all models for T_0 satisfy the Ramsey sentence.) (1) and (4) together may also be stated as the following principle of commutation:

(5) $\mathrm{MOD}\,(T \restriction L_0) = \mathrm{MOD}\,(T) \restriction L_0$.

This is the condition of *Ramsey Eliminability*, which has been studied in Przełęcki (1969), van Benthem (1978). In general, it is stronger than conservative extension. For, condition (3) is weaker than (4) (whose consequence it is). On *finite* structures, there is also a converse; but infinite models for T_0 may admit only of an ('elementary') *extension* which can be expanded to a model for T (in the presence of (3)). This amounts to postulating additional empirical objects in order to save the theory; in itself, a familiar procedure in science.

This approach can be generalized to the case where the operation relating T-models to L_0-structures is not simply restriction, but some other functor. E.g., if T_0 is interpretable in T through some translation τ, then τ will induce a functor from L_t-structures to τ-defined L_0-structures. (Conditions on such a functor inducing, conversely, such a syntactic translation are given in van Benthem and Pearce, 1984.)

The Ramsey account was refined in Sneed (1971) — so as to include a more sophisticated view of theoretical terms. The idea that an empirical situation becomes a model for the theory T by introducing merely *some* theoretical predicates satisfying the axioms seems too weak. Their denotation is to be determined in some stronger sense. But the obvious road, that of requiring *unique* expansions, leads to Carnapian reductionism: by Beth's Definability Theorem, all L_t-predicates would become explicitly L_0-definable on the basis of T. But then, this is not the correct analysis anyway. The point is that theoretical predicates are *constrained* by the fact that they have to behave 'regularly' across different empirical situations. E.g., mass values for the Earth should not differ according to the particular mechanical system one happens to be considering. Such constraints on simultaneous expansion of T_0-models may well enforce unique extension for the L_t-vocabulary without implying reducibility in the above sense.

Another, less formal part of Sneed's account concerns the very distinction between empirical and theoretical vocabulary. According to his well-known criterion, a term is to count as 'T-theoretical' if every attempt to determine its denotation presupposes a successful application of the theory T itself.

This concludes our survey of formal theory structure.

9.3. LOGICAL SEMANTICS AND FORMAL PHILOSOPHY OF SCIENCE

The preceding view of theories is not limited to empirical sciences. It is equally useful for deductive disciplines. Indeed, Ramsey's account had a predecessor in Hilbert's set-up for *mathematical* theories. These have a finitistic core of 'concrete' statements, with a set T_0 of *rules* of calculation. Again, these will be universal formulas. (One might even pursue an analogy between 'experiments' and 'sums'.) This core is streamlined through the introduction of an abstract theory T involving abstract notions, such as 'set' or 'infinite' totality. T generates higher proofs for elementary statements; and again, Hilbert's ideal was to establish that these were superfluous in principle (though certainly not in practice!): T was to be a conservative extension of T_0. For further connections with the above discussion, see van Benthem (1982a).

We now proceed to the inner sanctum of deductive science. The analogy to be pursued here is extremely simple. A logical semantic theory gives us a *truth definition* linking sentences in some *object language* to descriptions of semantic entities in some *metalanguage*. This is the basic Ramsey triad of 'dictionary', 'primary' and 'secondary system'.

Further details are supplied by Section 9.1. The *facts* of semantics are judgments of (in)validity of certain inferences (judgments about 'readings' being reducible to these). It has been noted that such judgments may change under theoretical pressure: but, a formal theory fixes a certain stage. (After all, most social changes do not occur very abruptly.) This formulation should be modified, as we have seen, to accommodate judgments of (in)validity of schemata. These express *laws* covering many individual facts ('types of outcome'). Note that there can be both 'positive' cases: $\varphi_1, \ldots, \varphi_n \vdash^+ \psi$, and 'negative' ones: $\varphi_1, \ldots, \varphi_n \vdash^- \psi$.

Now, the primary statements often come with a 'prima facie' semantics; for instance, in terms of 'meaning algebras'. Next, one searches for a deeper level of description, usually in some set-theoretic format. The latter will have its own 'working logic' of theoretical *axioms* T; and one check of a good choice is that one must be able to derive the laws, while avoiding all non-validities. Authors often forget about the invested theory T here, and talk as if the semantics justifies the laws for free; but this is obviously naive. For instance, a 'proof' of a predicate-logical axiom in Tarski semantics usually presupposes that very axiom, and some set theory besides.

This general point of view is also suggestive for logical semantics of natural language. There, all terms, whether theoretical or observational, occur in a single all-encompassing medium of discourse — and we are forced to think about possible *divisions* into L_0 and L_t-storeys. (Actually, this holds for science as well: there is often no natural order among scientific theories — and what is 'observational' vocabulary at one level may be 'theoretical' at another.) It is tempting to recall Sneed's criterion of theoreticity here, especially in connection with the remark about the circularity of logical justifications. Sneed makes theoreticity, in general, a context-dependent notion: the same term might be theoretical in theory T, and observational elsewhere. Could it be that the *logical constants* are distinguished, however, as those terms which are T-theoretical for every semantic theory T?

Many further speculations are inspired by the above. For instance, van Benthem (1983b) considers an alternative to current implementations of Frege's Principle of compositionality by explicit definitions for certain constructions; say, the comparative *-er*, in terms of component adjectives. Instead, it might be sufficient to have a combined adjective/comparative theory 'determining' the meaning of the comparative construction to a suitable degree. But, this goes beyond the boundaries of the present chapter. We now turn to one concrete example, viz. the semantics of propositional modal logic, to see where the preceding viewpoint takes us.

9.4. MODAL LOGIC: THEORETICAL TERMS

Possible worlds semantics for modal logic may be viewed as a theory in the above sense. For instance, let the laws T_0 consist of precisely all axioms in the so-called *minimal modal logic* K (cf. Section 5.2.). The primary semantics might use *modal algebras* in the usual sense, the secondary semantics ordinary *Kripke models*. No simple expansion is possible here from primary to secondary models. But, there exist well-known functors between these two kinds of semantic structure, as in the more general perspective mentioned in Section 9.2. (cf. van Benthem, 1985c). Here, we shall use a perspective that is somewhat closer to the simpler case.

Suppose that there is an intuitive picture already for the primary system, beyond an abstract algebra of meanings. Again, T_0 equals K; but, L_0-structures will now be models $M = \langle W, m, V \rangle$, where W is a universe of *possible worlds*, m a *modal operation* on subsets of W (cf. Section 5.2.), and V a *valuation* mapping proposition letters to the range of worlds where they hold. As it stands, m is a mere formal device — reflecting only whatever conditions T_0 chooses to impose. Here, these are the two K-axioms:

$$m(A \cup B) = m(A) \cup m(B) \qquad (\Diamond(\varphi \vee \psi) \leftrightarrow (\Diamond \varphi \vee \Diamond \psi))$$
$$m(\emptyset) = \emptyset \qquad (\Diamond \mathit{false} \leftrightarrow \mathit{false}).$$

Then, the secondary system has ordinary possible worlds frames $\langle W, R, V \rangle$ with an accessibility pattern R. Its language is the matching first-order one, having R, $=$ as well as unary predicate letters Q, one for each proposition letter q. The bridge between the two systems is given by the usual Kripke truth definition, which may be viewed as *translating* primary modal formulas into secondary ones, via the key clause

$$\Diamond q \mapsto \exists y(Rxy \wedge Qy).$$

Here, 'x' is a parameter for the world of evaluation. (E.g., the complex formula $\Diamond\Box p \rightarrow \neg\Box q$ would go to $\exists y(Rxy \wedge \forall z(Ryz \rightarrow Pz)) \rightarrow \neg\forall y(Rxy \rightarrow Qy)$.) Finally, the theory T may be taken to be just ordinary predicate logic in this case. For richer primary logics, additional laws for R will be needed — witness the S4-example in Section 5.2., which requires R to be transitive and reflexive. This fit between various modal logics and special model conditions is one of the main virtues of Kripke's enterprise.

In this perspective, what becomes of the earlier central questions concerning empirical theories? As it turns out, these return as well-known logical topics.

First, there was the *explanatory* part '$T_0 \subseteq T \upharpoonright L_0$'. This is precisely the content of the *soundness* part in the usual completeness theorems of modal logic: via the truth definition, all K-axioms follow from the possible worlds theory. Likewise, the *strictness* requirement '$T \upharpoonright L_0 \subseteq T_0$' corresponds to the *completeness* part. Thus, the ubiquitous logical *completeness theorems* have a natural motivation for the philosopher of science as well: they embody conditions of adequacy on empirical theories in semantics.

Now, there was also the stronger condition of *Ramsey eliminability* for theoretical terms. To establish the latter, every primary model is to be expanded to a model for the whole theory (here, just the truth definition) through the introduction of a suitable accessibility relation R. Here, the analysis of Section 5.2. becomes relevant. On *finite* universes, a standard definition produced such an R. So, if this restriction is accepted for semantics (cf. Section 1.2.), the theoretical term of Kripke semantics for K is strongly eliminable. In general, however, R can only be introduced in this way when m commutes with arbitrary (possibly infinite) disjunctions ('continuity'). Otherwise, obstacles may be encountered.

EXAMPLE. Let W be the set of natural numbers N, with the modal operation m defined by: $m(A) = A \cup \{0\}$ if A is infinite, $m(A) = A - \{0\}$ if A is finite. Both K-axioms are satisfied here, without infinite distributivity. (For, $(N=)m(N) \neq \bigcup_{n \in N} m(\{n\})$ $(= N - \{0\})$.) Yet, the latter would hold automatically if m were R-representable in the sense of the truth definition.

Actually, this discussion still has to be complicated somewhat. For, it may be more realistic to consider the modality m as given only on those sets of possible worlds which are the range of some proposition expressible in the primary language. In that case, an introduction of R has to use a stipulation different fom that of Section 5.2.: viz.

$$Rxy \text{ if, for all } A \text{ in the domain of } m, \text{ if } y \in A, \text{ then } x \in m(A).$$

These considerations do not exhaust the role of the usual (Henkin-type) modal completeness proofs. For, the latter may be viewed as being a set-theoretic representation method applied to primary (Lindenbaum-)algebras. As such, it is closer in spirit to the original set-up, however. (See also the account of the emergence of possible worlds in Section 7.9.)

Another aspect of the 'determination' of theoretical terms was the occurrence of *constraints* in Section 9.2. Primary K-models need not possess unique R-expansions, as is easy to see. In line with Sneed's suggestion, then, one would have to look for plausible requirements on accessibility relations across different universes of possible worlds — as was done, in fact, for *conditionals* in Section 4.7. The latter would be a good topic to analyze in the present light too, as various methodological remarks in Chapter 4 invite comparison with the present views.

