REPORT OF AN OSLO SEMINAR

IN

LOGIC AND LINGUISTICS

edited by

Jens Erik Fenstad

Contents

INTRODUCTION

Jens Erik Fenstad

In the fall of 1976 a psychologist (R. Rommetveit), a linguist (E. Hovdhaugen), a philosopher (D. Føllesdal), and a mathematical logician (J. E. Fenstad) decided to give a seminar together on language as seen from their various perspectives. This was an interdisciplinary venture, the topic turned out to be fashionable, and we had an overflow audience - at least until the mathematical logician had given some introductory lectures on Montague's PTQ. After that our audience was reduced to a more comfortable size. But enough interest had been generated for the seminar to continue. And it has served as a useful meeting place. The "hard core" has always been a group of linguists and logicians, but from time to time we have also had the active participation of philosophers, psychologists, and computer scientists.

After the grand opening we settled down to understanding the impact of the Montague paradigm on theoretical linguistics, and it was natural for us to focus on the semantics of natural languages and the interrelationship between syntax and semantics. A result of this learning phase was the lecture Models for natural languages [1] which I gave to the 4th Scandinavian Logic Symposium. It represents a logician's reaction to Montague's PTQ. As a logician I could make explicit the use Montague had made of the notion of generalized quantifier. (In addition to "surface" uses it was also pointed out that an unspecified generalized quantifier seems to lurk below the surface in passive constructions. If we assert that Mary is liked, we leave open whether by one, several, many, most.... Perhaps this explains why politicians love the passive voice.) I also pointed out how the Montague analysis could be given a more lexical bent. 'Love' as a bivalent verb has a natural interpretation as a relation between two individuals. This is the lexial "fact". If the compositionality requirements of your favorite syntactic analysis requires something different, you can always lift in type using the technique of λ-abstraction. Further an analysis of passive was given that was attuned to this lexical point of view and which seems to be similar in spirit to some current lexical treatments of passive.

The lecture concluded by noting several shortcomings of the Montague paradigm. The model frame $\mathfrak{A} = \langle A,F,I \rangle$ carries too little structure and needs to be enriched by further computational and/or geometric content. And I also emphasized that in a speaker/listener situation it is misguided to try to force all of the pragmatics into the index set I of the structure $\mathfrak{A}$. We have to deal with partial interpretations, conflicting interpretations, the building up of an interpersonal interpretation; and interpretations may be modified or determined by functionals which carry along the presuppositions introduced.

Except for a few specific points such as the remarks on generalized quantifiers much of what was said in the lecture was either expository or of a general and programmatic character. Later in our seminar we returned to and discussed further the connection between partiality and presuppositions. I introduced into the Montague framework a partial functional $R(\alpha;\beta)$ where α

is of arbitrary type and β is of type t, $R(\alpha;\beta)$ is then of the same type as α. The semantics is as follows:
$\|R(\alpha;\beta)\| = \|\alpha\|$ if $\|\beta\| = 1$, $\|R(\alpha;\beta)\|$ is undefined otherwise. (Note that e.g. $\|\beta\| = 1$ means that $\|\beta\|$ is defined and is equal to 1; with partiality there are "weak" and "strong" notions, this we do not enter into here.) The R-functional was applied by C. Fabricius-Hansen in Wieder éin wieder? Zur Semantik von wieder [2]. It is well-known that wieder (again) has both an external (repetitive) and an internal (restitutive) reading. In her paper C. Fabricius-Hansen gave a unified analysis within a Montague PTQ framework extending the partial R-functional to account for the associated presuppositions. (To indicate, very briefly, part of the repetitive wieder may be captured in the translation wieder' = $\lambda p.R(p;\underline{P}p)$, where $\underline{P}$ is the past tense operator. The full story, much improving on Dowty's account in Word Meaning and Montague Grammar, can be found in [2].)

An important event in the history of the seminar was the Workshop on Models for Natural Languages which we arranged in Oslo in the fall of 1980 and which helped us to establish broader international contacts. At that time some candidate theses were being written as an offshot of the activities in the seminar (a few each in lingustics, computer science, and mathematics). In this Report we have reproduced revised versions of the parts of three candidate theses which were submitted to the Department of Mathematics in 1982/83.

Helle Frisak Sem gives in the first part of her thesis, Quantifier Scope and Coreferentiality a modified treatment of the quantifier storage mechanism developed by R. Cooper in Quantification and Syntactic Theory. In particular, she eliminates the use of multi-valued functions in the semantic interpretation. Cooper's storage mechanism is successful in treating many questions of quantifier scope in natural languages, but there are definite limitations. Many of these can be overcome by using Hans Kamp's theory of Discourse Representation Systems, and in the second part of her thesis Helle Sem extends Kamp's analysis to cover the Cooper system and is thus able to obtain a satisfactory treatment of some of the examples that could not be handled by the storage method.

In the final part she studies the connection between Discourse Representation Systems and the theory of Situation Semantics as developed by J. Barwise and J. Perry.

Tore Langholm develops in his thesis, Some Tentative Systems Relating to Situation Semantics, some formal theories covering part of the semantics of situations. He studies first a propositional system L(S) based on a model structure $\langle \Omega, w, [\ ] \rangle$, where Ω is a non-empty set of situations, $w \in \Omega$ is the actual "world" and $[\ ]$ is a valuation on the prositional variables, $[p] = \langle [p]^+, [p]^- \rangle$, where $[p]^+, [p]^- \subseteq \Omega$ and $[p]^+$ is the set of situations which supports the truth of p. $[p]^-$ is the set of situations that definitely refutes the truth of p. We assume that $w \in [p]^+ \cup [p]^-$ and that $w \notin [p]^+ \cap [p]^-$, but in general neither $[p]^+ \cup [p]^- = \Omega$, nor $[p]^+ \cap [p]^- \neq \emptyset$, thus situations may both be partial and contradictory.

Langholm gives a complete axiomatization of L(S) and also proves similar results for a first-order theory L(SQ). In the final part he uses his systems to study naked infinitive perceptual reports. His thesis is a first and useful step toward a more comprehensive mathematical study of situation semantics.

Jan Tore Lønning develops in his thesis, Mass Terms and Quantification, a novel theory of the semantics of mass terms. The basic semantic entity is an atomless Boolean algebra and both mass noun phrases and the homogeneously referring verb phrases denotes elements in this algebra. This "simple" idea seems to get right the various troublesome features of previous accounts. To give some precision to his treatment Lønning constructs a fragment of English, a suitable logical formalism to correspond to this fragment, and the appropriate model theory. A mathematical investigation is carried out (completeness theorems, decision procedures). And looking at the formalism in different ways (i.e. using the appropriate representation theorems) he explains how it is possible (i.e. consistent) to either view mass terms as "individuals" (singular terms) or to view them as "properties" (general terms). Thus some of the earlier philosophical discussion do not seem too profitable. In the final sections of his thesis Lønning extends the fragment to include count nouns and amount terms.

An underlying premiss for our activities in the seminar has been our belief in the fruitfulness of the interaction between logic and linguistics. I hope that the work reported on in this Report gives substance to this belief. We plan to continue the Oslo Seminar and hope, in particular, to strengthen our competence in the computational aspects of natural language modeling.

Jens Erik Fenstad

[1] Fenstad, J.E., Models for Natural Languages, in: Hintikka et al. (eds.), Essays on Mathematical and Philosophical Logic, pp. 315-340.
D. Reidell Publ. Comp. Dordrecht 1978.

[2] Fabricius-Hansen, C., Wieder éin wieder? Zur Semantik von wieder, in: Bäuerle et al. (eds.), Meaning, Use, and Interpretation of Language, pp. 97-120.
W. de Gruyter, Berlin, New York, 1983.

SOME TENTATIVE SYSTEMS RELATING TO SITUATION SEMANTICS[1]

Tore Langholm

The motivation for considering the systems of this article is to be found in the field of Situation Semantics, which is currently being developed by Jon Barwise and John Perry. Our systems are not meant to catch the full complexity of the richer theory. Rather, they are intended as tools with which to study some of the more elementary properties of situations, and to provide the opportunity for a sort of reconnaissance trip into areas which eventually will have to be conquered by a more comprehensive formal theory.

1. THE PROPOSITIONAL SYSTEM L(S).

In every-day life, the truth-value of a proposition depends upon the situation at hand. In the first part we shall try to make this idea precise by studying a simple propositional language based on a semantics of situations.

1.1. Definition of L(S).

Symbols: propositional variables $p, q, p_1, q_1, \ldots$

connectives $\neg, \vee, \wedge, \Rightarrow$

auxiliary symbols (,)

Formation rules:

(1) All propositional variables are formulae.

(2) If A and B are formulae, so are $\neg A$, $(A \vee B)$, $(A \wedge B)$ and $(A \Rightarrow B)$.

By $(A \Rightarrow B)$ we want to express that B is true in every situation in which A is true. We have to discuss what logical principles ought to be valid. In [Barwise 1979] the following two rules are proposed:

(1) $(A \wedge B)$ is true on the basis of s just in case both A and B are true on the basis of s.

(2) $(A \vee B)$ is true on the basis of s just in case A or B or both are true on the basis of s.

This can also be inferred from the rules on pages 137 and 138 of [Barwise and Perry 1983]. If these rules are accepted, the following formulae and rules of inference become valid:

$$(A \wedge B) \Rightarrow (B \wedge A) \qquad (A \wedge B) \Rightarrow A$$

$$(A \vee B) \Rightarrow (B \vee A) \qquad A \Rightarrow (A \vee B)$$

$$\frac{A \Rightarrow B \quad A \Rightarrow C}{A \Rightarrow (B \wedge C)} \qquad \frac{A \Rightarrow C \quad B \Rightarrow C}{(A \vee B) \Rightarrow C}$$

As for negation, it appears obvious that A is true on the basis of s just in case $\neg A$ is false on the basis of s, and that $\neg A$ is true on the basis of s just in case A is false on the basis of s. This gives us the following principles:

$$A \Rightarrow \neg\neg A \qquad \neg\neg A \Rightarrow A$$

Also it seems reasonable to accept de Morgans laws in the present context:

$$\neg(A \wedge B) \Rightarrow (\neg A \vee \neg B)$$

$$(\neg A \vee \neg B) \Rightarrow \neg(A \wedge B)$$

$$\neg(A \vee B) \Rightarrow (\neg A \wedge \neg B)$$

$$(\neg A \wedge \neg B) \Rightarrow \neg(A \vee B)$$

1.1.1. The Formulae $(A \wedge \neg A) \Rightarrow B$ and $A \Rightarrow (B \vee \neg B)$ in Situation Semantics.

If no formula can be both true and false on the basis of the same situation, then $(A \wedge \neg A) \Rightarrow B$ is a valid principle for any A and B.

Motivated by a wish to start out with the broadest possible frame, we will, however, admit inconsistent situations, hence not accept the unrestricted validity of $(A \wedge \neg A) \Rightarrow B$. As we shall later see, inconsistent situations can be eliminated from our models by adding $(A \wedge \neg A) \Rightarrow B$ as an extra axiom scheme.

The principle $A \Rightarrow (B \vee \neg B)$ will be rejected on the grounds that a proposition may be totally foreign to a situation, and thus neither be true nor false on the basis of it.

From the above remarks, we construct the following truth tables for $\vee$ and $\wedge$, where t, nt, f and nf abbreviate true, not true, false and not false respectively.

(1)

A	B	$(A \vee B)$
t	t	t
t	nt	t
nt	t	t
nt	nt	nt

(2)

A	B	$(A \wedge B)$
t	t	t
t	nt	nt
nt	t	nt
nt	nt	nt

By de Morgans laws and the rules

(3) A true iff $\neg A$ false
A false iff $\neg A$ true

we derive

(4)

A	B	(A∨B)
nf	nf	nf
nf	f	nf
f	nf	nf
f	f	f

(5)

A	B	(A∧B)
nf	nf	nf
nf	f	f
f	nf	f
f	f	f

1.1.2. A Four-valued Logic for Situations.

We see that there are four possible truth-values for propositions with respect to situations, namely

true and not false (1)
true and false (↯)
neither true nor false (?)
false and not true (0)

From the above tables we derive the following (uniquely determined) truth tables for situations:

(6)

A	B	(A∨B)	(A∧B)
1	1	1	1
1	?	1	?
1	↯	1	↯
1	0	1	0
?	1	1	?
?	?	?	?
?	↯	1	0
?	0	?	0
↯	1	1	↯
↯	?	1	0
↯	↯	↯	↯
↯	0	↯	0
0	1	1	0
0	?	?	0
0	↯	↯	0
0	0	0	0

From rules (3) the truth table for negation is constructed:

(7)

A	$\neg A$
1	0
?	?
↯	↯
0	1

Observe that these truth tables are closed with respect to the set $\{0,1\}$. For these truth values they also agree with ordinary two-valued logic. To ensure that a situation defines an ordinary two element valuation, we therefore only have to require that each propositional variable is either true or false and not both.

1.2. The Semantics for L(S):

In our semantics the interpretation of a proposition will be the set of situations supporting its truth. More precisely, we assume that there is given a set Ω of situations. An interpretation is determined by a function from propositional variables to sets of situations. For every propositional variable p there will be two subsets $|p|^+$, $|p|^- \subseteq \Omega$, where $|p|^+$ is the set of situations which render p true and $|p|^-$ the set of situations which render it false. According to our discussion under 1.1.1 $|p|^+ \cap |p|^-$ may be non-empty (existence of "inconsistent" situations) and $|p|^+ \cup |p|^-$ may be a proper subset of Ω. We denote by $|p|$ the ordered pair $\langle |p|^+, |p|^- \rangle$, thus an interpretation $|\ |$ is a map from propositional variables to the set $\mathcal{P}(\Omega)\times\mathcal{P}(\Omega)$.

We shall further assume that Ω contains a particular element w - the "world situation" - which we single out as the key to truth and falsity. A proposition will be considered true or false just in case it is true or false on the basis of w. To ensure that all propositions are true or false and not both, we impose the requirements that

$$w \in |p|^{+} \cup |p|^{-} \quad \text{and} \quad w \notin |p|^{+} \cap |p|^{-}$$

for all propositional variables p.

With these preliminary explanations we come to the formal definition:

1.2.1. Definition of L(S) Model.

A model or structure for L(S) is a triple $\mathfrak{A} = \langle \Omega, w, |\ | \rangle$, where

(1) Ω is a non-empty set of situations.

(2) w is a distinguished element of Ω.

(3) $|\ |$ is a function from the set of propositional variables of L(S) into $\mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$.

We write $|p| = \langle |p|^{+}, |p|^{-} \rangle$ and require that

(4) $w \in |p|^{+} \cup |p|^{-}$ and $w \notin |p|^{+} \cap |p|^{-}$ for all propositional variables p of L(S).