Finally, all this attention to *eliminability* of theoretical terms does not imply that the latter is a very common, or even desirable feature in ordinary science. Non-conservation results are often just as important. Therefore, the so-called *incompleteness theorems* of more recent modal logic are equally relevant (cf. van Benthem, 1979). These will be the theme of the following section.

9.5. MODAL LOGIC: A RESEARCH PROGRAM THAT CANNOT FAIL?

When it arrived around 1960, possible worlds semantics constituted a clear conceptual advance. Existing syntactic modal theories turned out to be characterized in an enlightening way by means of simple elegant conditions on model classes. Since those days, this research program has become a routine enterprise, modelling a wide variety of intensional logics. With the introduction of the Henkin proof technique into the area, completeness results have become almost predictable — so much so, that various commentators have become worried. Could it be that this program, as conceived of today, is *irrefutable*, in the sense that *every* effectively axiomatized intensional logic can be provided with a complete possible worlds semantics? After all, there are so many parameters in the enterprise which can be (and have been) manipulated that almost any kind of fit seems feasible. Logicians allow themselves various types of model, containing alternative relations of different arities, special purpose conditions upon these, restrictions to 'distinguished worlds', and indeed do not shrink from modifying truth definitions to suit their purpose. With so many degrees of freedom, the positive statement that a certain logic has been 'semanticized' loses much informative content — and one would expect to learn more from a *negative* result.

Fortunately, these exist — and so, the possible worlds program is genuinely informative, having a potential for failure. To obtain this result, two ingredients are needed. One is the earlier analysis of what is meant by giving a 'possible worlds semantics' for a given logic, in terms of two degrees of freedom: choosing a truth definition, and selecting model conditions. The second ingredient is the earlier-mentioned incompleteness phenomenon, now assuming a wider significance.

In line with prevalent symbolism in the relevant literature, we shall now switch to the basic modality $\Box$ (*necessity*), with K-axioms $\Box(\varphi \wedge \psi) \leftrightarrow (\Box\varphi \wedge \Box\psi)$ and $\Box$ *true*. Giving a possible worlds semantics now involves the following. First, a correlation $\Box p \longmapsto \alpha(x, P, *)$ is to be fixed (intuitively, 'p holds at world x'), where P stands for the set of worlds validating p, and $*$ stands for any number of relations 'patterning' the universe of possible worlds in this particular semantics. In general, then, a *truth definition* τ sends modal formulas φ to formulas $\tau(\varphi)$ of the secondary semantic language, through the clauses

$$\tau(p) = Px,\ \tau(\neg\varphi) = \neg\tau(\varphi),\ \tau(\varphi \wedge \psi) = \tau(\varphi) \wedge \tau(\psi),$$
$$\tau(\Box\varphi) = \alpha(x, \tau(\varphi), *).$$

Note that this treatment of the classical connectives amounts to giving them their standard interpretation in the semantics. Variations are possible here as well (just as for $\Box$), witness current semantics for intuitionistic, relevant or quantum logic. These will not be considered here.

Right now, no special assumptions will be made about the format of α: it can be first-order or higher-order. In Section 4.8., a case was made for allowing only *first-order* clauses here — relocating further complexity to the models themselves. (Thus, for instance, a 'Beth semantics' for intuitionistic logic, which quantifies over 'barriers' of nodes across a tree of worlds, would be rephrased as a first-order semantics having models with both nodes and *branches* in the tree, together with explicit semantic conditions governing these.) This restriction will return in Section 9.6. below.

Next, the theoretical principles of T are reflected in the choice of a class K of intended secondary structures. Usually, K will be defined by some set of conditions C, again first-order or more complex. And one goal of the enterprise now becomes, as in Section 9.4., to explain primary validities:

$$\text{if } \varphi_1, \ldots, \varphi_n \vdash^+ \psi, \text{ then } C, \tau(\varphi_1), \ldots, \tau(\varphi_n) \models \tau(\psi),$$

and to explain non-validities:

$$\text{if } \varphi_1, \ldots, \varphi_n \vdash^- \psi, \text{ then } C, \tau(\varphi_1), \ldots, \tau(\varphi_n) \not\models \tau(\psi).$$

Often, this becomes even the tight fit of a completeness theorem:

$$\varphi_1, \ldots, \varphi_n \vdash^+ \psi \quad \text{iff} \quad C, \tau(\varphi_1), \ldots, \tau(\varphi_n) \models \tau(\psi).$$

In many cases, the consequence relation '$\models$' can be replaced by some notion $\vdash$ of deducibility for the secondary system. Then, giving a semantics becomes very much like the ordinary syntactic notion of providing a faithful *interpretation* for the primary theory in the secondary one. Some helpful observations emerge from this analogy. One is that ordinary interpretations often occur *relativized* to some definable subdomain of the target theory — and this, essentially, is the way to understand what happens when a class of 'distinguished worlds' is employed in certain variants of possible worlds semantics. (This possibility will not be envisaged here, as it does not affect our main results.) Another is the known *difficulty* in proving *non*-interpretability results between theories: a warning as to the complexity of the issue of 'non-semanticizability'.

The above formulation of our enterprise still borders on the trivial. For instance, for the standard τ in propositional modal logic, there is the 'general completeness' result that

$$\varphi_1, \ldots, \varphi_n \vdash_K \psi \quad \text{if} \quad \overline{\tau(\varphi_1)}, \ldots, \overline{\tau(\varphi_n)} \models \overline{\tau(\psi)};$$

where '$\vdash_K$' denotes derivability in the minimal modal logic, $\bar{\alpha}$ is the universal closure of α (for any first-order formula α), and '$\models$' is just first-order semantic consequence. But then, every modal logic L can be semanticized through the condition $C = \{\overline{\tau(\varphi)} \mid L \vdash_K \varphi\}$. This would not be a case of semantic *explanation*, however, but of mere *re-statement*. Accordingly, this cheap possibility will be blocked by the following stipulation:

> the conditions on the model class are to refer only to the predicates $*$ involved in explaining the intensional operators.

As it is, this is a little too stringent. For instance, in intuitionistic possible worlds semantics, one requires so-called 'heredity' for atomic predicates, but without allowing oneself the above liberties (cf. Section 7.6.). Such intermediate possibilities will not be considered here.

Evidently, the above perspective is a very abstract one. Usually, one does not think of a truth definition as something to be tampered with, or of a model class as a parameter to be set to some advantageous position. Nevertheless, it is this generality which gives one the 'strategic depth' to study and understand choices incorporated in the usual semantic proposals — if not to find alternatives.

For instance, the announced negative result can now be proved. Thomason (1974) contains an example of a modal logic for which no class of possible worlds frames exists such that the theorems of the logic coincide with the valid principles of that model class. But, here, this result is not yet good enough. Could not there be some 'non-standard' truth definition for modality providing a possible worlds modelling after all? In order to exclude even this, a stronger incompleteness result is needed, due to Gerson (1975). The above-mentioned logic (which contains the minimal modal logic K) is even incomplete with respect to classes of *neighbourhood frames.* The latter models generalize the usual Kripke frames $\langle W, R\rangle$ as follows. A 'neighbourhood relation' N relates worlds and sets of worlds, with the key clause in the truth definition being '$\Box\varphi$ is true at w iff $Nw[\![\varphi]\!]$'. (Notice the analogy here with the above general scheme $\alpha(x, P, *)$.) All Kripke frames can be regarded as neighbourhood frames in an obvious way, but not conversely. Thus, the Thomason logic has a stronger form of incompleteness than the usual one.

THEOREM. The Thomason logic has no complete possible worlds semantics.

Proof (Kit Fine). Suppose that this logic G had a complete semantics in the widest sense of this section. I.e., there exists a truth definition τ and a frame class K such that, for all modal formulas φ, φ is a theorem of G if and only if $F \vDash \tau(\varphi)$ for all F in K. (Here, $F \vDash \tau(\varphi)$ if $\tau(\varphi)$ holds everywhere in F, for every valuation of the atomic propositions.) Recall that τ involves a relation $\alpha(x, P, *)$ between worlds and sets of worlds, for interpreting the necessity operator. Now, for any frame F of the relevant type (i.e., that of K), let F^* be the neighbourhood frame with the same universe of possible worlds, and α for its neighbourhood relation. It is easy to see that $F \vDash \tau(\varphi)$ iff $F^* \vDash \varphi$ (in the ordinary sense), for all modal formulas φ. But then, $\{F^* \mid F \in K\}$ will be a class of neighbourhood frames whose validities match the G-theorems: quod non. □

This incompleteness example is artificial, having been made up to illustrate a phenomenon. (Its actual axiomatization would not be very illuminating for our purposes here.) Whether any *natural*, independently motivated intensional logics exhibit this behaviour remains an open question. But also, given at least this potential for failure, an *explanation* would be welcome of the descriptive success of possible worlds semantics so far.

9.6. THE RANGE OF POSSIBLE SEMANTICS

The preceding framework also suggests further questions. Perhaps the most basic of these concern the interplay of the two 'degrees of freedom' in semantic explanation. There is a certain trade-off between truth definition and model conditions, witness the following observation (van Benthem, 1984a).

EXAMPLE. The modal logic *KB*, consisting of the minimal modal logic *K* plus the *Brouwer* axiom *B*, can be modelled in the following two ways:

truth definition: $\forall y(Rxy \rightarrow Py)$ / model condition: R is symmetric,
truth definition: $\forall y((Rxy \vee Ryx) \rightarrow Py)$ / model condition: none.

Thus, a question arises as to the division of labour. Which part of the burden of explanation is to be borne by the truth definition?

As was argued in earlier sections, there is a case for keeping the latter *first-order*. In addition, in modal logic, one reasonable proposal is this. The truth defintion by itself should account for the validity of the *minimal* modal logic, whereas further modal principles should be taken care of by appropriate model conditions. Of course, the original Kripke truth definition does just this. Our question then becomes: which *range* of truth definitions would have worked equally well?