The function $|\ |$ gives an interpretation for the propositional variables. This will be extended to all formulae by the following definition:

1.2.2. Definition of the Interpretation $\|A\|_{\mathfrak{A}} = \langle \|A\|_{\mathfrak{A}}^{+}, \|A\|_{\mathfrak{A}}^{-} \rangle$.

(i) Propositional variables: $\|p\|_{\mathfrak{A}} = |p|$

(ii) Negation: $\|\neg A\|_{\mathfrak{A}}^{+} = \|A\|_{\mathfrak{A}}^{-}$

$\|\neg A\|_{\mathfrak{A}}^{-} = \|A\|_{\mathfrak{A}}^{+}$

(iii) Disjunction: $\|A \vee B\|_{\mathfrak{A}}^{+} = \|A\|_{\mathfrak{A}}^{+} \cup \|B\|_{\mathfrak{A}}^{+}$

$\|A \vee B\|_{\mathfrak{A}}^{-} = \|A\|_{\mathfrak{A}}^{-} \cap \|B\|_{\mathfrak{A}}^{-}$

(iv) Conjunction: $\|A \wedge B\|_{\mathfrak{A}}^{+} = \|A\|_{\mathfrak{A}}^{+} \cap \|B\|_{\mathfrak{A}}^{+}$

$\|A \wedge B\|_{\mathfrak{A}}^{-} = \|A\|_{\mathfrak{A}}^{-} \cup \|B\|_{\mathfrak{A}}^{-}$

The reader will see how these clauses correspond to the truth-tables above. When no confusion is likely to occur, we drop the index $\mathfrak{A}$ and write more simply $\|A\|$ for $\|A\|_{\mathfrak{A}}$.

(v) Strong Implication: From our previous explanations it follows that we want $(A\Rightarrow B)$ to be true, i.e. $w \in \|A\Rightarrow B\|^{+}$, just in case $\|A\|^{+} \subseteq \|B\|^{+}$. In other situations there seems to be no preferred way to define the truth-value of $(A\Rightarrow B)$ from $\|A\|$ and $\|B\|$. Considerations of simplicity have led us to the following choice:

$$\|A\Rightarrow B\| = \begin{cases} \langle\Omega,\emptyset\rangle, & \text{if } \|A\|^{+} \subseteq \|B\|^{+} \\ \langle\emptyset,\Omega\rangle, & \text{otherwise} \end{cases}$$

1.2.3. Definition of Validity.

A formula A of $L(S)$ is true in the model $\mathfrak{A}$, in symbols $\mathfrak{A} \models A$, iff $w \in \|A\|^{+}$. A is true or valid, in symbols $\models A$, if A is true in all models $\mathfrak{A}$ of $L(S)$.

Let as usual $A \supset B$ abbreviate $\neg A \vee B$. It follows from 1.2.2 that the following formulae are valid:

$$(A\Rightarrow B) \supset (C\Rightarrow(A\Rightarrow B))$$

$$\neg(A\Rightarrow B) \supset (C\Rightarrow\neg(A\Rightarrow B))$$

$$(A\Rightarrow B) \supset (\neg(A\Rightarrow B)\Rightarrow C)$$

$$\neg(A\Rightarrow B) \supset ((A\Rightarrow B)\Rightarrow C)$$

It should be pointed out that these formulae are very marginal to the theory. We could just as well have made them not well formed. Our chief concern lies with formulae of which no subformula is within the scope of more than one occurrence of $\Rightarrow$.

1.2.4. A Remark on Complete Sets of Connectives for L(S).

We close this section with some remarks on the truth-functional connectives $\vee$, $\wedge$ and $\neg$. They do not comprise a functionally complete set of connectives with respect to the four truth values 0, +, ⥮ and 1. This should be no cause for concern, since the contrary would be positively undesireable. If we had formulae corresponding to truth functions which are not closed with respect to the set $\{0,1\}$, w could not be made to define a two element valuation. However, it could be of interest to find a set of connectives complete with respect to the set of functions which preserve two-valued logic. Such a set is obtained by adding the connectives T and - with the following truth tables:

A	T(A)	-A
1	1	0
?	0	⥮
⥮	1	?
0	0	1

An absolutely complete set is obtained by adding (to $\{\wedge, \vee, \neg, T, -\}$) a constant u denoting ? (or $\langle\emptyset,\emptyset\rangle$ in the semantics of L(S).)

It is, however, difficult to find intuitive counterparts to these connectives (with the possible exceptions of T), and since the systems in this article are constructed with analysis of natural languages in mind, I cannot see any reason to include them. We could have been forced to introduce them in order to achieve a complete axiomatization. But this turned out not to be the case; hence they can safely be disregarded.

1.3. Valuation Structures.

Although the relationship between situations and propositions is defined as a membership relation, it will sometimes be useful to think of situations as valuations on propositions. Given a model $\mathfrak{A}$ of L(S) and a situation $s \in \Omega$ we define:

$$\|A\|_{\mathfrak{A},s} = \begin{cases} 1 & \text{if } s \in \|A\|^{+}_{\mathfrak{A}} \text{ and } s \notin \|A\|^{-}_{\mathfrak{A}} \\ \updownarrow & \text{if } s \in \|A\|^{+}_{\mathfrak{A}} \text{ and } s \in \|A\|^{-}_{\mathfrak{A}} \\ ? & \text{if } s \notin \|A\|^{+}_{\mathfrak{A}} \text{ and } s \notin \|A\|^{-}_{\mathfrak{A}} \\ 0 & \text{if } s \notin \|A\|^{+}_{\mathfrak{A}} \text{ and } s \in \|A\|^{-}_{\mathfrak{A}} \end{cases}$$

From the correspondence between truth tables (6) and (7) and definition 1.2.2 we see that for any $\mathfrak{A}$ and s, $\| \ \|_{\mathfrak{A},s}$ will be a valuation conforming to (6) and (7). Moreover, $\| \ \|_{\mathfrak{A},w}$ will be an ordinary two-valued valuation.

1.3.1. Definition of Simple Formulae.

A simple formula is a formula of L(S) without occurrences of the connective =>.

1.3.2. Definition of Valuation Structure.

A structure $\mathfrak{A} = \langle \Omega, w, | \ | \rangle$ is called a <u>valuation structure</u> if Ω is a set of valuations on the simple formulae of L(S) according to the truth tables (6) and (7), and for which

$$v(p) = \|p\|_{\mathfrak{A},v}$$

for all $v \in \Omega$ and all propositional variables p of L(S).

In other words, the situations of a valuation structure are valuations, and their valuations of propositional variables correspond to their memberships in the positive and negative extensions of the propositional variables. Any set of valuations (containing a two element valuation), together with a distinguished element in the set (being a two element valuation), uniquely determines a valuation structure. From the remarks above it follows that in a valuation structure $\mathfrak{A}$, $v(A) = \|A\|_{\mathfrak{A},v}$ for all simple formulae A.

1.3.3. Lemma.

Every L(S) structure is elementarily equivalent to a valuation structure.

The proof is straightforward.

1.3.4. Theorem.

A=>B, with A and B simple, is valid if and only if the joint truth table of A and B has 1 or $\mathfrak{t}$ in the column for B in any line in which A is assigned the value 1 or the value $\mathfrak{t}$.

Proof: It is immaterial to the truth of a formula of the type A=>B which two element valuation (if there are more) is the designated one. We also see that for any pair $\mathfrak{A}_1$, $\mathfrak{A}_2$ of valuation structures and pair A,B of simple formulae, if $\Omega_1 \subseteq \Omega_2$ and $\mathfrak{A}_2 \models A \Rightarrow B$, then $\mathfrak{A}_1 \models A \Rightarrow B$ follows. Now let $\mathfrak{B}$ be a valuation structure built up from all possible valuations on the simple formulae. By the remark above, if $\mathfrak{B} \models A \Rightarrow B$, A=>B is valid in any valuation structure, and hence in any structure.

From this line of reasoning a stringent proof can easily be constructed.

1.4. Axiomatization.

The following axiomatization is an adaption of the axiomatization of the System E_{fde} of Anderson and Belnap. The connection to E_{fde} will be explored in the next chapter.

Axiom schemes.

(A1) $(A \wedge B) \Rightarrow A$

(A2) $(A \wedge B) \Rightarrow B$

(A3) $\neg A \Rightarrow \neg(A \wedge B)$

(A4) $\neg B \Rightarrow \neg(A \wedge B)$

(A5) $A \Rightarrow (A \vee B)$

(A6) $B \Rightarrow (A \vee B)$

(A7) $\neg(A \vee B) \Rightarrow \neg A$

(A8) $\neg(A \vee B) \Rightarrow \neg B$

(A9) $A \Rightarrow \neg\neg A$

(A10) $\neg\neg A \Rightarrow A$

(A11) $\neg(A \wedge B) \Rightarrow (\neg A \vee \neg B)$

(A12) $(\neg A \wedge \neg B) \Rightarrow \neg(A \vee B)$

(A13) $(A \wedge (B \vee C)) \Rightarrow ((A \wedge B) \vee C)$

(A14) $(A \Rightarrow B) \supset ((A \Rightarrow C) \supset (A \Rightarrow (B \wedge C)))$

(A15) $(A \Rightarrow C) \supset ((B \Rightarrow C) \supset ((A \vee B) \Rightarrow C))$

(A16) $(A \Rightarrow B) \supset ((B \Rightarrow C) \supset (A \Rightarrow C))$

(A17) $(A \Rightarrow B) \supset (A \supset B)$

(A18) $(A \Rightarrow B) \supset (C \Rightarrow (A \Rightarrow B))$

(A19) $\neg(A \Rightarrow B) \supset (C \Rightarrow \neg(A \Rightarrow B))$

(A20) $(A \Rightarrow B) \supset (\neg(A \Rightarrow B) \Rightarrow C)$

(A21) $\neg(A \Rightarrow B) \supset ((A \Rightarrow B) \Rightarrow C)$

Rules of inference:

Modus Ponens: From A and $A \supset B$ we may infer B.

1.4.1. Theorem.

Every theorem is valid in every structure.

Proof: Validity of (A1)-(A13) is verifiable by 1.3.4. (A14)-(A16) are simple set-theoretic principles involving the interpretation rules for $\vee$ and $\wedge$. (A17) follows from the fact that $w \in \Omega$. The validity of (A18)-(A21) is a consequence of our particular choice for the interpretation of $(A \Rightarrow B)$. Finally, Modus Ponens clearly preserves validity.

1.4.2. Remark.

Since the axioms are given by axiom schemes (and M.P. is the only rule of inference), every substitution instance of a theorem will be a theorem. Hence we should expect the same for validity; the following reasoning shows this to be the case:

Since no restrictions are posed as to how the truth values of propositional variables are to be defined in various situations (except completeness and consistency for w), the interpretations of formulae in general cannot vary more freely than those of the propositional variables. This means that given a distribution of interpretations to some set of formulae, the same distribution can be attained by a corresponding set of propositional variables, and from this it is seen how a counter-example for some substitution instance A' of a formula A may be converted into a counter-example for A.

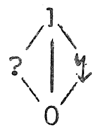
Note that this is not as trivial as it might perhaps seem. If, for instance, we had required that each propositional variable has a situation supporting its truth, substitution would not preserve validity.

1.5. Comparison to the System E of Entailment.

The semantics described in 1.2.1 and 1.2.2 gives L(S) some important features which makes it natural to compare it to the system E of entailment, which was constructed in order to catch the intuitive notions of entailment and relevancy.

1.5.1. An Equivalent Decision Procedure.

The decision procedure of 1.3.4 is equivalent to a slightly different one. We may define the partial ordering ⩽ on the set of truth values by letting $⩽£ if $ and £ are identical, or if £ is above $ in the figure below:



Now accept A=>B as valid if v(A)⩽v(B) for all valuations v.

The test may seem stricter than the first, since we do not allow <?,0>, <?,⇅> and <1,⇅>. However, because of the symmetry between ? and ⇅ in the truth tables; if there exists a v for which v(A) = ? and v(B) = 0, there is a v' for which v'(A) = ⇅ and v'(B) = 0. Similar remarks hold for <?,⇅>/<⇅,?> and <1,⇅>/<1,?>.

1.5.2. L(S) and E_{fde}.

In the system E, a first degree entailment formula is a formula A=>B with A and B containing only truth-functional connectives. The system E_{fde} (first degree entailment fragment of the calculus E) is characterized by the decision procedure above (cfr. [Anderson & Belnap 1975]). In other words, the set of valid L(S)-formulae within the set {A=>B|A,B simple} coincides with E_{fde}.

In proving the completeness of L(S), it would be a waste of effort to start from scratch rather than finding some way to benefit from the works of Anderson and Belnap. In "Entailment" they prove the following to be a complete axiomatization of E_{fde}.

Axiom schemes

$$(A \wedge B) \Rightarrow A$$

$$(A \wedge B) \Rightarrow B$$

$$A \Rightarrow (A \vee B)$$

$$B \Rightarrow (A \vee B)$$

$$(A \wedge (B \vee C)) \Rightarrow ((A \wedge B) \vee C)$$

$$A \Rightarrow \neg\neg A$$

$$\neg\neg A \Rightarrow A$$

Rules of inference:

$$\frac{A \Rightarrow B \qquad B \Rightarrow C}{A \Rightarrow C} \qquad \frac{A \Rightarrow B \qquad A \Rightarrow C}{A \Rightarrow (B \wedge C)} \qquad \frac{A \Rightarrow C \qquad B \Rightarrow C}{(A \vee B) \Rightarrow C} \qquad \frac{A \Rightarrow B}{\neg B \Rightarrow \neg A}$$

1.5.3. The System $L(S)_{fde}$.

We want to prove that the set of theorems of E_{fde} is cointained in the set of theorems of L(S). In order to do this, we define the systems $L(S)_{fde}$ in the following way:

Axiom schemes: (A1)-(A13).

Rules of inference:

$$\frac{A \Rightarrow B \qquad B \Rightarrow C}{A \Rightarrow C} \qquad \frac{A \Rightarrow B \qquad A \Rightarrow C}{A \Rightarrow (B \wedge C)} \qquad \frac{A \Rightarrow C \qquad B \Rightarrow C}{(A \vee B) \Rightarrow C}$$

1.5.4. Lemma.

A formula of the type A=>B with A and B simple is valid in L(S) iff it is valid in E_{fde}.

Proof: Since E_{fde} is characterized by the decision procedure of 1.5.1, this follows from the result of that section.

1.5.5. Lemma.

Every valid formula of E_{fde} is a theorem of E_{fde}.