Actually, we shall answer only one half of this question. To validate *at least K* (in a sense to be explained below), a truth definition must belong to a small 'Kripke family' of syntactic forms. To validate also *at most K* may well be a property possessed only by the original Kripke clause, and some marginal variants. For a start, recall the key K-axiom: $(\Box p \wedge \Box q) \leftrightarrow \Box(p \wedge q)$. In the light of Sections 1.4., 2.6., there are two well-known model-theoretic conditions here on modal operations $\Box$, as described by their truth schema $\alpha(x, P, *)$:

(1) α is *upward monotone* in P: if $\alpha(P)$ and $\forall x(Px \rightarrow P'x)$, then $\alpha(P')$;

(2) α is *conjunctive* in P: if $\alpha(P_1)$ and $\alpha(P_2)$, then $\alpha(P_1 \cap P_2)$.

(There are some innocent abuses of notation here.) What we are after now is a *preservation result* (cf. Section 2.5.). Which syntactic forms for α are enforced (modulo logical equivalence) by these two constraints? The obvious conjecture is this. The ordinary truth definition sets $\alpha(P) = \forall y(Rxy \rightarrow Py)$. Its only liberalization guaranteeing (1), (2) would seem to be

$$\alpha(P) = \forall y(\rho(x, y) \rightarrow Py), \quad \text{with } \rho \text{ any } P\text{-free formula.}$$

But, there is a subtle difficulty here with *infinite* models. E.g., the following sentence is conjunctive and upward monotone in P, without being reducible to the above form:

> 'R is a strict linear order, and $\forall y(Rxy \rightarrow \exists z(Ryz \wedge \forall u(Rzu \rightarrow Pu)))$'.

But then, this 'cofinal eventual truth' on linear orders lacks a stronger form of conjunction (compare the *continuity* of Section 5.2.):

> α is *conjunctive*$^+$ in P: if $\alpha(P_i)$ (all i in I), then α holds for the predicate $\bigcap_{i \in I} P_i$.

This new condition holds for the above scheme, while failing for the cofinality example. (Consider the structure $\langle N, < \rangle$ with $P_i = \{n \geqslant i \mid n \text{ in } N\}$ $(i = 1, 2, \ldots)$.) Although conjunctivity$^+$ was motivated by larger conjunctions than binary ones, it also has an extreme case, when I is *empty*. Then, the antecedent becomes vacuously true, while the consequent states that α must hold for the *universal* predicate (whose extension is the whole domain). Thus, the remaining K-axiom $\Box$*true* is subsumed as well.

THEOREM. A first-order formula $\alpha(P)$ is upward monotone and conjunctive$^+$ in P if and only if it is equivalent to one of the following form:

$$\forall x(\rho(x) \rightarrow Px), \quad \text{where } \rho \text{ is } P\text{-free.}$$

Proof. The 'if'-direction is immediate, with the 'only if'-direction containing the main result. The argument follows a standard model-theoretic pattern, with one new twist to it.

Consider $C(\varphi) = \{\beta \mid \varphi$ implies β, and β is of the form $\forall x(\rho(x) \rightarrow Px)$ as above$\}$.

CLAIM. $C(\varphi) \vDash \varphi$.

Once this has been shown, the required conclusion follows by compactness: the conjunction of some finite subset of $C(\varphi)$ will be equivalent to φ. Notice here that, e.g., $\forall x(\rho_1(x) \rightarrow Px) \wedge \forall x(\rho_2(x) \rightarrow Px)$ will be reducible to $\forall x((\rho_1 \vee \rho_2)(x) \rightarrow Px)$. Notice also that the case Pc, for individual constants c, is included, through the transformation $\forall x(x = c \rightarrow Px)$.

Now, let $\mathfrak{A}$ be any model for $C(\varphi)$. We shall find a model $\mathfrak{b}$ for φ, in which $\mathfrak{A}$ lies L-elementarily embedded, in such a special way that the truth of φ in $\mathfrak{b}$ may be transferred to $\mathfrak{A}$.

As a first step, consider the expanded model $(\mathfrak{A}, A)$ for the language $L \cup \{P\}$ (L is the P-free part), enriched with new individual constants $\bar{a}$ for each $a \in A$. Choose new unary predicates P_a for each object a in the complement of $P^{\mathfrak{A}}$.

LEMMA. $\mathrm{Th}_L(\mathfrak{A}, A) \cup \{\varphi(P_a), \neg P_a\bar{a} \mid a \in A - P^{\mathfrak{A}}\}$ is consistent.

Proof. Suppose it is not. Then, for some $a_1, \ldots, a_{n+m}$, the following set of sentences must be inconsistent:

$$\alpha(\bar{a}_1, \ldots, \bar{a}_n, \bar{a}_{n+1}, \ldots, \bar{a}_{n+m}), \quad \varphi(P_{a_1}), \ldots, \varphi(P_{a_n}),$$
$$\neg P_{a_1}\bar{a}_1, \ldots, \neg P_{a_n}\bar{a}_n.$$

In other words, by some suitable rearrangements, quantifying away irrelevant constants in α:

$$\alpha(x_1, \ldots, x_n), \varphi(P_1), \ldots, \varphi(P_n) \vDash P_1x_1 \vee \cdots \vee P_nx_n.$$

If $n = 1$, then this would mean that

$$\varphi(P) \vDash \forall x_1(\alpha(x_1) \rightarrow Px_1);$$

and hence the latter L-sentence holds in $\mathfrak{A}$ (as it belongs to $C(\varphi)$): contradicting the truth of $\alpha(\bar{a}_1)$, $\neg P\bar{a}_1$ in the latter model.

If $n > 1$, then a more extended argument is needed, involving repeated use of the Interpolation Lemma:

From the above implication, one obtains

$$\begin{aligned}&\varphi(P_1), \ldots, \varphi(P_{n-1}), \alpha(x_1, \ldots, x_n), \neg P_1x_1, \ldots,\\ &\neg P_{n-1}x_{n-1} \quad (\textstyle\sum, \text{for short})\\ &\quad \models \varphi(P_n) \to P_nx_n \ (\gamma).\end{aligned}$$

By interpolation, there exists some ψ_n not containing $P_1, \ldots, P_n$ or P such that

$$\textstyle\sum \models \psi_n \quad \text{and} \quad \psi_n \models \gamma.$$

So, we have $\varphi(P_n) \models \forall x_n(\psi(x_n) \to P_nx_n)$, or

$$\varphi(P) \models \forall x(\psi_n(x) \to Px).$$

Moreover, there remains

$$\varphi(P_1), \ldots, \varphi(P_{n-1}), \alpha \wedge \neg\psi_n \models P_1x_1 \vee \cdots \vee P_{n-1}x_{n-1};$$

to which the same argument may be applied, yielding

$$\varphi(P_{n-1}) \models \forall x_{n-1}(\psi_{n-1}(x_{n-1}) \to P_{n-1}x_{n-1}); \text{etc.}$$

In the end, one obtains the following list:

$$\begin{aligned}&\varphi(P) \models \forall x(\psi_n(x) \to Px), \ldots, \varphi(P) \models \forall x(\psi_2(x) \to Px), \text{ and}\\ &\varphi(P), \alpha \wedge \neg\psi_n \wedge \cdots \wedge \neg\psi_2 \models Px, \text{ i.e.,}\\ &\varphi(P) \models \forall x((\alpha \wedge \neg\psi_n \wedge \cdots \wedge \neg\psi_2) \to Px)\end{aligned}$$

(with a suitable existential quantification performed on the antecedent).

So, all these consequences must be true in $\mathfrak{A}$ (being in $C(\varphi)$). But, this contradicts the following facts about the original $a_1, \ldots, a_n$:

$$\alpha(\bar{a}_1, \ldots, \bar{a}_n), \neg P\bar{a}_2, \ldots, \neg P\bar{a}_n \text{ and yet } \neg P\bar{a}_1.$$

Remark. If the complement of $P^{\mathfrak{A}}$ is empty, the next stage of our argument does not get off the ground. But, in that case, $P^{\mathfrak{A}}$ is the universal predicate on $\mathfrak{A}$, for which φ holds automatically. But then, φ holds in $\mathfrak{A}$ — and we are done.

From the lemma, it follows that the mentioned set has a model $\mathfrak{b}$, together with an L-elementary embedding h from A to B, with the additional property that

$$\text{if not-} P\bar{a} \text{ in } \mathfrak{A}, \text{ then } \neg P_a\bar{a}, \varphi(P_a) \text{ hold in } \mathfrak{b}.$$

Next, consider $\mathfrak{b}$ as an L, P_a-model, for any a. Changing $\mathfrak{b}$ to $\mathfrak{b}^+$ (and hence also $h(\mathfrak{A})$ to $h(\mathfrak{A})^+$ by enlarging the P_a-denotation to $P_a^+ =_{\text{def}} B - \{h(a)\}$ leaves the following two facts undisturbed: — $\mathfrak{b}^+ \models \varphi(P_a)$ (because of *upward monotonicity*), — $h(\mathfrak{A})^+$ is an L-elementary submodel of $\mathfrak{b}^+$.

But then, as P_a^+ is atomically L-definable with a parameter in $h(\mathfrak{A})$

$$P_a^+ b \quad \text{iff} \quad b \neq h(a),$$

it follows that even

$$h(\mathfrak{A})^+ \text{ is an } L, P_a\text{-elementary submodel of } \mathfrak{b}^+.$$

Therefore, in particular, $\varphi(P_a)$ holds in $h(\mathfrak{A})^+$ (for all $a \in A - P^{\mathfrak{A}}$).

Now, by *conjunction*$^+$, these facts together imply that

$$(h(\mathfrak{A}), \bigcap_a (P_a^+ \cap h(\mathfrak{A}))) \models \varphi;$$

i.e., φ holds in the model consisting of the L-reduct of $h(\mathfrak{A})^+$, with a P-interpretation equal to $h[P^{\mathfrak{A}}]$. But, between this model and the original model $\mathfrak{A}$, h is an L, P-isomorphism (in addition to its being an L-isomorphism). Hence, φ holds in $\mathfrak{A}$ as well. □

Thus, we have arrived at one strong constraint upon possible modal truth definitions. Others might be added, this particular one could be changed — but the present pattern of argument will presumably remain applicable.

To conclude, here is another question of syntactic fine-structure, this time concerning the model conditions rather than the truth definition.

Suppose that a logic has been described using some truth definition τ and model conditions C. Now, one wants to change to some other truth definition τ': can the corresponding C' be determined? More in particular, for example, if τ belongs to the above family of syntactic forms, and C is a set of *first-order* model conditions, can this description always be 'normalized' to the ordinary Kripke clause (τ'), for some suitable set of first-order conditions C?

The obvious strategy here seems to be this. For each frame F in the original class, consider the Kripke frame F^* with the same universe of possible worlds, and $\lambda xy \cdot \rho(x, y)$ for its alternative relation. Then $F \models \tau(\varphi)$ iff $F^* \models \varphi$ in the standard sense, for all modal formulas φ — and hence the same modal logic is modelled by the $*$-image of the original model class.

But, the new class obtained is *projective* ('there exist predicates in the original sense such that . . .'), and hence it need not be first-order, even if the original one was.

EXAMPLE. Let C be the condition that R is a strict linear order, with τ of the form $\forall y(\rho(x, y) \rightarrow Py)$, where $\rho(x, y)$ expresses 'exactly one world lies between x and y, in that order'. Now, consider the frame F consisting of the integers in their usual order. F^* as defined above will consist of two disjoint copies of the integers, with the order of 'immediate succession'. Then, one such copy will obviously be elementarily equivalent to F^*. But, an easy argument shows that such a single copy cannot be a $*$-image of any strict linear order. Thus, our projective class is not first-order definable.

This counter-example is not conclusive — as a suitably enlarged elementary class C' might still do the job.

The area abounds in further questions.

CHAPTER 10

THE LOGIC OF SEMANTICS

Now that semantical investigations of natural language have established themselves as a recognized scientific activity, research material is accumulating which invites reflection. For, much clarity is still to be achieved as to the nature of semantic theory, explanation or evidence. One way to proceed here is by general philosophizing, much as in earlier periods. But a more concrete approach is also available at the present stage. There is room for a foundational component of semantics, consisting of a logical study of semantic theory. The grand aim of such a study was stated already; its specific contributions will concern such recurrent issues as the role of 'semantic constraints' on grammar, the nature of compositionality, the adequacy of first-order logical forms, or the content of 'semantic universals'. Many of these themes concern the relation between syntax and its interpretation; and indeed one may think of a mathematical semantics as an extension of traditional 'mathematical linguistics', which is essentially a science of syntax.

In this chapter, various developments in contemporary formal semantics will be brought together in one Grand Pattern of ever richer conceptions of semantic theory. At each stage, relevant logical questions are asked, and sometimes even answered. Not all relevant studies could be fitted into the present framework, but the presentation at least provides a coherent perspective upon much of Montague Grammar, philosophical logic, the recent generalized quantifier movement, as well as the incipient research into discourse representations.

When one asks just what a formal semantics for some part of language achieves, the answer turns out surprisingly difficult to state. Of course, at an initial stage, there was still so much elation at the discovery of flexible formal modes of description that such questions seemed irrelevant. And this may still be the dominant attitude: formal frameworks are still appearing on the market with so many conceptual parameters that they can model almost anything, and this is usually claimed to be a virtue. A comparison may be made here with the early days of *Logicism*, with the Montagovian classics as the counterpart of Principia Mathematica. There are dangers in this situation: logicism as the claim that mathematical discourse is formalizable is virtually *irrefutable*, and hence eventually sterile. (This point was raised already in Section 9.1.) The same fate looms large over contemporary formal philosophy of science, and it could also happen to (some would say it has already happened to) Montague Grammar. A fate even worse than death is conceivable too. People might not notice the problem, and become content with description instead of explanation, as in the time of Aristoteleanism. The analogy between contemporary semantics and medieval scholasticism, often drawn by critics, may not be completely off the mark!

But there is hope. After all, modern logic escaped its doom by developing a much richer range of questions about its subject matter, and making falsifiable claims about these — witness the case of Hilbert's Program in the foundations of mathematics, refuted by Gödel's Theorems. And indeed, similar developments may be discerned in contemporary semantics, witness the ascending ladder of goals for a semantic theory to be presented in this chapter, which has already given rise to heroic and instructive mistakes, i.e., to scientific progress.

10.1. SEMANTICS AS FAITHFUL DESCRIPTION

Semantic analysis should be true to the actual structure of sentences. On this much, most practitioners would seem to agree. This much is also about *all* that would be required by that pioneering school of Montague Grammar. For, consider its basic scheme for a semantic theory, as displayed in Figure 31.

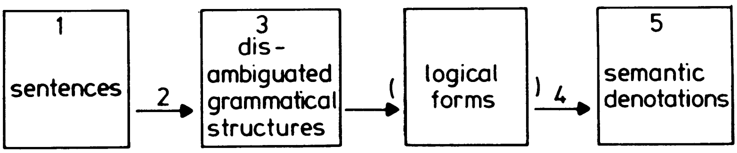

Fig. 31.

Even in the condensed version with direct interpretation of grammatical structures, the enterprise has five *degrees of freedom*, as we can choose our prima facie grammar (1), our grammatical analytical tools (3), the parsing relation (2), the semantic meanings (5) and the truth definition (4). Infallible success seems the reward of industry here.

In reality, things are not this easy, as *constraints* are operative. For instance, not just any semantic picture will be acceptable in 5: some will even demand first-order models. Nevertheless, much of this is implicit ideology. The only explicit constraint formulated by Montague himself concerned 4: semantic interpretation was to obey Frege's Principle of *Compositionality*. Now, this principle has been investigated thoroughly in an algebraic setting in Janssen (1983). The general outcome may be stated roughly as 'anything goes' — even though adherence to the principle often makes for elegance and uniformity of presentation.

The reason why can be stated in very simple terms. Suppose that some algebra $\langle A, \mathcal{O}\rangle$ with syntactic operations $\mathcal{O}$ has been chosen, representing disambiguated readings of linguistic items, and that, assuming the severest possible semantic constraints, some meaning algebra $\langle B, \mathcal{M}\rangle$ is prescribed in advance. Our task is then to see if there exists a *homomorphism* from the former into the latter algebra. (This is the algebraic version of Compositionality.) In an extreme case, even the connections between syntactic operations and semantic operations in $\mathcal{M}$ are fixed, and we have at least a respectable problem (which can be treated with model-theoretic methods). But usually, such a connection is not prescribed: indeed, any polynomially *definable* operation on B may be assigned, in principle, to any operation in $\mathcal{O}$. The extreme versatility of this device shows in the conjuring tricks performed using lambda-abstraction by many authors, in cases where the meaning algebra comes from some theory of types. Finally, even this presentation is too restricted; for actually, the syntactic algebra $\langle A, \mathcal{O}\rangle$ is *free* (being freely generated by the basic lexical items). What this algebraic assertion amounts to is this. The construction of $\langle A, \mathcal{O}\rangle$ is such that, given any connection of operations in $\mathcal{O}$ with semantic operations of the same number of arguments, an arbitrary map from basic lexical items to suitable semantic entities will be extendable to a homomorphism as required. Thus, by itself, *compositionality provides no significant constraint upon semantic theory.*

This conclusion does not exclude that having a compositional semantics as such could be useful. For instance, setting up things compositionally has proven a powerful technique in computer science for obtaining perspicuous pictures of manipulation of data, as well as a convenient format for proving 'correctness assertions' about program execution.

If more 'essential tension' is to be introduced, additional constraints will have to be formulated, drawing upon our fund of intuitive boundary conditions. Various suggestions to this effect may be found scattered in the literature. For instance, in Landman and Moerdijk (1981), the level 3 of grammatical analysis is restricted to that of categorial grammars with only syntactic rules of concatenation and substitution, while the interpretation 4 should be a homomorphism under the fixed convention that concatenation corresponds to functional application. And even further conditions might be proposed, both on the linguistic side (1, 2, 3) and on the logical side (4, 5). Cresswell (1973) contains the idea to let 2 consist of merely flattening 3-structures, removing auxiliary symbols; while there is a recurrent folklore idea to have 1 somehow consist of some context-free grammar. Conditions on 5 might assume the form of taking meaning algebras only from first-order predicate logic, or some simple type theory. Thus, there arises the conception of a *semantic hierarchy*, in analogy with the well-known grammatical hierarchy of mathematical syntax; starting with simple grammars, austere logics and tight connections between the two, becoming more liberal in its upper regions.

The study of this semantic hierarchy would be one of the first tasks for a systematic mathematical semantics. The earlier-mentioned Janssen (1983) gives a fair impression of universal algebraic techniques that may be useful here. For a case study, see the intermezzo below.

Some semanticists find the present perspective thoroughly uncongenial. The problem, for them, is not to delimit a range of acceptable solutions, but rather to find *the* correct one. Thus, it is reported by various observers that our semantic intuition can tell us what are 'the correct' truth conditions for a given syntactic operation. Given the amount of theoretical reconstruction in formal semantics, such acts of faith are rather gratuitous. There obviously is something like 'recognizing a good theory when we see one' in science, but the phenomenon needs careful scrutiny. For instance, one might apply criteria of *simplicity* to proposed truth definitions, just as in the philosophy of science. But that again will find its place as searching for the lowest possible levels in the above hierarchy.

Intermezzo: Homomorphisms and the Proper Form of Definitions

A sample question at the above level of generality is the following. Suppose that a language has some set of syntactic operations $\vec{f}$, with some constraints on these given by a set of equations T_0. Now, one wants to enrich the language with a new operation g, extending T_0 to T, either by explicit definition or via recursion equations, or yet otherwise. Furthermore, the original language was already interpreted compositionally: i.e., its syntactic algebra $\mathfrak{A}$ was mapped homomorphically into all algebras of the suitable similarity type. Now, here is our wish: addition of g should leave these already established interpretations 'undisturbed'. Which format of introduction for g is to be prescribed, in order to guarantee this?

First of all, the metaphor is to be translated into mathematical terms. Here is a first attempt:

> for all suitable algebras $\mathfrak{b}$, each $\vec{f}$-homomorphism from $\mathfrak{A}$ onto $\mathfrak{b}$ automatically remains an $\vec{f}$, g-homomorphism from $(\mathfrak{A}, g)$ onto $(\mathfrak{b}, g')$ *for some unique* operation g' on B. (HE)

In general, HE is difficult to handle, being 'local' in the algebra $\mathfrak{A}$. Therefore, here is a more global second attempt, in line with prevalent practice in this area:

> each $\vec{f}$-homomorphism from any $\mathfrak{A}$ in MOD (T) onto $\mathfrak{b}$ in MOD (T) is automatically an $\vec{f}$, g-homomorphism. (HE^+)

This formulation is very close in spirit to one in Ehrig *et al.* (1978), on the specification of 'abstract data types' in computer science.

The second version is more amenable to model-theoretic analysis, as was shown in van Benthem (1980). Notably, the special case of HE^+ with $\mathfrak{b} = \mathfrak{A}$ and the identity homomorphism running between them, amounts to *implicit definability* of g in T on the basis of $\vec{f}$, in the sense of Beth's Definability Theorem (cf. Chang and Keisler, 1973). The latter result then states that g must also be *explicitly definable* in T in terms of $\vec{f}$. I.e., there exists some first-order formula $\varphi(\vec{x}, y)$ in $\vec{f}$ such that $\forall \vec{x} y(g\vec{x} = y \leftrightarrow \varphi(\vec{x}, y))$ is derivable in T. But, such first-order definitions are too complicated for our equational theories — and besides, they do not guarantee HE^+ in their turn.