Proof: Cfr. [Anderson & Belnap 1975].

1.5.6. Lemma.

Every theorem of E_{fde} is a theorem of $L(S)_{fde}$.

Proof: All axioms of E_{fde} are axioms of $L(S)_{fde}$. In order to prove that any theorem of E_{fde} is a theorem of $L(S)_{fde}$, we need only to show that if A=>B is a theorem of $L(S)_{fde}$, so is ¬B=>¬A.

The proof os this (by induction on the derivation of theorems) is straightforward but lengthy, and we omit it here. Cfr. pp. 29-34 of my cand.scient. thesis.

1.5.7. Lemma.

Every theorem of $L(S)_{fde}$ is also a theorem of L(S).

Proof: All axioms of $L(S)_{fde}$ are axioms of L(S). The rules of inferense of $L(S)_{fde}$ are derivable in L(S) by Modus Ponens and (A14)-(A16).

1.5.8. Theorem.

Every valid formula $A \Rightarrow B$ of $L(S)$ with A and B simple is a theorem of $L(S)$.

Proof: This follows immediately from lemmas 1.5.4 to 1.5.7.

1.6. L(S) and Ordinary Two-valued Propositional Calculus.

1.6.1. Lemma.

Let A be a simple formula which is also a tautology of ordinary two-valued propositional calculus. If $p_1,\ldots,p_n$ are the propositional variables of A, then

$$((\neg p_1 \vee p_1)\wedge\ldots\wedge(\neg p_n \vee p_n)) \supset A$$

is a theorem of $L(S)$.

Proof: Validity of $((\neg p_1 \vee p_1)\wedge\ldots\wedge(\neg p_n \vee p_n)) \Rightarrow A$ is verified by 1.3.4. The rest then follows by 1.5.8, A17 and M.P..

1.6.2. Theorem.

If A is a tautology of ordinary two-valued propositional calculus, then A is a theorem of $L(S)$.

Proof: This follows from 1.6.1 by M.P. and iterated use of the fact that $(\neg p_i \vee p_i)$ and $A \supset (B \supset (A \wedge B))$ are theorems of $L(S)$.

1.6.2.1. Remark.

As the proof is carried out, it applies only to simple formulae. But any tautologous formula A is a substitution instance of a simple, tautologous formula A'. A proof of A' is easily converted into a proof of A by substituting throughout.

1.6.3. Definition.

$A_1,\ldots,A_n \vdash B$ means that B is a theorem of the system obtained by adding $A_1,\ldots,A_n$ to the axioms of L(S).

1.6.4. Theorem (Deduction Theorem).

If $A_1,\ldots,A_n \vdash B$, then $A_1,\ldots,A_{n-1} \vdash A_n \supset B$.

Proof: This property belongs to any system with M.P. as the only rule of inference, which contains all formulae of the types

$$(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C)) \quad \text{and} \quad A \supset (B \supset A)$$

as theorems.

1.7. p_1 Normal Form.

The reduction theorem of S5 states that every formula is equivalent to a formula in which no modal operator occurs within the scope of another modal operator. The similarity in structure between L(S) and S5, as opposed to T or S4, i.e. the fact that $\|A{=}{>}B\|^+$ is either Ω or $\emptyset$, leads us to expect something similar for L(S).

1.7.1. Definition.

A formula is in p_1 normal form if every subformula of the type A=>B has the following property: A is either simple or it is the formula $p_1{=}{>}p_1$. B is either simple, or it is the formula $\neg(p_1{=}{>}p_1)$. (p_1 is the first propositional variable.)

1.7.2. Theorem.

Let as usual $A \equiv B$ abbreviate $(A \supset B) \wedge (B \supset A)$. Given a formula A of $L(S)$, we can effectively find a formula A' of p_1 normal form such that $A \equiv A'$ is a theorem of $L(S)$.

The proof of this, which is by induction on the construction of formulae and makes essential use of axiom schemes (A18)-(A21), is straightforward but very tedious, and we omit it here. Cfr. pp. 48-61 of my cand.scient. thesis.

1.8. Theorem (Decision Procedure).

There is an effective procedure to check whether a formula of $L(S)$ is valid.

1.8.1. Restriction to a Subset of Formulae.

In view of 1.7.2, we need only prove decidability for formulae of p_1 normal form.

But we may simplify further. Relying on results of the ordinary propositional calculus, we know that there is an effective way to check whether an arbitrary formula F in p_1 normal form is a tautology ($(A \Rightarrow B)$ is considered atomary). If F is a tautology, it is valid. If it is not, there is an effective way to find a formula F' in conjunctive normal form which is tautologically equivalent to F. (And hence true in an $L(S)$ structure iff F is true.) Checking whether F' is valid in all structures, is equivalent to checking whether each of its conjuncts is valid in all structures.

Each conjunct looks like this (cfr. 1.8.1.1 - E and all indexed capital letters represents simple formulae.):

$$
\begin{array}{ll}
 & E \\
 & \vee(\bigvee_{i\leq m}\neg(G_i \Rightarrow H_i)) \\
 & \vee(\bigvee_{i\leq n}\neg((p_1 \Rightarrow p_1)\Rightarrow D_i)) \\
(1) & \vee(\bigvee_{i\leq r}\neg(L_i \Rightarrow \neg(p_1 \Rightarrow p_1))) \\
 & \vee(\bigvee_{i\leq s}(A_i \Rightarrow B_i)) \\
 & \vee(\bigvee_{i\leq k}((p_1 \Rightarrow p_1)\Rightarrow C_i)) \\
 & \vee(\bigvee_{i\leq \ell}(K_i \Rightarrow \neg(p_1 \Rightarrow p_1))) \qquad m,n,r,s,k,l\geq 0
\end{array}
$$

To check the validity of this is equivalent to checking the following (The special case $\Gamma\models$ should be read as $\models\neg\Gamma$, i.e. $\models\neg(\bigwedge_{A\in\Gamma} A)$):

$$
\begin{array}{lll}
 & \Gamma\models & (\bigvee_{i\leq s}(A_i \Rightarrow B_i)) \\
(2) & & \vee(\bigvee_{i\leq k}((p_1 \Rightarrow p_1)\Rightarrow C_i)) \\
 & & \vee(\bigvee_{i\leq \ell}(K_i \Rightarrow \neg(p_1 \Rightarrow p_1)))
\end{array}
$$

where Γ is the set

$$
\{\neg E\} \cup \{G_i \Rightarrow H_i\}_{i\leq m} \cup \{(p_1 \Rightarrow p_1)\Rightarrow D_i\}_{i\leq n} \cup \{L_i \Rightarrow \neg(p_1 \Rightarrow p_1)\}_{i\leq r}
$$

1.8.1.1. Remark.

In (1) of 1.8.1 we assume that the conjunct does not contain $(p_1 \Rightarrow p_1)\Rightarrow\neg(p_1 \Rightarrow p_1)$ or $\neg((p_1 \Rightarrow p_1)\Rightarrow\neg(p_1 \Rightarrow p_1))$.

When this fails, we have one of the following cases:

(i) The conjunct contains $\neg((p_1 \Rightarrow p_1)\Rightarrow\neg(p_1 \Rightarrow p_1))$ as a disjunct, and is hence valid.

(ii) The conjunct is the formula $(p_1 \Rightarrow p_1)\Rightarrow\neg(p_1 \Rightarrow p_1)$, and is hence invalid.

(iii) The conjunct properly contains $(p_1 \Rightarrow p_1)\Rightarrow\neg(p_1 \Rightarrow p_1)$ as a disjunct, and is equivalent to the formula obtained by deleting this disjunct.

1.8.2. Set-up for the Decision Procedure.

In the joint truth table of E and all G_i, H_i, D_i, L_i, A_i, B_i, C_i and K_i, delete all lines with one of the following properties:

(i) For some $i \leq m$, G_i has one of the values 1 or $1\!\!\downarrow$, and H_i has not.

(ii) For some $i \leq n$, D_i has a value distinct from 1 and $1\!\!\downarrow$.

(iii) For some $i \leq r$, L_i has a value distinct from 0 and $?$.

1.8.3. Decision.

The validity relation (2) holds if one of the following is true:

(i) None of the lines in the reduced truth table defines a two element valuation.

(ii) None of the lines in the reduced truth table does at the same time define a two element valuation, and assign $\neg E$ the value 1.

(iii) For some $i \leq s$, all lines which assign A_i the value 1 or $1\!\!\downarrow$, also assign B_i one of these values.

(iv) For some $i \leq k$, it is the case that all lines assign C_i 1 or $1\!\!\downarrow$.

(v) For some $i \leq \ell$, it is the case that all lines assign K_i 0 or $?$.

The procedure covers all possible special cases. If there is no formula E, the possibility of validity by (ii) vanishes. If Γ is empty or the second, third or fourth line in (1) does not exist, the reduction of the truth table is correspondingly less extensive. If the fifth, sixth or seventh line in (1) does not exist, the possibility of truth by (iii), (iv) or (v) vanishes. All lines in (1) cannot vanish at the same time, since then we would not have a formula in the first place.

1.8.4. Proposition.

If (2) holds, then the procedure concludes that (2) holds.

Proof: Suppose the procedure concludes that the validity relation does not hold. We then know that the reduced truth table contains a two element valuation. In case there is a formula E, there is a two element valuation supporting the truth of ¬E. For each formula $A_i \Rightarrow B_i$ there is a line which assigns A_i, but not B_i, 1 or $\downarrow$. For each C_i there is a line assigning C_i 0 or ?. For each L_i there is a line assigning L_i 1 or $\downarrow$.

These lines each defines a valuation, and from this set of valuations we construct a valuation structure, letting a two element valuation (which supports the truth of ¬E if there is such a formula) be the distinguished element. It is immediate that this constitutes a counterexample to (2).

1.8.5. Proposition.

If the procedure concludes that (2) holds, then (2) holds.

Proof: Suppose (2) does not hold. There is then a counter-example $\mathfrak{A}$ which by 1.3.3 is elementary equivalent to a valuation structure $\mathfrak{A}'$. All valuations in $\mathfrak{A}'$ correspond to lines in the reduced truth table. For if a valuation is present in the counter-example (which must validate Γ), it has defied all reductions on the previous page. For each of the formulae $A_i \Rightarrow B_i$, $(p_1 \Rightarrow p_1) \Rightarrow C_i$ and $K_i \Rightarrow \neg(p_1 \Rightarrow p_1)$, there are situations in $\mathfrak{A}'$ which prevent them from being true; and hence lines in the reduced truth table which prevent the decision procedure from according (2) truth.

1.8.6. Corollary.

From the formulation of the decision procedure, it immediately follows that a validity relation of the kind indicated in (2) holds just in case

$$\models \neg\Gamma$$

or for some i,

$$\Gamma \models A_i \Rightarrow B_i$$

or $\Gamma \models (p_1 \Rightarrow p_1) \Rightarrow C_i$

or $\Gamma \models K_i \Rightarrow \neg(p_1 \Rightarrow p_1)$

This result will be useful when we prove completeness in the next section.

1.9. Theorem (Completeness).

Every valid formula of $L(S)$ is a theorem of $L(S)$.

1.9.1. Proposition.

Completeness follows if we are able to show the following implications ($\vdash \neg\Gamma$ should be read as $\vdash \neg(\bigwedge_{A\in\Gamma} A)$):

(1) $\Gamma \models A \Rightarrow B \longrightarrow \Gamma \vdash A \Rightarrow B$

(2) $\models \neg\Gamma \longrightarrow \vdash \neg\Gamma$

(3) $\Gamma \models (p_1 \Rightarrow p_1) \Rightarrow C \longrightarrow \Gamma \vdash (p_1 \Rightarrow p_1) \Rightarrow C$

(4) $\Gamma \models K \Rightarrow \neg(p_1 \Rightarrow P_1) \longrightarrow \Gamma \vdash K \Rightarrow \neg(p_1 \Rightarrow p_1)$

when Γ is the set ($m \geq 0$, and E, D, L do not necessarily exist):

$$\{E\} \cup \{(p_1 \Rightarrow p_1) \Rightarrow D\} \cup \{L \Rightarrow \neg(p_1 \Rightarrow p_1)\} \cup \{G_i \Rightarrow H_i\}_{i \leq m}.$$

(All capital letters represent simple formulae).

Proof: By 1.7.2 and 1.6.2 we only need to show that validity entails provability for formulae of type (1) of 1.8.1. (1.8.1.1 applies to provability as well as validity.) By 1.6.4, this is equivalent to showing that (2) implies the corresponding provability relation. By 1.8.6 this is equivalent to proving (1)-(4) above for Γ of the type specified in 1.8.1.

The assumption that Γ contains at most one formula of the type $(p_1 => p_1) => D$ is justified by the fact that

$$(((p_1=>p_1)=>D_1)\wedge((p_1=>p_1)=>D_2))\equiv((p_1=>p_1)=>(D_1\wedge D_2))$$

is a theorem of $L(S)$. Similarly for $L=>\neg(p_1=>p_1)$.

1.9.2. <u>Proposition</u>.

Implication (1) of 1.9.1 is true of $L(S)$.

Proof: By adopting a technique used in the completeness proof of "de Morgan Implication" (which is the same as E_{fde}) in [Makinson 1973], it can be shown that if $\Gamma \not\models A=>B$, then there must exist a valuation v which satisfies $v(D)\in\{1,\updownarrow\}$, $v(L)\notin\{1,\updownarrow\}$ and $v(G_i)\in\{1,\updownarrow\}\longrightarrow v(H_i)\in\{1,\updownarrow\}$, but not $v(A)\in\{1,\updownarrow\}\longrightarrow v(B)\in\{1,\updownarrow\}$. Also, if $\Gamma \not\models A=>B$, then Γ must be consistent. Hence, also $\{E\wedge D\wedge\neg L\} \cup \{G_i \supset H_i\}_{i\leq n}$ is consistent. Since $L(S)$ contains the ordinary propositional calculus, this implies the existence of a two element valuation w such that $w(E\wedge D\wedge\neg L) = 1$ and $w(G_i) = 1 \longrightarrow w(H_i) = 1$. It is easily checked that the valuation structure based on $\{v,w\}$ with w as the distinguished element constitutes a counterexample to $\Gamma\models A=>B$.

<u>1.9.3. Proposition.</u>

Implication (2) of 1.9.1 is true of L(S).

Proof: If $\models \Gamma$, then $\Gamma \models (p_1 \vee \neg p_1) \Rightarrow (p_1 \wedge \neg p_1)$. $\Gamma \vdash (p_1 \vee \neg p_1) \Rightarrow (p_1 \wedge \neg p_1)$ now follows by 1.9.2. Since $\neg((p_1 \vee \neg p_1) \Rightarrow (p_1 \wedge \neg p_1))$ is a theorem, we must then have $\vdash \neg\Gamma$.

<u>1.9.4. Lemma.</u>

Let Π be a subset of $\{E\} \cup \{L \Rightarrow \neg(p_1 \Rightarrow p_1)\} \cup \{G_i \Rightarrow H_i\}_{i \leq n}$ (i.e. Π does not contain $(p_1 \Rightarrow p_1) \Rightarrow D$). If Π has a model, then $\Pi \not\models (p_1 \Rightarrow p_1) \Rightarrow C$ for any simple C.

Proof: Suppose $\mathfrak{A} \models \Pi$ for a valuation structure $\mathfrak{A}$. A new valuation structure $\mathfrak{A}'$ is derived from $\mathfrak{A}$ by adding to the situation domain of $\mathfrak{A}$ the constant valuation assigning all simple formulae the value ?. E, $L \Rightarrow \neg(p_1 \Rightarrow p_1)$ and all $G_i \Rightarrow H_i$ remain unchanged, but $\mathfrak{A}' \not\models (p_1 \Rightarrow p_1) \Rightarrow C$.

<u>1.9.5. Proposition.</u>

Implication (3) of 1.9.1 is true of L(S).

Proof: Suppose $\Gamma \models (p_1 \Rightarrow p_1) \Rightarrow C$. There are two possible cases, corresponding to whether Γ has a model or not. If Γ has no model, $\vdash \neg\Gamma$ follows by 1.9.3, and hence $\Gamma \vdash (p_1 \Rightarrow p_1) \Rightarrow C$ since L(S) contains the propositional calculus. If Γ has a model, $((p_1 \Rightarrow p_1) \Rightarrow D) \in \Gamma$ for some D by the lemma. Since $((p_1 \Rightarrow p_1) \Rightarrow D) \supset (((p_1 \Rightarrow p_1) \Rightarrow C) \equiv (D \Rightarrow C))$ is a theorem, the rest now follows by 1.9.2.

1.9.6. Proposition.

Implication (4) of 1.9.1 is true of L(S).

The proof is similar to the proof of 1.9.5, and uses a lemma similar to 1.9.4.

1.10. Closed Strctures.

A set E of valuations is closed if it has the following property (we suppose all propositional variables occur in the sequence $\langle p_i \rangle_{i \in \mathbb{N}}$):

For any valuation u if for all n there is a valuation $v_n \in E$ such that $v_n(p_i) = u(p_i)$ for all $i \leq n$, then $u \in E$.

A structure is closed if its situation domain corresponds to a closed set of valuations.

1.10.1. Topological Compactness.

The expression "closed set of valuations" is topologically motivated. We define the metric d on the set of valuations in the following way:

$$d(v_1, v_2) = \sum_{v_1(p_i) \neq v_2(p_i)} 2^{-i}.$$

A set of valuations will then be closed in the sense above if and only if it is closed with respect to the topology defined by d. This topology is also a product topology of a very trivial, compact topology. Hence it is itself compact. This means that it has the following important property:

If the intersection of every finite subfamily of a family of closed sets is non-empty, the intersection of the family itelf is non-empty.

As will become apparent from the proof of the next theorem, the notion of topological compactness is closely related to the notion of logical compactness.

1.10.2. Theorem (Logical Compactness).

A set of formulae has a model if and only if each finite subset has a model.

Proof: This follows if we can show that every consistent set of formulae has a model.

1.10.2.1. Contruction of the Model $\mathfrak{A}_\Delta$.

Suppose the set Π is consistent. We extend Π to a maximal consistent set Δ. Let V be the set of all valuations on the simple formulae. For each simple formula A, we define F_A to be the set $\{v \in V \mid v(A) \in \{1, 1\!\!\downarrow\}\}$. Hence F_A is both open and closed. We now define Ω_Δ as follows:

$$\Omega_\Delta = \bigcap_i (\tilde{F}_{A_i} \cup F_{B_i}) \cap (\bigcap_i F_{C_i}) \cap (\bigcap_i \tilde{F}_{D_i})$$

$\{A_i \Rightarrow B_i\}_i$ is the set of all formulae $A \Rightarrow B$ in Δ (A, B simple).
$\{C_i\}_i$ is the set of simple formulae C for which $((p_1 \Rightarrow p_1) \Rightarrow C) \in \Delta$.
$\{D_i\}_i$ is the set of simple formulae D for which $(D \Rightarrow \neg(p_1 \Rightarrow p_1)) \in \Delta$.
The simple formulae contained in Δ define a two element valuation w. It is easily checked that $w \in \Omega_\Delta$. We now want to show that the valuation structure $\mathfrak{A}_\Delta$ based on Ω_Δ with w as the distinguished element, is a model for Δ.

1.10.2.2. Proposition.

For all simple A, $A \in \Delta$ iff A is true in $\mathfrak{A}_\Delta$.

Proof: A is true in $\mathfrak{A}_\Delta$ iff $w \in \|A\|^+_{\mathfrak{A}_\Delta}$, i.e. iff $w(A) = 1$, i.e. iff $A \in \Delta$.

1.10.2.3. Proposition.

If $(A \Rightarrow B) \in \Delta$, then $A \Rightarrow B$ is true in $\mathfrak{A}_\Delta$.

Proof: This follows immediately from the construction of Ω_Δ.

1.10.2.4. Proposition.

If $A \Rightarrow B$ is true in $\mathfrak{A}_\Delta$, then $(A \Rightarrow B) \in \Delta$.

Proof: Suppose $A \Rightarrow B$ is true in $\mathfrak{A}_\Delta$. This means that $\Omega_\Delta \cap F_A \cap \tilde{F}_B = \emptyset$. By topological compactness, it now follows that $\Omega_{\Delta_0} \cap F_A \cap \tilde{F}_B = \emptyset$ for some finite subset Δ_0 of Δ. Clearly under these conditions the set $\Delta_0 \cup \{\neg(A \Rightarrow B)\}$ can have no model. This is because $\mathfrak{A}_{\Delta_0}$ is in a sense a maximal model of Δ_0, and if a model of Δ_0 contains a counterexample to $A \Rightarrow B$, this counterexample must be present in $\mathfrak{A}_{\Delta_0}$. By completeness $\Delta_0 \cup \{\neg(A \Rightarrow B)\}$ must then be inconsistent, i.e. $\Delta_0 \vdash A \Rightarrow B$. Accordingly $\Delta \vdash A \Rightarrow B$, and by maximal consistency $(A \Rightarrow B) \in \Delta$.

1.10.2.5. Proposition.

For every formula A of $L(S)$, $A \in \Delta$ iff $\mathfrak{A}_\Delta \models A$.

Proof: By 1.7.2, we only need to show this for formulae of the types A, $(A \Rightarrow B)$, $(p_1 \Rightarrow p_1) \Rightarrow A$, $A \Rightarrow \neg(p_1 = p_1)$ with A and B simple. We have already showed it for the two former types. The proofs for the two latter types are similar to 1.10.2.3 and 1.10.2.4.

1.10.3. Theorem.

Every structure is elementarily equivalent to a closed valuation structure.

Proof: The proof of 1.10.2 applies.

1.11. Sideview to Modal Logic.

1.11.1. The Systems L(CS) and L(CS)+.

Consider the following structural constraint:

(1) $|p|^{+} \cap |p|^{-} = \emptyset$ for all p.

This corresponds to a ban on inconsistent situations, and a narrowing of attention to (strong) three element valuations.

When L(S) is used to describe such structures, we call it L(CS). (Language of Consistent Situations.)

A complete axiomatization of L(CS) is obtained by adding to (A1)-(A21) the axiom scheme

(A22) $(A \wedge \neg A) \Rightarrow B$.

An interesting group of L(CS) structures are those which satisfy the following:

(2) Given a situation s and a finite set $\{q_1, \ldots, q_n\}$ of propositional variables such that $\|q_i\|_s \neq ?$ for all $i \leq n$. For any propositional variable p there is then a situation t for which $\|p\|_t \neq ?$ and $\|q_i\|_s = \|q_i\|_t$ for all $i \leq n$.

When L(CS) is used to describe such structures, we call it L(CS)+. A complete axiomatization of L(CS)+ is obtained by adding to (A1)-(A22) the axiom scheme:

(A23) $(((A \wedge B) \Rightarrow\!> \neg A) \wedge ((A \wedge \neg B) \Rightarrow\!> \neg A)) \supset (A \Rightarrow\!> \neg A)$

These completeness results are easily derived from the completeness theorem for L(S). The proofs by themselves do not reveal any interesting properties, and we omit them here. Cfr. pp. 94-96 of my cand.scient. thesis.

1.11.2. Lemma.

Every L(CS)+ structure is elementary equivalent to a closed L(CS)+ structure.

Proof: Let $\mathfrak{A}$ be and L(CS)+ structure. $\mathfrak{A}$ is also an L(S) structure, and hence elementarily equivalent to a closed L(S)-structure $\mathfrak{A}'$. All A22 and A23 axioms are true in $\mathfrak{A}$, and therefore in $\mathfrak{A}'$. Since every L(S) structure satisfying all the A22 axioms is an L(CS) structure, and every L(CS) structure satisfying all A23 axioms is an L(CS)+ structure, $\mathfrak{A}'$ must be an L(CS)+ structure.

1.11.3. Definition.

s is a subsituation of t if

$s \in |p|^+ \longrightarrow t \in |p|^+$ and

$s \in |p|^- \longrightarrow t \in |p|^-$

for all propositional variables p.

1.11.4. Definition.

A situation is a possible world if it defines a two element valuation.

1.11.5. Lemma.

A closed L(CS) structure is an L(CS)+ structure iff every situation in the structure is a subsituation of a possible world in the structure.

Proof: Obviously, an L(CS) structure in which every situation is a subsituation of a possible world, is an L(CS)+ structure. Let s be an arbitrary situation in a closed structure for L(CS)+. Let $q_1, \dots, q_n, \dots$ be an enumeration of those propositional variables q for which $|q|_s \neq ?$. Since $\mathfrak{A}$ is an L(CS)+ structure, for any n there is a sequence $\langle t^n_r \rangle_r$ of situations in Ω which all agree with s for $q_1, \dots, q_n$, and which converges towards a possible world. This possible world must then be a member of Ω. Hence we have shown that

$$F_{A_n} \cap \Omega \cap PW \neq \emptyset \quad \text{for all} \quad n$$

where A_n is $Q_1 \wedge \dots \wedge Q_n$ ($Q_i = q_i$ if $|q_i|_s = 1$, $Q_i = \neg q_i$ if $|q_i|_s = 0$) and PW is the set of possible worlds. Since PW is closed, $(\bigcap_n F_{A_n}) \cap \Omega \cap PW \neq \emptyset$ by compactness, and there is a possible world in Ω which agrees with s for all q_i, i.e. in relation to which s is a subsituation. This completes the proof.

1.11.6. Definition.

$\Box A$ is an abbreviation of $\neg A \Rightarrow \neg(A \Rightarrow A)$. An M-formula is a formula which is built up using the connectives $\Box$, $\vee$, $\wedge$ and $\neg$ only. $\Box A$ is true iff A is false in no situation. We want to compare this operator to the necessity operator of S5.

1.11.7. Lemma.

If the M-formula A is valid in L(S), then it is valid in S5.

Proof: From a set of two element valuations on simple formulae, and a distinguished member of the set, both an S5 structure and an L(S) structure can be built. The rules for interpretation of formulae in the connectives $\Box$, $\vee$, $\wedge$ and $\neg$ are seen to coincide, hence a formula in these connectives is true in both the S5 structure and the L(S) structure, or false in both.

Since all S5 structures are of this kind, any M-formula which is valid in L(S), is also valid in S5. The converse does not hold, however, since the formula $\Box(A \supset B) \supset (\Box A \supset \Box B)$ is valid in S5 but not in L(S). On the other hand, it is valid in L(CS)+. This will follow from a theorem below.

1.11.8. Lemma.

If the M-formula A is valid in S5, then it is valid in all closed L(CS)+ structures.

Proof: Let $\mathfrak{A}$ be an arbitrary, closed L(CS)+ structure, and let $\mathfrak{A}\upharpoonright PW$ be the structure obtained by removing from $\mathfrak{A}$ all situations which are not possible worlds. We must show that $\mathfrak{A}$ and $\mathfrak{A}\upharpoonright PW$ are elementarily equivalent with respect to the set of M-formula.

By 1.11.5 there is a function $f:\Omega \to \Omega\upharpoonright PW$ such that s is a subsituation of f(s) for all s. Since we must have f(w) = w, and $\|A\|_{\mathfrak{A},w} \neq ?$ for all A, the lemma will follow from the proposition below:

1.11.8.1. Proposition.

For all situations s, if $\|A\|_{\mathfrak{A},s} \neq ?$ and A is an M-formula, then $\|A\|_{\mathfrak{A},s} = \|A\|_{\mathfrak{A}\upharpoonright PW,f(s)}$.

Proof: We show this by induction on the construction of M-formulae.

Basis: For propositional variables, this is immediate from the definition of subsituations.

The induction steps for the truth-functional connectives are straightforward.

Consider $\Box A$. If there is a counterexample in $\mathfrak{A}$, i.e. an s such that $\|A\|_{\mathfrak{A},s} = 0$, $\|A\|_{\mathfrak{A}\upharpoonright PW,f(s)} = 0$ by the induction hypothesis, and so there is also a counter-example in $\mathfrak{A}\upharpoonright PW$. Suppose there is a counter-example in $\mathfrak{A}\upharpoonright PW$, i.e. an u such that $\|A\|_{\mathfrak{A}\upharpoonright PW,u} = 0$. Since $u \in PW$, $\|A\|_{\mathfrak{A},u} \in \{0,1\}$. Suppose (for reductio ad absurdum) that $\|A\|_{\mathfrak{A},u} = 1$. Then $\|A\|_{\mathfrak{A}\upharpoonright PW,f(u)} = 1$ by the induction hypothesis. But $f(u) = u$ (since $u \in PW$), so this gives a contradiction. Hence we must have $\|A\|_{\mathfrak{A},u} = 0$, and so there is a counter-example in $\mathfrak{A}$.

We have now seen that $\Box A$ is true in $\mathfrak{A}$ iff it is true in $\mathfrak{A}\upharpoonright PW$. Accordingly, since $\|\Box A\| = \langle \Omega, \emptyset \rangle$ or $\|\Box A\| = \langle \emptyset, \Omega \rangle$ in all structures, $\Box A$ is either true in all situations in both $\mathfrak{A}$ and $\mathfrak{A}\upharpoonright PW$, or it is false in all situation in both. This completes the induction step for $\Box$, and hence the proof by induction.