A form of definition which does have the latter property is *polynomial definability*: T derives $\forall \vec{x}(g\vec{x} = t(\vec{x}))$, for some $\vec{f}$-term t (ED). But, the latter again is too strong for HE^+: a counter-example was found by Kees Doets (see the above-mentioned paper).

One promising weakening of polynomial definability is the following: there exists some finite set ϵ of equations in $\vec{f}$ and the variables $\vec{x}$, y

such that (i) $g\vec{x} = y$ is derivable from $T + \epsilon$, and (ii) ϵ itself is derivable from T, $g\vec{x} = y$ (ED^+). The latter form of definability does imply HE^+; but it is still too strong for the purpose. The real solution eventually emerges from a standard model-theoretic analysis:

THEOREM. The homomorphism extension condition HE^+ holds if and only if T proves the definition $(\forall \vec{x} y(g\vec{x} = y \leftrightarrow \theta(\vec{x}, y))$ for some *positive* first-order formula φ in $\vec{f}$.

Nevertheless, there is now a certain interest to the converse direction too. Which semantic conditions match the various notions of definability encountered along the way? For equational theories, here are the answers:

THEOREM. ED holds if and only if $\vec{f}$-subalgebras of models of T are always closed under the operation g.

THEOREM. ED^+ holds if and only if each $\vec{f}$-homomorphism from an *$\vec{f}$-subalgebra* of a model for T into a model for T is already an $\vec{f}$, g-homomorphism.

Here, for tuples $\vec{a}$ in that subalgebra, *if* $g(\vec{a})$ belongs to the subalgebra as well, *then* $F(g(\vec{a})) = g(\overrightarrow{F(a)})$ for the $\vec{f}$-homomorphism F.

All these results employ only preservation of the relevant theories under the formation of *subalgebras* and *direct products*; and hence they also hold for *universal Horn sentences*, the standard format in the theory of abstract data types — which has obvious connections with the present area.

Finally, on this topic of extension by definitions, the connection between the two theories (or 'specifications') T_0 and T itself raises some further questions. In line with the discussion of Section 9.2., one might require that T be a *conservative extension* of T_0, either in the weaker syntactic sense, or the stronger semantic sense of 'Ramsey Eliminability'.

Finally, we return to the original *local* extension principle HE. We can make some progress here by noting that frameworks such as Montague's typically employ *free* syntactic algebras, which are 'initial' in the class of models for their equational theory:

THEOREM. If $\mathfrak{A} = \langle A, \vec{f}\rangle$ is freely generated by some *infinite* $X \subseteq A$ such that each $\vec{f}$-homomorphism from $\mathfrak{A}$ into a model for the equational theory of $(\mathfrak{A}, g)$ is already an $\vec{f}$, g-homomorphism, then g is $\vec{f}$-polynomially definable in $(\mathfrak{A}, g)$.

The converse of this proposition is immediate. In case the set of generators is finite, however, counter-examples exist to the theorem; as was pointed out by W. Peremans.

This concludes our discussion of 'successive definition' — illustrating the flavour of universal algebraic methods in this area.

10.2. SEMANTICS EXPLAINS INFERENCE

The previous account cannot contain all there is to semantics. After all, natural language *itself* does pretty well as a faithful description of natural language. If we are to formalize, this had better be for some purpose, if we are to avoid sterile cartography. Indeed, there is no reason to deviate from common scientific practice here, which is to formalize as little as necessary for any given purpose.

One goal of semantic description stressed by various authors in the area is the explanation of (in)valid inference. Semantic intuition provides us with a number of observations on validity or fallaciousness of certain proposed arguments, and these are to be accounted for in the formal semantics. As Barbara Partee once remarked: "inferences are becoming part of the linguistic data". Here, the viewpoint of Section 9.3. becomes appropriate. There is a prima facie language L_0, with a partial relation $\Rightarrow^+$ of valid inference, as well as a partial relation $\Rightarrow^-$ of invalid inference, representing the above intuitive evidence. Then, one introduces a theoretical language L_t, with an independent notion of derivability $\vdash$; and an *explanation* is to be found in the form of a map τ taking L_0-sentences to L_t-sentences in such a way that

$$A \Rightarrow^+ B \text{ only if } \tau(A) \vdash \tau(B),$$
$$A \Rightarrow^- B \text{ only if } \tau(A) \nvdash \tau(B).$$

How much of a constraint is all this? Again, the answer is 'surprisingly little'. Here is a sample result.

THEOREM. For countable L_0, $\langle L_0, \Rightarrow^+, \Rightarrow^-\rangle$ is embeddable into predicate logic in the above sense if and only if $\Rightarrow^-$ is disjoint from the reflexive transitive closure of $\Rightarrow^+$.

Proof. It suffices to show that any countable reflexive transitive structure can be isomorphically embedded into the relational structure consisting of all predicate-logical formulas with the derivability relation. But, this may be seen as follows. Any countable reflexive transitive relation may be represented as a countable partial order of 'indifference classes'. Any such partial order may, in its turn, be embedded in a countable Boolean algebra. And finally, any countable Boolean algebra can be embedded into the countably infinite atomless one. But, that is precisely the above structure of predicate logic (with an infinite vocabulary). Moreover, each formula has an infinite indifference class of logical equivalents, and hence there is enough room to accommodate our original structure. □

We will not go into the combinatorially more complex case where $\Rightarrow^+$, $\Rightarrow^-$ are relations between finite sets of premises and a conclusion. (For more interesting results about the above predicate-logical structure, see Mason, 1984.)

Simple though it may be, the above theorem illustrates two points. One is that vexed methodological quarrels, such as the one about the 'adequacy of predicate logic', can be resolved into various precise problems, to which answers may be found (not necesarily the same one in all cases). The other is that we are forced methodologically to become clearer about further features of the semantic enterprise, if we are to escape the conclusion.

One obvious move in the latter spirit is to restrict the proposed connection τ to, say, *effective* or recursive operations. More in line with the preceding, one may impose the condition of *compositionality*. Thus, τ is to do the job of explaining inference, while respecting syntactic structure. Formally, we are now looking for compositional embeddings from $\langle A, \mathcal{O}, \Rightarrow^+, \Rightarrow^- \rangle$ into $\langle B, \mathcal{M}, \vdash, \nvdash \rangle$. But, this is precisely the perspective of 'philosophical logic', with its search for completeness theorems via suitable truth definitions. And that again, has been investigated thoroughly in the preceding chapter, with the outcome that there are indeed constraints here: not all given intuitions about inference can be 'semanticized', subject to the present constraints.

10.3. SEMANTIC UNIVERSALS

In recent years, richer conceptions of the aims of semantics may be discerned in various publications. For instance, our study of generalized quantifiers in this book has focused upon special categories of expression (noun phrases, determiners, etc.), discovering a whole fine-structure of Montague Grammar, so to speak. Most conspicuously, this has led to the formulation of general conjectures about the occurrence of, or connections between certain types of determiner in all human languages. This development of 'semantic universals' not only enriches the Montagovian fund of semantic themes, it also provides a promising rapprochement between earlier 'fragmentary' approaches in formal semantics and more 'global' linguistic habits of description — aiming at finding higher regularities across one or several human languages.

Now, by itself, the generalized quantifier framework is just a medium of precise description. Contrary to some avowed opinions, it does not carry any explanatory power as such. The discovery that the denotation of a noun phrase such as *nobody's fool* exhibits such a familiar mathematical structure as an 'ideal of sets' by itself means nothing. (Compare Piaget's often repeated observation that the logical operations of identity, negation and converse, mastered by fourteen-year olds, give rise to 'Klein's 4-Group'. Nobody has ever been able to fit this into some significant theory.) But, the framework invites further enquiry, and to that topic we now turn.

As we have seen in Sections 1.4., 2.6., generalized quantifiers have been used as a vehicle for proposing semantic universals concerning observed 'systematic gaps' in natural language. For instance, Zwarts has conjectured that *no human language has determiners denoting strict partial orders.* Barwise and Cooper also proposed more sophisticated dependencies, such as *all persistent determiners in human languages are monotone.*

When examples such as these first reached traditional semantic circles (still dwelling in the realm of Sections 10.1., 10.2. above), the immediate reaction was scepsis (or hilarity). There was the unfamiliar terminology ('universal' carries nebulous philosophical connotations), there were also unhappy experiences in the past with such grand aims (see below) — but there was also sheer inability to recognize this activity for what it is: an attempt to formulate semantic regularities or laws, just as in any science. The above examples may be viewed as empirical laws, say like Kepler's — or better (in view of their humble nature), Bode's Law in astronomy.

The relevant criticism, then, is not that people should dare to put forward such proposals, but rather how we are to evaluate them. What is still lacking at present is that other component which makes science out of systematic observation, viz. some kind of explanatory background theory. (After all, Kepler's laws were only explained by Newton's mechanics.) Awareness of this point shows in several publications in the area. Thus, Barwise and Cooper provide some background speculation about their proposed universals in terms of *psychology*, that great consolation of so many a semanticist in need of foreign support. Keenan and Stavi refer to broad conceptions of *efficiency* of natural language, another pious favourite of many students of semantics. Perhaps least substantially, van Benthem has drawn upon 'conceptual intuitions'.

The claim of the present section is that these vague speculations can be investigated systematically. A logical foundational study can contribute its usual insights here: providing a conceptual apparatus, as well as 'limitative results'. Especially, it will tell us how much of the proposed universals is due to logic (and mathematics); i.e., to pure deduction already; and (thus) where the border line lies with real empirical content. This function is particularly important in view of earlier experiences with 'linguistic universals', which usually turned out to be either tautologous, or incorrect.

Of course, one might also give a more optimistic interpretation of this phenomenon. Perhaps, the realm of linguistics realizes the old metaphysical Principle of *Plenitude*: every logical possibility is realized in some natural language, somehow, sometime, somewhere.

How far pure *logic* takes us has been amply demonstrated in Chapters 1 and 2. Notably, the above Zwarts universal turned out to be true as a matter of logic, whereas others did not. No more need be said on this topic here.