1.11.9. Theorem.

If A is a M-formula, then A is valid in L(CS)+ iff it is valid in S5.

Proof: This follows from 1.11.7, 1.11.8 and 1.11.2.

1.11.10. Remark.

The theorem states that S5 is in a sence a subsystem of L(CS)+. The converse, however, does not hold. It is not possible to define => from the connectives □, ∨, ∧ and ¬. In the proof above, it was shown that a M-formula cannot distinguish between $\mathfrak{A}$ and $\mathfrak{A}\restriction PW$ when $\mathfrak{A}$ is a closed structure for L(CS)+. If => were defineable using these connectives only, no formula of L(S) could distinguish between $\mathfrak{A}$ and $\mathfrak{A}\restriction PW$. But this is clearly not so. If $\mathfrak{A}$ contains a situation s for which $|p|_s$ = ?, ¬((p=>p)=>(p∨¬p)) is true in $\mathfrak{A}$ but not in $\mathfrak{A}\restriction PW$.

2. EXTENDING L(S) TO QUANTIFICATIONAL LOGIC - THE SYSTEM L(SQ).

We will now consider a quantificational extension of L(S). More or less sophisticated modes of quantification (e.g. generalized quantifiers) could be used, but for the sake of simplicity we have chosen the ordinary first order quentifiers.

However, the question of how to define such quantifiers turns out to have no obvious answer. If we associate with each situation a set of individuals - those present in the situation, we have the following alternatives: Is $\forall xA$ true in the situation s when A is true in s for all individuals of s, or should it be all individuals whatsoever? L(SQ) contains two distinct universal quantifiers, each corresponding to one of the alternatives above.

2.1. Definition of L(SQ).

Symbols:

A countably infinite set of individual variables	$x_1, x_2, \ldots$
For each n and m such that $m \leq n$, a countably infinite set of n,m-ary relation symbols	$R_1, R_2, \ldots$
Logical symbols	$\neg, \wedge, \vee, \Rightarrow, \forall$, a
Auxiliary symbols	(,)

By an n-ary relation symbol we will understand a relation symbol which is n,m-ary for some m. The difference between n,m-ary and n,k-ary relation symbols will only become apparent through the semantics. Syntactically they are interchangeable, though it will be possible to decide from the appearance of a relation symbol whether it is n,m-ary or n,k-ary (by indices or otherwise).

Formation rules:

If $x_1,\dots,x_n$ are variables and R is an n-ary relation symbol, then $R(x_1,\dots,x_n)$ is a formula.

If A_1 and A_2 are formulae, so are $(A_1 \wedge A_2)$, $(A_1 \vee A_2)$, $\neg A_1$ and $(A_1 \Rightarrow A_2)$.

If A is a formula and x is a variable, $\forall xA$ and axA are formulae.

2.2. The Semantics for L(SQ).

2.2.1. Definition of L(SQ) Model.

A Model or Structure for L(SQ) is a quintuple $\langle \Omega, w, D, dom, |\ | \rangle$

(1) Ω is the set of situations

(2) w is a distinguished element of Ω

(3) D is the set of individuals

(4) dom is a function from Ω into $\mathcal{P}(D)$.

The function dom assigns to each situation its domain of individuals, which is a subset of D.

(5) $dom(w) = D$

(6) $|\ |$ is a function from the set of relation symbols into $\bigcup_{n \in N} (\mathcal{P}(\Omega \times D^n) \times \mathcal{P}(\Omega \times D^n))$ such that if R is n-ary, then $|R| \in \mathcal{P}(\Omega \times D^n) \times \mathcal{P}(\Omega \times D^n)$.

$|R|^+$ is to be interpreted as the set of those n+1-tuples $\langle s, a_1, \dots, a_n \rangle$ such that $a_1,\dots,a_n$ bear the R-relation to each other in the situation s.

$|R|^-$ is to be interpreted as the set of those $n+1$-tuples $\langle s, a_1, \ldots, a_n \rangle$ such that $a_1, \ldots, a_n$ explicitly don't bear the R-relation to each other in the situation s.

(7) $(\{w\} \times D^n) \cap |R|^+ \cap |R|^- = \emptyset$ and
$(\{w\} \times D^n) \subseteq (|R|^+ \cup |R|^-)$.

This corresponds to requirement (4) of 1.2.1.

(8) If R is n,m-ary and $\langle s, a_1, \ldots, a_m, \ldots, a_n \rangle$ is a member of $|R|^+ \cup |R|^-$, then $\{a_1, \ldots, a_m\} \subseteq \mathrm{dom}(s)$.

In an n,m-ary relation symbol, the first m argument places are reserved for agents which are to be present for the truth or falsity of an instance of this relation to be determined on the basis of a situation.

A variable assignment for a structure $\mathfrak{A}$ we define to be a function from the set of variables into the set of individuals in $\mathfrak{A}$

2.2.2. Definition of the Interpretation $\|A\|_{\mathfrak{A},g}$.

For a given structure $\mathfrak{A} = \langle \Omega, w, D, \mathrm{dom}, |\;| \rangle$ and a variable assignment g, the interpretation function $\|\;\|_{\mathfrak{A},g}$ is defined by induction in the following way:

Basis:

$$\|R(x_1, \ldots, x_n)\|^+_{\mathfrak{A},g} = \{s \mid \langle s, g(x_1), \ldots, g(x_n) \rangle \in |R|^+\}$$

$$\|R(x_1, \ldots, x_n)\|^-_{\mathfrak{A},g} = \{s \mid \langle s, g(x_1), \ldots, g(x_n) \rangle \in |R|^-\}$$

Induction step: For $\wedge$, $\vee$, $\neg$ and $\Rightarrow$, the induction step is defined as in 1.2.2.

$\|\forall x A\|^+_{\mathfrak{A},g}$ is the set of those $s \in \Omega$ such that $s \in \|A\|^+_{\mathfrak{A},g'}$ for every g' which differs from g only in the value at x.

$\|\exists x A\|^{+}_{\mathfrak{A},g}$ is the set of those $s \in \Omega$ such that $s \in \|A\|^{+}_{\mathfrak{A},g'}$ for some g' which differs from g only in the value at x.

$\|ax A\|^{+}_{\mathfrak{A},g}$ is the set of those $s \in \Omega$ such that $s \in \|A\|^{+}_{\mathfrak{A},g'}$ for every g' which differs from g only in the value at x, and for which $g'(x) \in dom(s)$.

$\|ex A\|^{+}_{\mathfrak{A},g}$ is the set of those $s \in \Omega$ such that $s \in \|A\|^{+}_{\mathfrak{A},g'}$ for some g' which differs from g only in the value at x, and for which $g'(x) \in dom(s)$.

$\exists xA$ is defined as $\neg\forall x\neg A$

exA is defined as $\neg ax\neg A$.

2.2.3. Truth and Validity.

A formula A is true in $\mathfrak{A}$ with respect to g if $w_{\mathfrak{A}} \in \|A\|^{+}_{\mathfrak{A},g}$.

A formula is false if its negation is true. It is easily verified that every formula is either true or false, and not both.

A formula is valid if it is true in every stucture with respect to every variable assignment.

2.3. Axiomatization.

The following is a complete axiomatic basis for the set of valid formulae of L(SQ).

Axioms All substitution instances of valid L(S) formulae are axioms. Since there is a decision procedure for L(S), it is effectively decideable whether a formula is an axiom of this kind. Also, all formulae represented by the following schemes are axioms: ($A \Leftrightarrow B$ is an abbreviation of $(A \Rightarrow B) \wedge (B \Rightarrow A)$. $A(y/x)$ is the formula obtained from A by substituting y for every free occurrence of x.)

(Ax1)	$\forall x(A \Rightarrow B) \supset (\exists x A \Rightarrow B)$	x is not free in B
(Ax2)	$\forall x(B \Rightarrow A) \supset (B \Rightarrow \forall x A)$	
(Ax3)	$\forall x(B \vee A) \Rightarrow (B \vee \forall x A)$	
(Ax4)	$((\exists x A) \wedge B) \Rightarrow \exists x(A \wedge B)$	
(Ax5)	$\forall x A \Rightarrow A(y/x)$	in A, x does not occur within the scope of any quantifier binding y
(Ax6)	$A(y/x) \Rightarrow \exists x A$	
(Ax7)	$\neg \forall x A \Rightarrow \exists x \neg A$	
(Ax8)	$ax(B \vee A) \Rightarrow (B \vee axA)$	x is not free in B
(Ax9)	$((exA) \wedge B) \Rightarrow ex(A \wedge B)$	
(Ax10)	$((axA) \wedge C) \Rightarrow A(y/x)$	In A, x does not occur within the scope of any quantifier binding y
(Ax11)	$(A(y/x) \wedge C) \Rightarrow exA$	

In Ax10 and Ax11, C is the formula $R(z_1,\dots,z_n)$ or the formula $\neg R(z_1,\dots,z_n)$, with R n,m-ary and with y among $z_1,\dots,z_m$.

(Ax12) $\neg axA \Leftrightarrow ex\neg A$

Rules of inference:

<u>Modus Ponens</u> $\vdash A$ and $\vdash A \supset B$ yield $\vdash B$.

<u>Generalization</u> $\vdash A$ yields $\vdash \forall x A$.

<u>a-introduction</u> $\vdash A \supset ((B \wedge P(x)) \Rightarrow C)$ yields $\vdash A \supset (B \Rightarrow axC)$ when A, B and C do not contain the unary relation symbol P, and x is not free in A or B.

<u>e-introduction</u> $\vdash A \supset ((C \wedge P(x) \Rightarrow B)$ yields $\vdash A \supset (exC \Rightarrow B)$ when A, B and C do not contain the unary relation symbol P, and x is not free in A or B.

a-introduction and e-introduction are not rules of inference in the sence that the consequent is true in $\mathfrak{A}$ if the premiss is. They merely state that if the premiss is true in all structures, then so is the consequent.

2.3.1. Theorem.

Every theorem of L(SQ) is valid in L(SQ).

Proof: The axioms are clearly valid, and the first two rules of inference are easily seen to preserve validity. We show this to be the case also for a-introduction. The proof for e-introduction is similar.

Suppose $A \supset (B \Rightarrow axC)$ is false in $\mathfrak{A},g$. Then $(B \Rightarrow axC)$ is false, and so there is an $s \in \|B\|^+_{\mathfrak{A},g}$ such that $s \notin \|axC\|^+_{\mathfrak{A},g}$. By the interpretation rule for a, $s \notin \|C\|^+_{\mathfrak{A},g'}$ for a g' differing with g at most at x, and for which $g'(x) \in dom(s)$. Let $\mathfrak{A}'$ be a structure identical to $\mathfrak{A}$, except that P denotes a "presence predicate", i.e. $|P|^+ = \{\langle s,a \rangle \mid a \in dom(s)\}$. Hence $s \in \|P(x)\|^+_{\mathfrak{A}',g'}$. Since neither of A, B or C contains P, a shift from $\mathfrak{A}$ to $\mathfrak{A}'$ does not alter the interpretation of any of these formulae. Also, A and B do not contain x, so a shift from g to g' does not alter the interpretation of these formulae. Hence $s \in \|B\|^+_{\mathfrak{A}',g'}$, $s \in \|P(x)\|^+_{\mathfrak{A}',g'}$ and $s \notin \|C\|^+_{\mathfrak{A}',g'}$. This means that $(B \wedge P(x)) \Rightarrow C$ is false in $\mathfrak{A}',g'$. Furthermore, A is true, so $\mathfrak{A}',g'$ constitutes a counter-example to $A \supset ((B \wedge P(x) \Rightarrow C)$.

2.3.2. Derived Rules of Inference.

2.3.2.1. L(S) Entailment.

If $A'_1,\ldots,A'_n$ and B' are substitution instances (by the same substitutions) of the L(S) formulae $A_1,\ldots,A_n$ and B, and $A_1,\ldots,A_n \vdash_{L(S)} B$, then $A'_1,\ldots,A'_n \vdash_{L(SQ)} B'$

Proof: This follows immediately from the deduction theorem of L(S), the fact that every substitution instance of a valid L(S) formula is an axiom, and the rule of Modus Ponens.

2.3.2.2. Simplified a- and e-Introduction.

The rules we obtain by deleting the "A" in the rules of a- and e-introduction, are easily seen to follow from these rules. We will sometimes refer to these simpler rules also as a- and e-introduction.

2.4. Definition.

Γ is consistent if $\neg(A_1 \wedge \ldots \wedge A_n)$ is a theorem for no finite subset $\{A_1, \ldots, A_n\}$ of Γ.

2.5. Theorem (Completeness).

Every consistent set of formulae in L(SQ) has a model.

The central ideas of the proof below are borrowed from the completeness proof of Modal LPC in [Hughes & Cresswell 1972]. Central to the proof is the notion of E-formulae:

2.5.1. Definition of E-formulae.

Formulae of the following kinds are called E-formulae: (P is 1,1-ary and there are no occurrence of P or free occurrences of y other han those displayed. We also assume that y does not originally occur in C.)

(i) $\exists xA \supset A(y/x)$

(ii) $(((A \wedge \forall xC) \Rightarrow B) \wedge (A \Rightarrow (B \vee C(y/x)))) \supset (A \Rightarrow B)$

(iii) $(((C(y/x) \wedge A) \Rightarrow B) \wedge (A \Rightarrow ((\exists xC) \vee B))) \supset (A \Rightarrow B)$

(iv) $(((A \wedge axC) \Rightarrow B) \wedge ((A \wedge P(y)) \Rightarrow (B \vee C(y/x)))) \supset (A \Rightarrow B)$

(v) $(((C(y/x) \wedge A \wedge P(y)) \Rightarrow B) \wedge (A \Rightarrow ((exC) \vee B))) \supset (A \Rightarrow B)$

They are said to be E-formulae with respect to the variable which is here represented by y, and (for categories iv and v) the 1,1-ary relation symbol which is here represented by P.

Two E-formulae which are identical, except that they are E-formulae with respect to different variables or relation symbols, are said to be of the same E-form.

2.5.2. Lemma.

If Π is a consistent set of formulae, then so is $\Pi \cup \{E\}$ for any E-formula E with respect to a variable and (for categories iv and v) a 1,1-ary relation symbol which do not occur in any formula of Π.