If some *mathematics* is brought in (actually, some elementary arithmetic), several pleasing insights may be added. For instance, mere attention to orders of magnitude turned out to be relevant in the work of Keenan and Stavi (cf. Section 1.3.), treating the question how natural language manages to attain maximal expressive power with a minimal basic vocabulary. Here, we consider another example. Some of the 'psychological' speculations in Barwise and Cooper (1981) concerning 'minimal verification' of determiner statements admit of mathematical statement and proof.

Consider a set A with n elements. Some minimal numbers of individuals in A whose B-behaviour one needs to know, in order to confirm or refute the statement QAB for quantifiers Q, are given in the following table:

Determiner	Confirm	Refute	Total
all	n	1	$n+1$
some	1	n	$n+1$
most	$[\frac{1}{2}n]^{+}$	$[\frac{1}{2}n]_{+}$	$n+1$
all but at most one	$n-1$	2	$n+1$
precisely one	n	2	$n+2$

Here, $[m]^{+}$ means: the first integer greater than m, and $[m]_{+}$ the first integer no smaller than m.

Further examples confirm the suspicion that $n + 1$ is a lower bound of complexity. Let us tighten up some definitions. Call x_n *confirmation-minimal* with respect to n, if x_n is minimal with respect to the following property: 'there exists a couple (k_1, k_2) in Q (recall the representation of quantifiers in the Tree of Numbers; cf. Section 2.2.) with $k_1 + k_2 = x_n$ which already decides Q at level n, in the sense that *any* (n_1, n_2) with $n_1 + n_2 = n$ and $k_1 \leqslant n_1$, $k_2 \leqslant n_2$ must belong to Q'. 'Refutation-minimality' is defined analogously (y_n).

THEOREM. For each n, $x_n + y_n \geqslant n + 1$.

Proof. The decomposition (k_1, k_2) of x_n which belongs to Q induces $n - x_n + 1$ adjacent couples (n_1, n_2) belonging to Q, and y_n likewise induces $n - y_n + 1$ adjacent couples outside of Q. In all, there are $n+1$ couples (n_1, n_2) at level n, and hence $n - x_n + 1 + n - y_n + 1 \geqslant n + 1$; i.e., $n + 1 \leqslant x_n + y_n$. □

Note also that, if such minimal x_n, y_n exist with $x_n + y_n = n + 1$, then their induced inside/outside sequences occupy the complete n-row for the quantifier in the number tree.

Now, Barwise and Cooper suggest that basic quantifiers will be of minimal count complexity. One, rather liberal way of phrasing this idea is to require the existence of minimal confirmation/refutation numbers for each cardinality $n > 1$. Using the tree of numbers, all possibilities of this sort may be surveyed geometrically:

THEOREM. The quantifiers of minimal count complexity are precisely those for which each row consists of (at most) two adjacent inside/outside segments — in such a way that, going down the tree, the extreme 'inside' position can only move one step toward the left or the right at a time.

This geometric pattern can also be described differently. Moving along any of the three main directions in the Tree (/, –, \) will produce at most one change of truth value. This amounts to *Strong Continuity* for quantifiers QAB, with respect to varying A (cf. Section 1.4.), but also $A - B$, $A \cap B$. (A proof may be found in van Benthem, 1985a, using Variety.) The latter condition of continuity was already used in Section 8.3. as one aspect of 'computability' for quantifier denotations.

An interesting consequence establishes a connection with another basic condition on determiners:

COROLLARY. All quantifiers of minimal count complexity are (either upward or downward) monotone.

This result is obvious from the above description of patterns.

There is an amazing variety of minimal count quantifiers, as the above +/– border may have $2^{\aleph_0}$ distinct shifting patterns. Thus, there arises the need for some fine-structure. One way to get this is by imposing *regularity* of the shifting pattern, in line with the Uniformity intuitions of Section 2.4. At a first level, the top shift pattern repeats itself — which leaves only four non-degenerate possibilities: *no, all, some, not all*; being the Square of Opposition. At a second level, the pattern of the two top shifts repeats itself — inducing four additional cases: *most, not most, least, not least*; being the best-behaved higher-order quantifiers. At higher levels, complexity increases rapidly. But here, we shall stop, recalling Emmon Bach's dictum that, in linguistics, the only significant numbers are 1, 2, *n*.

Finally, let us consider a perspective which is not purely logical or mathematical, but rather concerned with *information* (cf. Section 4.2.). Suppose that $D_E AB$ has been stated on the basis of present partial knowledge: say, A, B are the denotations of predicates X, Y in E, as far as we know them. Then, *two* types of increase of information are possible. We may learn about new individuals, enlarging E to E', or also obtain new X, Y-information about old individuals, enlarging A, B

to A', B'. It seems reasonable to expect that human communication requires basic terms which are 'stable' under such vicissitudes.

Stability under extension of individuals may be formulated thus. If $E \subseteq E'$, A', $B' \subseteq E$ with $A' \cap E = A$, $B' \cap E = B$, then $D_E AB$ only if $D_{E'} A' B'$. Its corresponding picture is given in Figure 32.

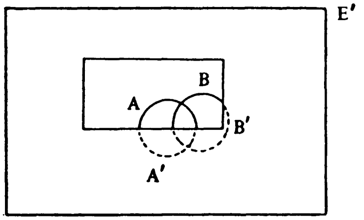

Fig. 32.

It seems preferable to decompose this assertion into two aspects. One is merely (one half of) the earlier principle of *Extension* (cf. Section 1.3.), the other reduces to addition of information concerning already available individuals. Now, stability under extension of information of the latter kind amounts to *upward persistence*, for the argument A, and *upward monotonicity*, for B. The intuitive connection of these two phenomena may be the proper background for the earlier-mentioned universal that 'persistence implies monotonicity'.

This information perspective suggests other notions too. For instance, call a determiner *communicative* if, whenever two people know universes E, E' with $D_E AB$, $D_{E'} A'B'$ such that (upon comparing notes) in the intersection $E \cap E'$, A and A', B and B' coincide, then after pooling their experience, they will have $D_{E \cup E'}(A \cup A')(B \cup B')$. Again, this seems a rather useful type of expression to possess in human languages (see Figure 33).

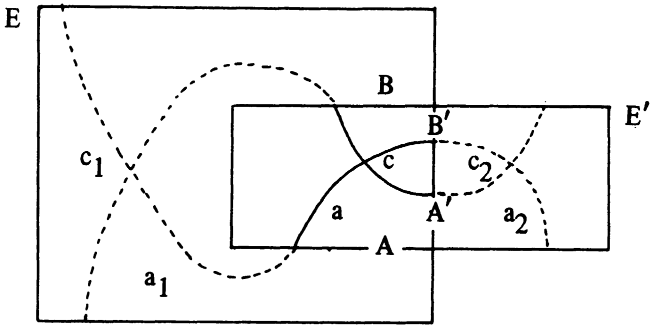

Fig. 33.

For quantifiers, the previous techniques determine the precise content of this class (cf. van Benthem, 1984e): communicative quantifiers are all persistent, and hence first-order definable. In fact, an effective enumeration of their forms can be found using the tree of numbers.

A more systematic theory of information remains to be developed.

The examples so far might suggest that semantic universals will be tied up with determiners, or at least specific lexical items. This is just an accident of presentation, however. For instance, general category structure as such also suggests broad regularities across human languages — witness Keenan and Faltz (1985), which also makes an interesting junction with the existing body of knowledge on comparative linguistics. Moreover, the general perspective in Sections 7.8., 7.9. on 'semantic mechanisms' (such as type change, inference or intensionalization) and their interaction, provides another rich source of general conjectures about natural language.

10.4. DYNAMICS OF INTERPRETATION

Another pronounced tendency in contemporary semantics is its attention for the dynamical aspects of interpretation. The Montagovian scheme is static. But, how are interpretation functions from linguistic items to models actually constructed? Perhaps the most substantial theory to have arisen out of this movement is that of Kamp (1981), in which every grammatical structure generates a *discourse representation* providing the necessary clues (anaphoric and otherwise) for actual interpretation in models.

Even as a medium of description, a separate component of discourse representation offers many heuristic insights (cf. Section 1.1.). Moreover, it also has some interesting theoretical aspects. For instance, the intuitive idea behind the enterprise is that *truth* of a sentence φ in model M will now reduce to *embeddability* of its associated discourse representation $DR(\varphi)$ in M (being a 'small picture' of part of a large world). As was pointed out in van Benthem and van Eyck (1982), this particular idea would make truth of φ preserved under *extensions* of the model M — and so, the intuitive account only works for purely *existential* sentences φ. Accordingly, more sophisticated implementations are needed, and provided.

But again, our earlier point remains, that description is not yet explanation. For instance, empirical observations about bounds on anaphoric dependencies, or co-occurrence restrictions tend to get copied into the definition of discourse representations — whereas one would like to see some independent account. These are early days, of course — and, e.g., Cooper has suggested that the present dynamic framework will be suitable, again, for formulating new types of semantic universal. These might take the form of predicting that only such-and-such anaphoric mechanisms will be realized in all human languages. Moreover, in the spirit of this book, plausible *constraints* could be sought on representational semantics, providing the necessary creative tension.

A suggestion as to the most profitable kind of background theory is found in van Benthem (1983b). Grammatical structures can be any non-circularly interpretable graphs (cf. Section 8.4.); for convenience, say, phrase structure trees. One then needs some processing strategy for these; in the simplest case, a processing order — for which various mathematical options exist ('depth first' or 'breadth first' modes of search; cf. Hintikka, 1979, Barwise and Perry, 1983). In combination with other basic acts, these might form the basis of the desired independent account of anaphora and scope. In this section, a more traditional topic of this ilk will be discussed, however; in line with Chapter 7.

Present dynamic accounts of interpretation revive an old question concerning the relation between *grammaticality* and *interpretability*. According to current wisdom, this is a non-issue: the two classes of expressions coincide, by definition. But in reality, there are divergences galore. *Martina or Claudia met* is grammatical, yet uninterpretable; while *Lucas walks not* is ungrammatical, yet interpretable. Thus, the interplay between *independent* accounts of interpretability and grammaticality will provide interesting comparisons. One might even expect 'completeness' and 'incompleteness' theorems of a kind. Thus, in a sense, we are back with the old issue of recognition versus generation of language.

One simple pilot example will illustrate the idea. Let the 'grammatical' sentences of propositional logic be the Polish ones. Now, call a string of symbols in $\neg$, $\wedge$, p, q, r 'interpretable' if it can be processed to a truth value expression by associating $\neg$, $\wedge$ to the right in the usual way. It is evident that the two classes of expression will coincide. Now, let us liberalize the notion of interpretability to the (plausible) case where we can also associate $\neg$, $\wedge$ to the left, and $\wedge$ in the middle: thus, operators can pick adjoining arguments in any way they please. (Compare the related Lambek treatment of this example in Section 7.4.) Though elementary, the following type of result is instructive:

THEOREM. The interpretable expressions of propositional logic are precisely all permutations of the grammatical ones.