Proof: Suppose (for reductio ad adsurdum) that $\Pi \cup \{E\}$ is inconsistent. This means that $H \supset \neg E$ is a theorem, where H is the conjunction of some formulae in Π. Furhter suppose E is of type iv, i.e. E is the formula

$$(((A \wedge axC) \Rightarrow B) \wedge ((A \wedge P(y)) \Rightarrow (B \vee C(y/x)))) \supset (A \Rightarrow B)$$

Since $\vdash H \supset \neg E$, we must have

$\vdash H \supset \neg(A \Rightarrow B)$,

$\vdash H \supset ((A \wedge axC) \Rightarrow B)$

and $\vdash H \supset ((A \wedge P(y)) \Rightarrow (B \vee C(y/x)))$

We want to show that the last entails

$$\vdash H \supset (A \Rightarrow (B \vee axC))$$

This is proved in the following manner:

(1) $H \supset ((A \wedge P(y)) \Rightarrow (B \vee C(y/x)))$ theorem by the assumption

(2) $H \supset (A \Rightarrow ay(B \vee C(y/x)))$ 1, a-introd.

(3)	$H \supset (A \Rightarrow (B \vee ayC(y/x)))$	2, Ax8, L(S) ent.
(4)	$(ayC(y/x) \wedge P(x)) \Rightarrow C$	Ax10
(5)	$ayC(y/x) \Rightarrow axC$	4, a-intro.
(6)	$H \supset (A \Rightarrow (B \vee axC))$	3, 5, L(S) ent.

From $\vdash H \supset \neg(A \Rightarrow B)$

and $\vdash H \supset ((A \wedge axC) \Rightarrow B)$

and $\vdash H \supset (A \Rightarrow (B \vee axC))$,

$\vdash \neg H$ follows by the rule of L(S) entailment.

Hence Π is inconsistent, contrary to the assumption. This means that $\Pi \cup \{E\}$ must be consistent when E is of type iv. The proofs for E-formulae of types i-iii and v are similar.

2.5.3. Lemma.

Every consistent set of formulae (in an L(SQ) language $\mathcal{L}$) can be extended to a maximal consistent set (in an L(SQ) language $\mathcal{L}'$ which extends $\mathcal{L}$) which contains at least one E-formula of each E-form (in $\mathcal{L}'$).

Proof: Let there be given a consistent set Γ of formulae. We extend the language by an infinite list of 1,1-ary relation symbols, and an infinite list of variables. Since the original language did contain infinitely many of both these sorts of symbols, Γ is consistent also in this language. We now want to extend Γ to a consistent set which contains at least one formula from each E-form of the new language. There are just countably many E-forms, hence they can be given an enumeration.

We now define a corresponding sequence $\langle A_n \rangle_n$ of E-formulae in the following way:

Infinitely many variables do not occur in Γ, and only finitely many occur in $\{A_m\}_{m\leq n}$. Similar remarks hold for 1,1-ary relation symbols. Hence there is a variable y and a 1,1-ary relation symbol P which do not occur in Γ or $\{A_m\}_{m\leq 1}$. Let A_{n+1} be the E-formula of form n+1 with respect to y (and, in cases iv and v, to P). Since Γ is consistent, it now follows by induction and use of lemma 2.5.2, that $\Gamma \cup \{A_m\}_{m\leq n}$ is consistent for all n. Hence $\Gamma \cup \{A_m\}_{m\in\mathbb{N}}$ is also consistent.

In order to obtain the set prescribed by the lemma, we now only need to extend $\Gamma \cup \{A_m\}_{m\in\mathbb{N}}$ to a maximal consistent set. For this, the standard procedure used for consistent sets of first order predicate logic is applicable.

2.5.4. Definitions.

An atomary formula is of the type $R(x_1,\ldots,x_n)$.

A quasi-atomary formula is either a proper atomary formula, or it is of one of the types A=>B, $\forall$xA or axA.

Hence all formulae are built up from quasi-atomary formulae using truth-functional connectives only.

2.5.5. Lemma.

Every maximal consistent set Δ which contains at least one formula for each E-form, has a model.

2.5.5.1. Construction of the Model $\mathfrak{A}$.

Let V be the set of four element valuations which treat quasi-atomary formulae as atomary. We only demand of elements in V that they conform with truth tables 6 and 7 of 1.1.2. As an example, no requirements are posed concerning the relationship between $v(\forall xA)$ and $v(A)$. On V we may define a topology in the same manner as described in 1.10.1.

For each fomula A, we define F_A to be the set $\{v \in V \mid v(A) \in \{1, \rlap{1}{\downarrow}\}\}$. F_A is then both open and closed. The subset V' of V is defined as follows:

$$V' = \bigcap_{(A \Rightarrow B) \in \Delta} (\tilde{F}_A \cup F_B).$$

V' is the intersection of closed sets, hence it is closed itself. Ω_0 is the subset of those $v \in V'$ with the following property: For every formula A and every variable x:

- if $v(\exists xA) \in \{1, \rlap{1}{\downarrow}\}$, there is a y for which $v(A(y/x)) \in \{1, \rlap{1}{\downarrow}\}$
- if $v(\forall xA) \notin \{1, \rlap{1}{\downarrow}\}$, there is a y for which $v(A(y/x)) \notin \{1, \rlap{1}{\downarrow}\}$
- if $v(\text{e}xA) \in \{1, \rlap{1}{\downarrow}\}$, there is a y and a 1,1-ary P for which $v(P(y)) \in \{1, 0, \rlap{1}{\downarrow}\}$ and $v(A(y/x)) \in \{1, \rlap{1}{\downarrow}\}$
- if $v(\text{a}xA) \notin \{1, \rlap{1}{\downarrow}\}$, there is a y and a 1,1-ary P for which $v(P(y)) \in \{1, 0, \rlap{1}{\downarrow}\}$ and $v(A(y/x)) \notin \{1, \rlap{1}{\downarrow}\}$.

Finally, to obtain the set Ω, we delete from Ω_0 the two constant valuations on ? and $\rlap{1}{\downarrow}$ respectively.
The structure $\mathfrak{A}$ is defined as follows:

(1) Δ defines a two element valuation w ($w(A) = 1$ iff $A \in \Delta$). We let this w be the distinguished element.

(2) The domain of situations is $\Omega \cup \{w\}$. (Ω is the Ω above.) We later show that $w \in \Omega$.

(3) The domain of individuals is the set of variables in the language.

(4) $\mathrm{dom}(v) = \{x \mid$ there is a n,m-ary relation symbol R and a m-tuple $\langle z_1,\ldots,z_m\rangle$ which contains x, such that $v(R(z_1,\ldots,z_m,y_1,\ldots,y_{n-m}))$ is in $\{0,1,\updownarrow\}$ for some $\langle y_1,\ldots,y_{n-m}\rangle\}$.

(5) $|R|^+ = \{\langle v,x_1,\ldots,x_n\rangle \mid v(R(x_1,\ldots,x_n))\in\{1,\updownarrow\}\}$

$|R|^- = \{\langle v,x_1,\ldots,x_n\rangle \mid v(R(x_1,\ldots,x_n))\in\{0,\updownarrow\}\}$.

Clearly, $\mathfrak{A}$ satisfies all requirements of 2.2.1. The assignment g is defined as the identity map on the set of variables.

<u>2.5.5.2. Proposition.</u>

If $v\in\Omega$, $(G\Rightarrow H)\in\Delta$ and $v(G)\in\{1,\updownarrow\}$, then $v(H)\in\{1,\updownarrow\}$. This follows immediately from the construction of Ω.

<u>2.5.5.3. Proposition.</u>

If $(\neg(G\Rightarrow H))\in\Delta$, then there is a $v\in V'$ such that $v(G)\in\{1,\updownarrow\}$ and $v(H)\notin\{1,\updownarrow\}$.

Proof: Suppose there is no $v\in V'$ such that $v(G)\in\{1,\updownarrow\}$ and $v(H)\notin\{1,\updownarrow\}$. By the definition of V', this means that

$$F_G \cap \tilde{F}_H \cap \Big(\bigcap_{(A\Rightarrow B)\in\Delta} (\tilde{F}_A \cup F_B)\Big) = \emptyset.$$

By topological compactness, there is a finite subset $\{A_i\Rightarrow B_i\}_{i\leqslant n}$ of Δ such that

$$F_G \cap \tilde{F}_H \cap \Big(\bigcap_{i\leqslant n} (\tilde{F}_{A_i} \cup F_{B_i})\Big) = \emptyset.$$

But then

$$\Big(\bigwedge_{i\leqslant n} (A_i\Rightarrow B_i)\Big) \supset (G\Rightarrow H)$$

is a substitution instance of a valid $L(S)$ formula, and by consistency of Δ, $(\neg(G\Rightarrow H))\notin\Delta$.

2.5.5.4. Proposition.

If $(\neg(G_0 \Rightarrow H_0)) \in \Delta$, then there is a $v \in \Omega$ such that $v(G_0) \in \{1, \rlap{1}\vee\}$ and $v(H_0) \notin \{1, \rlap{1}\vee\}$.

Proof: Suppose $(\neg(G_0 \Rightarrow H_0)) \in \Delta$. The set of formulae of the types $\exists xA$, exA, $\forall xA$ and axA can be enumerated. Suppose the first is exA. From the construction of Δ, we know there is a variable y and a 1,1-ary relation symbol P such that

$$(((A(y/x) \wedge G_0 \wedge P(y)) \Rightarrow H_0) \wedge (G_0 \Rightarrow ((\mathrm{ex}A) \vee H_0))) \supset (G_0 \Rightarrow H_0)$$

is contained in Δ. By maximal consistency of Δ,

$$(\neg((A(y/x) \wedge G_0 \wedge P(y)) \Rightarrow H_0)) \in \Delta$$

$$\text{or} \quad (\neg(G_0 \Rightarrow ((\mathrm{ex}A) \vee H_0))) \in \Delta.$$

We choose one, and call it $\neg(G_1 \Rightarrow H_1)$. Since $(\neg(G_1 \Rightarrow H_1)) \in \Delta$, we have $F_{G_1} \cap \tilde{F}_{H_1} \cap V' \neq \emptyset$ by 2.5.5.3. We also see that

$(F_{G_1} \cap \tilde{F}_{H_1}) \subseteq (F_{G_0} \cap \tilde{F}_{H_0})$.

The procedures for $\forall$, $\exists$, and a are similar. Also for these, suitable E-formulae are present in Δ. This process is repeated by induction. At step $n+1$ we form

$$\neg(G_{n+1} \Rightarrow H_{n+1}) \qquad \text{from} \qquad \neg(G_n \Rightarrow H_n)$$

and the $n+1$'th formula in the $(\exists xA$, exA, $\forall xA$, axA$)$-sequence in the same way as

$$\neg(G_1 \Rightarrow H_1) \text{ was formed from } \neg(G_0 \Rightarrow H_0)$$

and the first element in the $(\exists xA$, exA, $\forall xA$, axA$)$-sequence.

For all n we will have

$$F_{G_n} \cap \tilde{F}_{H_n} \cap V' \neq \emptyset$$

and $$(F_{G_{n+1}} \cap \tilde{F}_{H_{n+1}}) \subseteq (F_{G_n} \cap \tilde{F}_{H_n}).$$

By compactness

$$(\bigcap_n (F_{G_n} \cap \tilde{F}_{H_n})) \cap V' \neq \emptyset$$

follows.

It follows from the construction of $\{\neg(G_n \Rightarrow H_n)\}_n$ that

$$(\bigcap_n (F_{G_n} \cap \tilde{F}_{H_n})) \cap V' \subseteq \Omega_0.$$

Together, these results yield

$$(\bigcap_n (F_{G_n} \cap \tilde{F}_{H_n})) \cap \Omega_0 \neq \emptyset.$$

Clearly, no element of $\Omega_0 \setminus \Omega$ is contained in $F_{G_n} \cap \tilde{F}_{H_n}$. Hence also

$$(\bigcap_n (F_{G_n} \cap \tilde{F}_{H_n})) \cap \Omega \neq \emptyset$$

follows.

Thus $F_{G_0} \cap \tilde{F}_{H_0} \cap \Omega \neq \emptyset$, and there is a $v \in \Omega$ for which $v(G_0) \in \{1, \mathfrak{V}\}$ and $v(H_0) \notin \{1, \mathfrak{V}\}$.

2.5.5.5. Proposition.

$F_A \cap \Omega \subseteq F_B \cap \Omega \neq \emptyset$ if and only if $(A \Rightarrow B) \in \Delta$.

This follows from 2.5.5.2 and 2.5.5.4.

2.5.5.6. Proposition.

$w \in \Omega$.

Proof: Suppose $w \notin V'$. Then there are formulae A and B such that $w \in F_A \cap \tilde{F}_B$ and $(A \Rightarrow B) \in \Delta$. But $(A \Rightarrow B) \supset (A \supset B)$ is a theorem, and by maximal consistence $(A \supset B) \in \Delta$, i.e. $w(A \supset B) = 1$, contradicting $w \in F_A \cap \tilde{F}_B$. Hence $w \in V'$.

In order to show that this w is also in Ω_0, suppose $w(\exists xA) = 1$. There is an y such that $((\exists xA) \supset A(y/x)) \in \Delta$, and hence $w(A(y/x)) = 1$. $exA \supset \exists xA$ is easily seen to be a theorem of $L(SQ)$. Hence also $w(exA) = 1 \longrightarrow w(A(y/x)) = 1$. For w the two other requirements for elements of Ω_0 are reducible to the two just considered, since w is a two element valuation. Hence we have shown $w \in \Omega_0$. Since $w(p) \in \{0,1\}$ for all p, $w \in \Omega$ immediately follows.

2.5.5.7. Proposition.

$\|A\|_{\mathfrak{A},g,v} = v(A)$ for all $v \in \Omega$.

The proof is by induction on the construction of formulae.

Basis: For atomary formulae this follows immediately from the definition of $|\ |$.

Induction step: For truth-functional connectives, the induction step follows from the correspondence between the valuations and the inductive definition of $\|\ \|$.

=>: We first show that $v(A \Rightarrow B) \in \{0,1\}$ for all $v \in \Omega$. There must be some formula C for which $v(C) \in \{1, \lightning\}$ (v is not the constant valuation on ?. Hence $v(C') \in \{1, \lightning, 0\}$ for some C', and so $v(C' \vee \neg C') \in \{1, \lightning\}$). Since $C \Rightarrow ((A \Rightarrow B) \vee \neg(A \Rightarrow B))$ is a theorem, $v(A \Rightarrow B) \in \{1, \lightning, 0\}$ by 2.5.5.5. There is also some formula D for which $v(D) \in \{0, ?\}$. $((A \Rightarrow B) \wedge \neg(A \Rightarrow B)) \Rightarrow D$ is a theorem, so $v(A \Rightarrow B) \in \{1, ?, 0\}$. Combining these results, we obtain $v(A \Rightarrow B) \in \{0,1\}$. Also, $(A \Rightarrow B) \supset (C \Rightarrow (A \Rightarrow B))$ and $(\neg(A \Rightarrow B)) \supset (C \Rightarrow \neg(A \Rightarrow B))$ are theorems, so $v(A \Rightarrow B) = w(A \Rightarrow B)$. Hence we only need to show that $w \in \|A \Rightarrow B\|^+$ iff $(A \Rightarrow B) \in \Delta$. Now $w \in \|A \Rightarrow B\|^+$ iff $\|A\|^+ \subseteq \|B\|^+$. By the induction hypothesis, $\|A\|^+ \subseteq \|B\|^+$ iff $(F_A \cap \Omega) \subseteq (F_B \cap \Omega)$. But $(F_A \cap \Omega) \subseteq (F_B \cap \Omega)$ iff $(A{=}{>}B) \in \Delta$.

Hence the induction step for => is proved.

a: Suppose $v(axA) \notin \{1,\updownarrow\}$. Since $v \in \Omega$, and hence satisfies the requirements of 2.5.5.1, there is a y and a 1,1-ary relation symbol P such that $v(P(y)) \in \{1,0,\updownarrow\}$ and $v(A(y/x)) \notin \{1,\updownarrow\}$. By the induction hypothesis, this means that $\|A(y/x)\|_{\mathfrak{A},g,v} \notin \{1,\updownarrow\}$. Then $\|A\|_{\mathfrak{A},g',v} \notin \{1,\updownarrow\}$ for the g' identical to g, except that $g'(x) = y$. Since $y \in dom(v)$, $\|axA\|_{\mathfrak{A},g,v} \notin \{1,\updownarrow\}$.

Now suppose $v(axA) \in \{1,\updownarrow\}$. Further suppose $y \in dom(v)$, i.e. that $v(R(z_1,\ldots,z_n)) \in \{1,\updownarrow,0\}$ where y is among $z_1,\ldots,z_m$ and R is n,m-ary. Since $((axA) \wedge R(z_1,\ldots,z_n)) {=}{>} A(y/x)$ and $((axA) \wedge \neg R(z_1,\ldots,z_n)) {=}{>} A(y/x)$ are contained in Δ, $v(A(y/x)) \in \{1,\updownarrow\}$. By the induction hypothesis, $\|A(y/x)\|_{\mathfrak{A},g,v} \in \{1,\updownarrow\}$. This means that $\|A\|_{\mathfrak{A},g',v} \in \{1,\updownarrow\}$ for all g' differing with g at most at x, and for which $g'(x) \in dom(v)$. Or, in other words, $\|axA\|_{\mathfrak{A},g,v} \in \{1,\updownarrow\}$. Hence we have shown that $v(axA) \in \{1,\updownarrow\}$ iff $\|axA\|_{\mathfrak{A},g,v} \in \{1,\updownarrow\}$. By analogous arguments, it follows that $v(exB) \in \{1,\updownarrow\}$ iff $\|exB\|_{\mathfrak{A},g,v} \in \{1,\updownarrow\}$. And hence, by Ax12, $v(axA) \in \{0,\updownarrow\}$ iff $\|axA\|_{\mathfrak{A},g,v} \in \{0,\updownarrow\}$. Together, these results yield

$$\|axA\|_{\mathfrak{A},g,v} = v(axA).$$

The induction step for $\forall$ is similar. This concludes the proof of 2.5.5.7.

2.5.5.8. Proposition.

$\mathfrak{A}$ is a model for Δ.

Proof: By 2.5.5.6 and 2.5.5.7, $\|A\|_{\mathfrak{A},g,w} = w(A)$, i.e. $w \in \|A\|^{+}_{\mathfrak{A},g}$ iff $A \in \Delta$.

2.6. A Remark on a- and e-Introduction.

2.6.1. Plain a- and e-Introduction.

The rule $\vdash A \supset (B \Rightarrow C)$ / $\vdash A \supset (B \Rightarrow axC)$ x is not free in A or B

we call plain a-introduction. The analogous simlification of e-introduction we call plain e-introduction. They are simpler than a- and e-introduction, and do also preserve validity. They are not chosen, however, since they would be too weak. There are cases for which $A \supset ((B \wedge P(x)) \Rightarrow C)$ is a teorem, but not $A \supset (B \Rightarrow C)$, even when P occurs in neither of A, B or C. This is so because of Ax10 and Ax11. We want to show that a- and e-introduction cannot be replaced by their "plain" counterparts, and define L_0 to be the system in which this replacement is made.

2.6.2. Generalized Semantics.

In the generalized semantics for L(SQ) dom is a function of two entities. One coordinate is a situation, the other an individual variable. For each variable x and situation s, dom(s,x) must satisfy all requirements of dom(s) in an ordinary structure. axA is true in a situation s iff all $a \in$ dom(s,x) have the property of A. (For a formal definition, substitute dom(s,x) for dom(s) in the definitions of $\|axA\|^{+}_{\mathfrak{A},g}$ and $\|exA\|^{+}_{\mathfrak{A},g}$ in 2.2.2.)

It should be noted that dom is a function of the variable x and not its interpretation. Thus ax really binds the variable x, since dom(s,x) is independent of the interpretation of x.

2.6.3. Lemma.

(axR(x))=>(ayR(y)) is not a valid formula of L(SQ) with the generalized semantics.

Proof: If $\langle s,a\rangle \in |R|^+$ for all $a \in dom(s,x)$, but $\langle s,b\rangle \notin |R|^+$ for a $b \in dom(s,y)$, then $\| axR(x) \|^+ \nsubseteq \| ayR(y) \|^+$.

2.6.4. Lemma.

Every theorem of L_0 is valid in the generalized semantics for L(SQ).

Proof: All axioms of L(SQ) are easily checked to be valid. Modus Ponens and Generalization clearly preserve validity. In order to show that plain a-introduction preserves validity, suppose that $A \supset (B \Rightarrow axC)$ is false in the structure $\mathfrak{A}, g$. Then $w \in \| A \|^+_{\mathfrak{A},g}$, $s \in \| B \|^+_{\mathfrak{A},g}$ and $s \notin \| axC \|^+_{\mathfrak{A},g}$ for a situation s in the structure. By the definition of $\| axC \|^+_{\mathfrak{A},g}$, $s \notin \| C \|^+_{\mathfrak{A},g'}$ for a g' differing with g at most at x. Since neither A nor B contains x, $A \supset (B \Rightarrow C)$ must be false in $\mathfrak{A}, g'$.
The proof for plain e-introduction is similar.

2.6.5. Theorem.

L_0 is not a complete axiomatization of L(SQ).

Proof: This follows from 2.6.3, 2.6.4 and the fact that (axR(x))=>(ayR(y)) is a valid formula of L(SQ).

3. APPLICATION TO NAKED INFINITIVE PERCEPTION REPORTS

3.1. Preliminary remarks

Much of Barwise's early work on Situation Semantics emerged from his study of perception reports, cfr. [Barwise 1979] [Barwise 1981]. The "situational" ideas on perception outlined there have been incorporated in the theory of [Barwise & Perry 1983].

We will now try to formulate some of these notions within the framework developed so far in this article. Due to the comparative simplicity of L(S) and L(SQ), only a rough approximation can be expected.

In the following we will assume that the reader is acquainted with [B. & P. 83]. From the example on page 186, the following rule for naked infinitive seeing can be inferred. (Barwise and Perry do not in fact state this as a general rule, and express the need for caution when more complex sentences are embedded. However, we try it here as an approximation.):

> $d,c \, \| \text{SEES } \phi \| \, a,e$
> iff there is a c.o.e. e' and a location
> l = c(SEES) temporally overlapping that
> of the discourse situation such that
> $d,c \, \| \phi \| \, l,e'$
> In e: at l: seeing,a,e'; yes

Now everything considered in this article is insensitive towards connections and discourse situations. Also, there is nothing corresponding to locations. This leads us to the following simplification:

> $\| \text{SEES } \phi \| \, a,e$
> iff there is an e' such that
> $\| \phi \| \, e'$
> In e: seeing,a,e'; yes

Finally we drop the individual coordinate in the seeing relation. This can be interpreted as a restriction of our study to the perceptions of one specific individual. This simplification is just a convenience and not really necessary. But it is also of little consequence and can easily be amended. If we also switch from "$\|\phi\| e$" to "$e \in \|\phi\|^+$", we now obtain:

$e \in \| \text{SEES } \phi \|^+$

iff there is an $e' \in \|\phi\|^+$

such that In e: seeing,e';yes

This is what we intend to represent within the setting of L(S).

Now the scheme "In e: seeing,e';yes" is one which cannot possibly be defined in L(S), since the situations of this system are primitives and not sets. However, we can get something very similar by so to speak lifting the seeing relation out of the situations, and define a

$\Sigma^+ \subseteq \Omega \times \Omega$

such that $\langle e,e' \rangle \in \Sigma^+$ corresponds to In e: seeing,e';yes

The interpretation of $\neg$SEES ϕ is something which Barwise and Perry say nothing about. The following I feel is a plausible suggestion:

$e \in \| \text{SEES } \phi \|^-$

iff In e: seeing,e';no for every e' (possibly within some further specified set) such that

$e' \in \|\phi\|^+$

Now

In e: seeing e';no is something different from

not(In e: seeing e';yes)

Hence we need one more relation Σ^- such that $\langle e,e'\rangle \in \Sigma^-$ corresponds to

In e: seeing, e': no

3.2. Syntax and semantics of L(See*)

Symbols

propositional variables	$p, q, p_1, q_1, \ldots$
connectives	$\neg, \vee, \wedge, \Rightarrow$, See
auxiliary symbols	(,)

Formation rules

(1) Every propositional variable is a formula.

(2) If A and B are formulae, then so are $(A \Rightarrow B)$, $(A \vee B)$, $(A \wedge B)$ and $\neg A$.

(3) If A is a formula without occurrences of See or $\Rightarrow$, then See(A) is a formula.

A structure for L(See*) is a quadruple $\langle \Omega, w, |\,|, \Sigma \rangle$

(1) Ω is the set of situations. These are primitives.

(2) w is a distinguished element of Ω

(3) $|\,|$ is a function from the set of propositional variables into $\mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$

(4) $w \in |p|^+ \cup |p|^-$ and $w \notin |p|^+ \cap |p|^-$

(5) Σ is an ordered pair $\langle \Sigma^+, \Sigma^- \rangle$ with both elements being sets of ordered pairs of situations.

(6) $(\{w\} \times \Omega) \cap \Sigma^+ \cap \Sigma^- = \emptyset$ and
$(\{w\} \times \Omega) \subseteq \Sigma^+ \cup \Sigma^-$

The last is just the requirement that for every t, $\langle w,t \rangle$ is a member of one, but not both, of Σ^+ and Σ^-. This corresponds to (4) and is necessary in order to ensure that See(A) is either true or false and not both.

For a given structure $\mathfrak{A}=\langle\Omega,w,|\ |,\Sigma\rangle$, the interpretation function $\|\ \|_{\mathfrak{A}}$ is defined exactly as for L(S) on pages II.6 and II.7, with the following addition:

(vii) $s\in\|See(A)\|^{+}$ iff there is a $t\in\Omega$ such that $\langle s,t\rangle\in\Sigma^{+}$ and $t\in\|A\|^{+}$.

(viii) $s\in\|See(A)\|^{-}$ iff $\langle s,t\rangle\in\Sigma^{-}$ for all $t\in\|A\|^{+}$.

A is <u>true</u> in $\mathfrak{A}$ if and only if $w\in\|A\|^{+}_{\mathfrak{A}}$.

3.3. Lemma.

Every well-formed substitution instance in L(See *) of a valid L(S) formula is valid in L(See *).

Proof: Remark 1.4.2 is applicable.

3.4. Axiomatization.

A complete axiomatic characterization of the set of valid formulae of L(See *) is obtained by adding to the derivation rule and axiom schemes of L(S) the following axiom schemes:

(S1) $(A \Rightarrow B) \supset (See(A) \Rightarrow See(B))$

(S2) $(A \Rightarrow B) \supset (\neg See(B) \Rightarrow \neg See(A))$

(S3) $See(A \vee B) \supset (See(A) \vee See(B))$

(S4) $\neg(See(A) \vee See(B)) \Rightarrow \neg See(A \vee B)$

(S5) $See(A) \Rightarrow \neg(A \Rightarrow \neg(A \Rightarrow A))$

(S6) $(A \Rightarrow \neg(A \Rightarrow A)) \Rightarrow \neg See(A)$

3.5. Theorem (Validity).

Every theorem is true in every structure.

Proof: By lemma 3.3, all formulae indicated by the axiom schemes (A1)-(A21) are valid in L(See $\star$). Validity of S1-S6 follows immediately from the interpretation rules for the connectives.

<u>3.6. Theorem (Completeness).</u>

Every consistent set of formulae in L(See $\star$) has a model.

We prove this from the corresponding result for L(S):

<u>3.6.1. Construction of a model $\mathfrak{A}'$ for a consistent set Γ.</u>

Let Γ be a consistent set of formulae in L(See $\star$). We extend the language by a countably infinite set of new propositional variables, one for each simple formula of the original language. The propositional variable corresponding to the simple formula A is denoted by p_A. To obtain the set Γ', we substitute p_A for every occurrence of See(A) in formulae of Γ.

Π is the set of all formulae of the types

$$(A \Rightarrow B) \supset (p_A \Rightarrow p_B)$$

$$(A \Rightarrow B) \supset (\neg p_B \Rightarrow \neg p_A)$$

$$p_{(A \vee B)} \Rightarrow (p_A \vee p_B)$$

$$\neg(p_A \vee p_B) \Rightarrow \neg p_{(A \vee B)}$$

$$p_A \Rightarrow \neg(A \Rightarrow \neg(A \Rightarrow A))$$

$$(A \Rightarrow \neg(A \Rightarrow A)) \Rightarrow \neg p_A$$

Since Γ is consistent in L(See $\star$), $\Gamma' \cup \Pi$ must be consistent in L(S). $\Gamma' \cup \Pi$ then has a model $\mathfrak{A} = \langle \Omega, w, |\,| \rangle$ which is a closed valuation structure. In this structure we define an object $\Sigma = \langle \Sigma^+, \Sigma^- \rangle$ such that $\mathfrak{A}' = \langle \Omega, w, |\,|, \Sigma \rangle$ is an L(See $\star$) model for Γ:

$$\Sigma^+ = \{\langle s,t \rangle \mid t \notin \cup \{\| A \|^+ \mid s \notin \| p_A \|^+\}$$

$$\Sigma^- = \{\langle s,t \rangle \mid t \in \cup \{\| A \|^+ \mid s \in \| p_A \|^-\}.$$

3.6.2. Proposition.

$\mathfrak{A}'$ is an L(See *) structure.