Proof. First, one shows that $\alpha \neg \beta$ is interpretable if and only if $\alpha\beta$ is. Next, one observes that a sequence of proposition letters and conjunction symbols is interpretable if and only if there is precisely one more of the former than of the latter. □

Thus, this 'completeness result' assumes the form

$$\text{INT} = \text{TRANSF (GRAMM)};$$

reviving associations with the early days of transformational grammar. Indeed, there is a whole spectrum to be studied, between the extremes of TRANSF $= \emptyset$ and TRANSF $=$ PERM, the set of all permutations.

A proper setting for such investigations is the earlier *categorial grammar*. For, this would seem to be the most suggestive account of actual interpretation, which also lends itself quite easily to progressive liberalizations of the association process. And that, of course, was exactly what was investigated in Chapter 7 — leading to several relevant results about the variety of constructions and readings which can be obtained for linguistic expressions. For instance, in the most general Lambek calculus, all permutations of interpretable phrases were themselves interpretable (Section 7.3.). Even this extreme case has been claimed to occur in some human languages (cf. Bach, 1984; be it with one 'fixed point'). And of course, with our looser views of grammaticality and interpretability, the stability of meaningfulness under an amount of permutation (whether grammaticality is lost or not), is a phenomenon which a linguistic semantics ought to *explain*, rather than ignore. But most natural languages will have their characteristic restrictions here, reflecting the degree of rigidity of word order imposed. Thus, an interesting principle of linguistic classification arises, in terms of constraints on categorial type change rules, reflecting various degrees of interpretative freedom.

Categorial grammar, of course, does not exhaust all dynamic aspects of natural language interpretation. For instance, higher-order 'interpretative strategies' for text and discourse have remained outside the scope of the present study altogether. Here is where the boundary comes in sight between the study of denotations and that of linguistic 'control', to borrow a term from computer science (cf. van Benthem, 1986c). The next important task lies certainly straight ahead!

10.5. CONCLUSION

Contemporary semantics offers a set of issues, arising from solid research contributions, which invite further logical study. In this chapter, some of these have been arranged in one coherent picture of the nature of semantic theory. No doubt, many further suitable topics will arise in the near future. Thus, there is a subject matter for a logic of semantics. Moreover, its methods also lie at hand: we have used elements of universal algebra, model theory, finite combinatorics, and indeed any respectable formal method. It is the mixture of philosophical concerns about the structure and scope of semantics, and their translation into definite methodological questions, which constitutes one of the main charms of this foundational study. Chronic quarrels may be replaced by clear-cut questions for research. Nevertheless, I am well aware of the earlier controversy surrounding the utility of mathematical linguistics, which could easily be raised here once again. On this point, I would only say this. Logical foundational studies do not replace empirical content. They rather make it clear what is valid *a priori* and what not, thus enabling us to focus more sharply on empirical content as well. The phenomenon of 'natural language' is a conglomerate of brute facts and eternal truths: indeed, the two poles of the human condition.

REFERENCES

Bach, E.: 1984, 'Some Generalizations of Categorial Grammars', in Landman and Veltman (eds.), 1984, pp. 1—23.

Bach, E., R. Oehrle, and D. Wheeler (eds.): 1986, *Proceedings Conference on Categorial Grammar, Tucson 1985*, Reidel, Dordrecht (Studies in Linguistics and Philosophy).

Barwise, J. and R. Cooper: 1981, 'Generalized Quantifiers and Natural Language', *Linguistics and Philosophy* **4**, 159—219.

Bäuerle, R., C. Schwarze, and A. Von Stechow (eds.): 1983, *Meaning, Use and Interpretation of Language*, De Gruyter, Berlin.

Bealer, G.: 1983, 'Completeness in the Theory of Properties, Relations and Propositions', *Journal of Symbolic Logic* **48**, 415—426.

van Benthem, J.: 1978, 'Ramsey Eliminability', *Studia Logica* **37**, 321—336.

van Benthem, J.: 1979, 'Syntactic Aspects of Modal Incompleteness Theorems', *Theoria* **45**, 63—77.

van Benthem, J.: 1980, *Universal Algebra and Model Theory: Two Excursions on the Border*, report ZW79—08, Mathematisch Instituut, Rijksuniversiteit, Groningen.

van Benthem, J.: 1981, 'Why is Semantics what?', in Groenendijk, Janssen and Stokhof (eds.), 1981, pp. 29—50.

van Benthem, J.: 1982a, 'The Logical Study of Science', *Synthese* **51**, 431—472.

van Benthem, J.: 1982b, *The Logic of Time*, Reidel, Dordrecht (Synthese Library, Vol. 156).

van Benthem, J.: 1983a, 'Determiners and Logic', *Linguistics and Philosophy* **6**, 447—478.

van Benthem, J.: 1983b, 'Five Easy Pieces', in ter Meulen (ed.), 1983, pp. 1—17.

van Benthem, J.: 1983c, *Tenses in Real Time*, report 83—26, Dept. of Mathematics, Simon Fraser University, Burnaby. (To appear in *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 1986.)

van Benthem, J.: 1983d, *The Semantics of Variety in Categorial Grammar*, report 83—29, Dept. of Mathematics, Simon Fraser University, Burnaby. (To appear in van Benthem, Buszkowski and Marciszewski (eds.), 1986.)

van Benthem, J.: 1984a, 'Correspondence Theory', in Gabbay and Guenthner (eds.), Vol. II, 1984, pp. 167—247.

van Benthem, J.: 1984b, 'Foundations of Conditional Logic', *Journal of Philosophical Logic* **13**, 303—349.

van Benthem, J.: 1984c, 'Possible Worlds Semantics: A Research Program That Cannot Fail?', *Studia Logica* **43**, 379—393. (An earlier version appeared as report 83—27, Dept. of Mathematics, Simon Fraser University, Burnaby.)

van Benthem, J.: 1984d, 'Questions about Quantifiers', *Journal of Symbolic Logic* **49**, 443—466.

van Benthem, J.: 1984e, 'The Logic of Semantics', in Landman and Veltman (eds.), 1984, pp. 55—80.

van Benthem, J.: 1985a, 'A Linguistic Turn: New Directions in Logic', to appear in Weingartner (ed.), 1986.
van Benthem, J.: 1985b, *A Manual of Intensional Logic*, CSLI Lecture Notes 1, Center for the Study of Language and Information, Stanford.
van Benthem, J.: 1985c, *Modal Logic and Classical Logic*, Bibliopolis, Naples (Indices, Vol. 4).
van Benthem, J.: 1985d, *Semantic Automata*, CSLI Report 85—27, Center for the Study of Language and Information, Stanford. Also to appear in Groenendijk, de Jongh and Stokhof (eds.), 1986.
van Benthem, J.: 1985e, 'The Lambek Calculus', to appear in Bach, Oehrle and Wheeler (eds.), 1986.
van Benthem, J.: 1985f, 'The Varieties of Consequence, According to Bolzano', to appear in *Studia Logica*. (An earlier version appeared as Tech Report 84—6, Center for the Study of Language and Information, Stanford, under the title 'Lessons from Bolzano'.)
van Benthem, J.: 1985g, 'Towards a Computational Semantics', to appear in Cooper, Engdahl and Gärdenfors (eds.), 1986.
van Benthem, J.: 1986a, 'Meaning: Interpretation and Inference', to appear in Synthese Library (M. dalla Chiara and G. Toraldo di Francia (eds.), 'Proceedings Workshop on Theories of Meaning, Florence 1985').
van Benthem, J.: 1986b, 'Strategies of Intensionalization', to appear in *Filozófiai Figyelö* (L. Pólos (ed.), 'Festschrift for Imre Ruzsa').
van Benthem, J.: 1986c, 'The Ubiquity of Logic in Natural Language', to appear in W. Leinfellner (ed.), *Proceedings Tenth International Wittgenstein Symposium, Kirchberg 1985*.
van Benthem, J. and H. Doets: 1983, 'Higher-Order Logic', in Gabbay and Guenthner (eds.), Vol. I, pp. 275—329.
van Benthem, J. and J. van Eyck: 1982, 'The Dynamics of Interpretation', *Journal of Semantics* **1**, 3—20.
van Benthem, J. and A. ter Meulen (eds.): 1985, *Generalized Quantifiers in Natural Language*, Foris, Dordrecht (GRASS series, Vol. 4).
van Benthem, J. and D. Pearce: 1984, 'A Mathematical Characterization of Interpretation between Theories', *Studia Logica* **43**, 295—303.
van Benthem, J., W. Buszkowski, and W. Marciszewski (eds.): 1986, *Categorial Grammar*, John Benjamin, Amsterdam.
Büchi, J.: 1960, 'Weak Second-Order Arithmetic and Finite Automata', *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* **6**, 66—92.
Burgess, J.: 1981, 'Quick Completeness Proofs for Some Logics of Conditionals', *Notre Dame Journal of Formal Logic* **22**, 76—84.
Buszkowski, W.: 1982, *Lambek's Categorial Grammars*, dissertation, Instytut Matematyki, Uniwersytet Im. Adama Mickiewicza, Poznań.
Carnap, R.: 1956, 'The Methodological Character of Theoretical Concepts', in Feigl and Scriven (eds.), 1956, pp. 33—76.
Chellas, B.: 1980, *Modal Logic: an Introduction*, Cambridge University Press, Cambridge
Cooper, R., E. Engdahl, and P. Gärdenfors (eds.): 1986, *Proceedings Workshop on Generalized Quantifiers, Lund 1985*, Reidel, Dordrecht (Studies in Linguistics and Philosophy).

Cresswell, M.: 1973, *Logics and Languages*, Methuen, London.

Davidson, D. and G. Harman (eds.): 1972, *Semantics of Natural Language*, Reidel, Dordrecht (Synthese Library, Vol. 40).

van Deemter, K.: 1985, 'Generalized Quantifiers: Finite versus Infinite', in van Benthem and ter Meulen (eds.), 1985, pp. 145—159.

Doets, H., 1983: see van Benthem and Doets, 1983.

Ehrig, H., H-J. Kreowski, and P. Padawitz: 'Stepwise Specification and Implementation of Abstract Data Types', in *Proceedings ICALP 5 (1978)*, Springer, Berlin, Lecture Notes in Computer Science, Vol. 62, pp. 205—226.

Englebretsen, G.: 1981, *Three Logicians*, van Gorcum, Assen.

van Eyck, J.: 1982: see van Benthem and van Eyck, 1982.

van Eyck, J.: 1985, *Aspects of Quantification in Natural Language*, dissertation, Filosofisch Instituut, Rijksuniversiteit, Groningen.

Feferman, S.: 1969, 'Persistent and Invariant Formulas for Outer Extensions', *Compositio Mathematica* **28**, 29—52.

Feigl, H. and M. Scriven (eds.): 1956, *Minnesota Studies in the Philosophy of Science*, Vol. 1, University of Minnesota Press, Minneapolis.

Flynn, M.: 1983, 'A Categorial Theory of Structure-Building', in G. Gazdar, E. Klein, and G. Pullum (eds.), 1983, pp. 139—174.

Gabbay, D. and F. Guenthner (eds.): 1983, *Handbook of Philosophical Logic*, Vol. I, Reidel, Dordrecht.

Gabbay, D. and F. Guenthner (eds.): 1984, *Handbook of Philosophical Logic*, Vol. II, Reidel, Dordrecht.

Gallin, D.: 1975, *Intensional and Higher-Order Modal Logic*, North-Holland, Amsterdam (Mathematics Studies, Vol. 19).

Gazdar, G., E. Klein, and G. Pullum (eds.): 1983, *Order, Concord and Constituency*, Foris, Dordrecht.

Geach, P.: 1972, 'A Program for Syntax', in Davidson and Harman (eds.), 1972, pp. 483—497.

Gerson, M.: 1975, 'The Inadequacy of Neighbourhood Semantics for Modal Logic', *Journal of Symbolic Logic* **40**, 141—148.

Ginsburg, S.: 1966, *The Mathematical Theory of Context-Free Languages*, McGraw-Hill, New York.

Ginsburg, S. and E. Spanier: 1966, 'Semi-Groups, Pressburger Formulas and Languages', *Pacific Journal of Mathematics* **16**, 285—296.

Goldblatt, R.: 1976, 'Metamathematics of Modal Logic', *Reports on Mathematical Logic* **6**, 4—77 and **7**, 21—52.

Groenendijk, J. and M. Stokhof: 1984, *Studies on the Semantics of Questions and the Pragmatics of Answers*, dissertation, Filosofisch Instituut, Universiteit van Amsterdam, Amsterdam.

Groenendijk, J., T. Janssen, and M. Stokhof (eds.): 1981, *Formal Methods in the Study of Language*, Mathematical Centre, Amsterdam (Mathematical Centre Tracts 135/136).

Groenendijk, J., D. de Jongh, and M. Stokhof (eds.): 1986, *Information, Interpretation and Inference*, Foris, Dordrecht (GRASS series, Vol. 5), to appear.

Harper, W., R. Stalnaker, and G. Pearce (eds.): 1981, *Ifs*, Reidel, Dordrecht.

Hintikka, J.: 1979, 'Quantifiers in Natural Language: Some Logical Problems', in Saarinen (ed.), 1979, pp. 81—117.

Hoeksema, J.: 1983, 'Plurality and Conjunction', in ter Meulen (ed.), 1983, pp. 63—83.
Hoeksema, J.: 1984, *Categorial Morphology*, dissertation, Nederlands Instituut, Rijksuniversiteit, Groningen.
Hopcroft, J. and J. Ullman: 1979, *Introduction to Automata Theory, Languages and Computation*, Addison-Wesley, Reading (Mass.).
Hughes, G. and M. Cresswell: 1968, *Introduction to Modal Logic*, Methuen, London.
Janssen, T.: 1983, *Foundations and Applications of Montague Grammar*, dissertation, Mathematical Centre, Amsterdam.
Kamp, H.: 1981, 'A Theory of Truth and Semantic Representation', in Groenendijk, Janssen and Stokhof (eds.), 1981, pp. 277—322. (Reprinted in GRASS series, Vol. 2, *Truth, Interpretation and Information*, Foris, Dordrecht, 1984, pp. 1—41.)
Keenan, E. and L. Faltz: 1985, *Boolean Semantics for Natural Language*, Reidel, Dordrecht (Synthese Language Library, Vol. 23).
Keenan, E. and L. Moss: 1985, 'Generalized Quantifiers and the Expressive Power of Natural Language', in van Benthem and ter Meulen (eds.), 1985, pp. 73—124.
Keenan, E. and Y. Stavi: 1982, 'A Semantic Characterization of Natural Language Determiners', to appear in *Linguistics and Philosophy*.
Ladusaw, W.: 1979, *Polarity Sensitivity as Inherent Scope Relations*, dissertation, University of Texas, Austin.
Lambek, J.: 1958, 'The Mathematics of Sentence Structure', *American Mathematical Monthly* **65**, 154—170.
Landman, F. and I. Moerdijk: 1981, *Morphological Features and Conditions on Rules in Montague Grammar*, Amsterdam Papers in Formal Grammar, Vol. 3, Centrale Interfaculteit, Universiteit van Amsterdam, Amsterdam.
Landman, F. and F. Veltman (eds.): 1984, *Varieties of Formal Semantics*, Foris, Dordrecht (GRASS series, Vol. 3).
Lenzen, W.: 1980, *Glauben, Wissen und Wahrscheinlichkeit*, Springer, Wien.
Lewis, D.: 1973, 'Counterfactuals and Comparative Similarity', in Harper (ed.), 1981, pp. 57—85.
Lewis, D.: 1981, 'Ordering Semantics and Premise Semantics for Counterfactuals', *Journal of Philosophical Logic* **10**, 217—234.
Lindström, P.: 1966, 'First-Order Predicate Logic with Generalized Quantifiers', *Theoria* **32**, 186—195.
Löbner, S.: 1984, 'Quantification as a Major Module of Natural Language Semantics', to appear in Groenendijk, de Jongh and Stokhof (eds.), 1986.
Mason, I.: 1983, 'Undecidability of the Metatheory of the Propositional Calculus', *Journal of Symbolic Logic* **50** (1985), 451—457.
McNaughton, W. and S. Papert; 1971, *Counter-free Automata*, MIT Press, Cambridge (Mass.).
Mellor, D. H. (ed.): 1978, *Foundations, Essays in Philosophy, Mathematics and Economics*, Routledge & Kegan Paul, London.
Mendelson, E.: 1964, *Introduction to Mathematical Logic*, Van Nostrand, Princeton.
ter Meulen, A. (ed.): 1983, *Studies in Model-theoretic Semantics*, Foris, Dordrecht (GRASS series, Vol. 1).
Montague, R.: 1974, *Formal Philosophy* (R. Thomason, ed.), Yale university Press, New Haven.
Moortgat, M.: 1984, 'Functional Composition and Complement Inheritance', to appear in *Proceedings Conference on Meaning and the Lexicon* (*Cleves 1983*), Nijmegen.

Partee, B.: 1985, 'Syntactic Categories and Semantic Types', to appear in Groenendijk, de Jongh and Stokhof (eds.), 1986.
Partee, B. and M. Rooth: 1983, 'Generalized Conjunction and Type Ambiguity', in Bäuerle, Schwarze and Von Stechow (eds.), 1983.
Pearce, D., 1984: see van Benthem, J. and D. Pearce, 1984.
Przełecki, M.: 1969, *The Logic of Empirical Theories*, Routledge & Kegan Paul, London.
Quine, W.: 1966, *Selected Logic Papers*, Random House, New York.
Ramsey, F.: 1929, 'Theories', in Mellor (ed.), 1978, pp. 101—125.
Saarinen, E. (ed.): 1979, *Game-theoretical Semantics*, Reidel, Dordrecht (Synthese Language Library, Vol. 5).
Scott, D.: 1971, 'On Engendering an Illusion of Understanding', *Journal of Philosophy* **68**, 787—808.
Segerberg, K.: 1971, *An Essay in Classical Modal Logic*, Filosofiska Studier, University of Uppsala, Uppsala.
Simon, T. and R. Scholes (eds.): 1982, *Language, Mind and Brain*, Lawrence Erlbaum, Hilldale (N.J.).
Smoryński, C.: 1984, 'Modal Logic and Self-Reference', in Gabbay and Guenthner (eds.), Vol. II, 1984, pp. 441—495.
Sneed, J.: 1971, *The Logical Structure of Mathematical Physics*, Reidel, Dordrecht (Synthese Library, Vol. 35).
Sommers, F.: 1982, *The Logic of Natural Language*, Cambridge University Press, Cambridge.
Suppes, P.: 1982, 'Variable-free Semantics with Remarks on Procedural Extensions', in Simon and Scholes (eds.), 1982, pp. 21—31.
Thomason, S.: 1974, 'An Incompleteness Theorem in Modal Logic', *Theoria* **40**, 30—34.
Thijsse, E.: 1983, *Laws of Language*, M. A. Thesis, Filosofisch Instituut, Rijksuniversiteit, Groningen.
Thijsse, E.: 1985, 'Counting Quantifiers', in van Benthem and ter Meulen (eds.), 1985, pp. 125—144.
Väänänen, J.: 1980, 'A Quantifier for Isomorphisms', *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* **26**, 123—130.
Veltman, F.: 1985, *The Logic of Conditionals*, dissertation, Filosofisch Instituut, Universiteit van Amsterdam, Amsterdam.
Weingärtner, P. (ed.): 1986, *Proceedings 7th International Congress of Logic, Methodology and Philosophy of Science, Salzburg 1983*, North-Holland, Amsterdam (Studies in Logic), to appear.
Westerståhl, D.: 1982, 'Logical Constants in Quantifier Languages', to appear in *Linguistics and Philosophy* (1985).
Westerståhl, D.: 1984, 'Some Results on Quantifiers', *Notre Dame Journal of Formal Logic* **25**, 152—170.
Westerståhl, D.: 1985, 'Quantifiers in Natural Language and Formal Languages', to appear in Gabbay and Guenthner (eds.), 1985 (Vol. IV).
Zwarts, F.: 1981, 'Negatief-polaire Uitdrukkingen', *Glot* **4**, 35—132.
Zwarts, F.: 1983, 'Determiners: A Relational Perspective', in ter Meulen (ed.), 1983, pp. 37—62.
Zwarts, F.: 1986, *Model-theoretic Semantics and Natural Language: The Case of Modern Dutch*, dissertation, Nederlands Instituut, Rijksuniversiteit, Groningen.

INDEX OF NAMES

INDEX OF SUBJECTS