Proof: We must check that (6) in 3.2 holds. Since $\mathfrak{A}$ is an L(S) structure, $w \notin \|p_A\|^+$ iff $w \in \|p_A\|^-$. Hence from the definition of Σ, $\langle w,t\rangle \in \Sigma^+$ iff $\langle w,t\rangle \notin \Sigma^-$, and this is just a reformulation of (6).

3.6.3. Proposition.

$\|See(A)\|^+ \subseteq \|p_A\|^+$.

Proof: Suppose $s \in \|See(A)\|^+$. Then $\langle s,t\rangle \in \Sigma^+$ for a $t \in \|A\|^+$. By definition of Σ^+, this implies that $s \in \|p_A\|^+$.

3.6.4. Proposition.

$\|p_A\|^+ \subseteq \|See(A)\|^+$.

Proof: Suppose $s \in \|p_A\|^+$, and suppose (for reductio ad absurdum) that $s \notin \|See(A)\|^+$, i.e. $\langle s,t\rangle \notin \Sigma^+$ for all $t \in \|A\|^+$. By the definition of Σ^+, this means that $\|A\|^+ \subseteq \bigcup \{\|B\|^+ | s \notin \|p_B\|^+\}$. Since $p_A \Rightarrow \neg(A \Rightarrow \neg(A \Rightarrow A))$ is true, $\|A\|^+ \neq \emptyset$. By compactness, $\|A\|^+ \subset (\|B_1\|^+ \cup \ldots \cup \|B_n\|^+) = \|B_1 \vee \ldots \vee B_n\|^+$ for a finite subset $\{B_1,\ldots,B_n\}$ of $\{B | s \notin \|p_B\|^+\}$. By the truth of

$(A \Rightarrow (B_1 \vee \ldots \vee B_n)) \supset (p_A \Rightarrow p_{(B_1 \vee \ldots \vee B_n)})$, $s \in \|p_{(B_1 \vee \ldots \vee B_n)}\|^+$ follows. But by repeated use of the true scheme $p_{(C_1 \vee C_2)} \Rightarrow (p_{C_1} \vee p_{C_2})$ this must lead to a contradiction, since $s \notin \|p_{B_i}\|^+$ for all i, $1 \leq i \leq n$.

3.6.5. Proposotion.

$\|p_A\| = \|See(A)\|$.

Proof: This follows from 3.6.3, 3.6.4 and corresponding results for the negative interpretations.

3.6.6. Proposition.

$\mathfrak{A}'$ is a model for Γ.

Proof: This follows immediately from 3.6.5.

3.7. Inspection of valid principles.

3.7.1. Veridicality.

L(See *) is clearly too weak, since it allows the possibility of seeing false situation; i.e.

$$\text{See}(A) \supset A$$

is not necessarily true. The addition of this as an extra axiom scheme corresponds to the following extra constraint on the structures:

If $\langle w,t\rangle \in \Sigma^+$ and $t\in|p|^+$ $(t\in|p|^-)$
then $w\in|p|^+$ $(w\in|p|^-)$.

In such a system, both principles

A: If b sees ϕ then ϕ

D: If b sees $\neg\phi$ then b doesn't see ϕ

of chapter 8 in [Barwise and Perry 1983] become valid.

The constraint above can be strengthened to the following:

(i) If $\langle s,t\rangle \in \Sigma^+$ and $t\in|p|^+$ $(t\in|p|^-)$ then $s\in|p|^+$ $(s\in|p|^-)$

(ii) If $s\in|p|^+$ and $t\in|p|^-$, then $\langle s,t\rangle, \langle t,s\rangle \in \Sigma^-$.

This corresponds to the two extra axiom schemes:

$$\mathrm{See}(A) \Rightarrow A \qquad \neg A \Rightarrow \neg \mathrm{See}(A).$$

3.7.2. Distribution Principles.

The principles

E: If b sees ϕ and ψ, then b sees ϕ and b sees ψ

F: If b sees ϕ or ψ, then b sees ϕ or b sees ψ

which correspond to

$$\mathrm{See}(\phi \wedge \psi) \supset (\mathrm{See}(\phi) \wedge \mathrm{See}(\psi))$$

$$\mathrm{See}(\phi \vee \psi) \supset (\mathrm{See}(\phi) \vee \mathrm{See}(\psi))$$

are both valid in L(See *). The first follows from S1, the second from S3.

3.7.3. Principles Concerning Identity and Quantification.

Barwise and Perry also put forward the following principles:

B: If b sees $\phi(t_1)$ and t_1 is t_2 then b sees $\phi(t_2)$

C: If b sees ϕ(the Π) then there is something$_1$ such that b sees $\phi(it_1)$

G: If b sees ϕ(a Π) then there is a Π_1 such that b sees $\phi(it_1)$.

None of these are expressible in L(See *). However the following very rough approximations:

$$\mathrm{See}(\phi(x_1)) \wedge x_1 = x_2) \supset \mathrm{See}(\phi(x_2))$$

$$\mathrm{See}(\exists x \phi) \supset \exists x\, \mathrm{See}(\phi)$$

are both expressible and true in the language L(See Q) obtained by combining L(SQ) and L(See *), and adding identity. (There are several alternatives for the definition of $\| x = y \|$, buth that does not concern us here as long as $w \in \| x = y \|^{+}_{\mathfrak{A}, g}$ iff $g(x) = g(y)$.)

The language L(See Q) was studied in my cand.scient. thesis, in particular, a complete axiomatization was given. This will be the topic of a further publication.

4. APPLICATION TO THE CONCEPT OF POSSIBILITY.

The definition of the M-operator in section 1.11 enabled us to investigate the connections between the formal systems L(S) and S5. But there is more than this to be said about the notion of possibility in situation semantics, even within the narrow frame of the present systems.

The concept of possibility represented in systems like S5, is the concept of philosophical possibility; "possibly A" means "A could have been true" or "A does not describe an impossible constellation of facts". However, in ordinary discourse another interpretation is intended at least as often. When our information is incomplete, we often use "possibly A" as a synonyme of "perhaps A is the case". In the terminology of situation semantics, such a usage may correspond approximately to

(1) "The present situation may be part of a situation in which A is the case."

Could such a notion be represented within possible world semantics? First note that $\neg A$ and Poss(A) could never be true at the same time. Then suppose Poss(A) is true in a possible world w. Since w defines a total valuation on the formulae, either A or $\neg A$ is true. It cannot be $\neg A$, hence A follows. If we also accept the principle $A \Rightarrow Poss(A)$, the result is that we cannot distinguish between A and Poss(A), and we have a collapse to the propositional or predicate calculus. This argument cannot be used against L(S), since

$$Poss(A) \Rightarrow (A \vee \neg A)$$

need not be true.

Before we try to represent Poss(A) (as defined in (1)) within L(S), we must decide how to understand "may be part of". One interpretation is our technical term subsituation (cfr. 1.11.3). "This situation may be part of t" should then be read as something like "Everything true in this situation is also true in t, so on the basis of this situation we cannot exclude the possibility that t is factual.".

This could be a viable approximation, but I fear it may turn out to be a little confining, for instance in the presence of the concept of a situation's associated domain of individuals. In order to grant ourselves a little flexibility, we therefore leave open the possibility that there can be something inherent in a situation which prevents some other situation from being a possible extension, independently of other facts in the two situations. Hence subsituation is a necessary, but not sufficient, condition.

Since we choose "possibly part of" to be a basic notion, we must add something to an ordinary L(S) structure in order to represent this notion. One candidate for this added feature is simply a relation Π^+ on the set of situations. In order to represent also "cannot be part of" and a gap in between, we need one more relation Π^-. Poss(A) is then defined to be true in s iff s may be extended to a situation t in which A is true; i.e. iff there is a $t \in \|A\|^+$ such that $\langle s,t\rangle \in \Pi^+$. Correspondingly, Poss(A) is false in s iff s cannot be extended to any situation t in which A is true; i.e. iff $\langle s,t\rangle \in \Pi^-$ for all $t \in \|A\|^+$.

By what we have said so far, we have defined a language L(Poss ⋆) which is isomorphic to L(See ⋆). Hence we have a complete axiomatization at hand. So far we have done nothing to ensure that s is a subsituation of t if $s\Pi^{+}t$. Hence we do exactly that, and add the structural constraint:

(2) If $s\Pi^{+}t$, then s is a subsituation of t.

A new axiomatization is then obtained by adding the axiom

$$\mathrm{Poss}(A) \wedge B \Rightarrow \mathrm{Poss}(A \wedge B).$$

A counterpart for the negative extension could be

(3) If $s \in |p|^{+}$ and $t \in |p|^{-}$ then $s\Pi^{-}t$ and $t\Pi^{-}s$ and the axiom

$$\neg A \Rightarrow \neg \mathrm{Poss}(A).$$

It could also be argued that $\neg\mathrm{Poss}(A)$ is the same as $\neg A$, so that

$\|\mathrm{Poss}(A)\|^{-}$ should be defined as $\|A\|^{-}$

and we could dispense with Π^{-}.

Other reasonable additions are the axiom $A \Rightarrow \mathrm{Poss}(A)$ and the requirement that Π^{+} be reflexive.

We note that by the isomorphy of L(Poss $\star$) to L(See $\star$),

$$\mathrm{Poss}(A) \Rightarrow \neg(A \Rightarrow \neg(A \Rightarrow A))$$

i.e. $\mathrm{Poss}(A) \Rightarrow M(A)$ is valid.

Several augmentations of the L(Poss $\star$) structures could be studied. An interesting step would be to allow formulae like Poss(Poss(A)), and to discuss which additions should be made to make valid the right formulae of this extended language.

It could also be argued that L(CS), rather than L(S), would provide the proper framework, since the truth of Poss(A) in a consistent situation should entail the truth of A in a consistent extension.

The formula

$$Poss(A) \Rightarrow (Poss(A \wedge B) \vee Poss(A \wedge \neg B))$$

could also be of interest, perhaps even $L(CS)^+$ would be the best framework for a discussion of this formula.

Note.

1 This article was based on excerpts from my cand.scient. thesis of the same title, which was written at the University of Oslo under the supervision of Professor Jens Erik Fenstad. I am indebted to him for encouragements and valuable advice during the work on my thesis and in the composition of this article.

REFERENCES

[Anderson and Belnap 1975] Anderson, A.R. and Belnap, N.D. Entailment. (The Logic of Relevance and Necessity.) Princeton 1975.

[Barwise 1979] Barwise, Jon. Perception and Inference.

[Barwise 1981] Barwise, Jon. Scenes and Other Situations. The Journal of Philosophy LXXVIII (7): 369-397, July, 1981.

[Barwise and Perry 1983] Barwise, Jon and Perry, John. Situations and Attitudes. MIT Press, 1983.

[Huges and Cresswell 1972] Hughes, G.E. and Cresswell, M.J. An Introduction to Modal Logic. Methuen, 1972.

[Makinson 1973] Makinson, D.C. Topics in Modern Logic. Methuen 1973.

MASS TERMS AND QUANTIFICATION*

Contents

Jan Tore Lønning

0. Introduction.

Some typical examples of mass terms are: water, gold and cheese. They are separated from typical count terms like man, house and book, both on semantic and syntactic grounds e.g. in the determiners they take. But the dividing line is not that sharp. Some words are said to belong to both classes e.g. cake, typical mas terms like water has occurences as count terms, and it is even claimed (Pelletier (1975)) that every count noun can occur as a mass term. We shall not consider this claim, but from the mere fact that it has been stated, we can conclude that it may be better to look upon the distinction between mass terms and count terms as a distinction between different types of noun occurences in sentences, and not as a lexical distinction, even though many nouns nearly always occur in only one of the two categories.

I shall not try to classify all noun occurences according to this dichotomy, nor shall I discuss whether such a classification is possible. What I shall do, is to separate out two smaller classes of noun occurences which are clear cut examples of mass terms and count terms, respectively, and whose semantics will occupy us in the sequel. The mass terms will be nouns (or nouns together with modifying adjectives or restrictive clauses) in indefinite singular preceded by one of the determiners all, some, much, most, little, a little or an amount term like two liters of, less than one kilo of and so on. The whole noun phrase including the determiner or amount term will be called a mass noun phrase. In comparison, typical count terms are nouns (or modified nouns) in singular preceded by a, an, or every, or in plural preceded by many, most, few, a few, two, three, ... more than two, less than three and so on.

Definite descriptions will also be of concern to us since there are semantic differences between expressions like the red car and expressions like the water that John drank. Intuitively the first one is generated from a count term, the latter from a mass term. However, since we have abandoned the idea of the mass/count dichotomy as a lexical distinction we have also lost the possibility to distinguish these two definite descriptions on purely syntactic grounds. That is maybe not so bad, because it is possible to give at least the latter one a second ("countable") reading. We shall be careful not to use the technical term mass noun phrase to any occurrence of definite descriptions for reasons to be given in section 1.

Many studies of mass terms (e.g. Quine (1960), Parsons (1970)) take as the starting point occurences of mass terms without determiners e.g. in the sentence Water is wet. There are several reasons for why we deviate from this practice. First, of course, mass terms with determiners are semantically interesting in themselves. Second, it is possible, as we have seen, to separate out a class of mass terms with determiners syntactically from a class of count terms with determiners. This may not be possible for mass terms without determiners, at least not in a language like Norwegian where kjøre bil (drive a car) has the same form as drikke vann (drink water). Third, if one wants to study the semantics of mass terms with no determiners, one runs into problems which are independent of the mass/count distinction. Water is wet has a form very similar to Horses are kind and one is thus lead into problems with so-called generics. On the other hand, in the study of mass terms with determiners one can learn from the study of count terms with determiners. We shall particularly profit from Barwise and Cooper's (1981) study of generalized quantifiers. Fourth, most analyses of the sentence Water is wet (e.g. Parsons (1970), Montague (1973 b)) propose to paraphrase it as All water is wet or Most water is wet, thus in effect reducing the study to the cases we propose to consider.

Why is the analysis of quantification and mass terms problematic? Why cannot this problem be treated along the same lines as "normal" quantification, letting water as a mass term refer to all quantifies of the world's water, All water is wet be true if and only if every quantity of water is wet, and Some water is wet be true if and only if some quantity of water is wet? I think that this approach works if one restricts oneself to the logical determiners all and some, but one gets into problems when more general mass noun phrases and definite descriptions are taken into consideration.[1]

A definite description like the red car has meaning in a context only if there is one and only one red car present. But if John drank one quantity of water there were many more quantities that he also drank, still the description the water that John drank has a clear meaning. There has been two ways to tackle this. The first approach, proposed e.g. by Burge (1972), is often called mereology and stems from Leonard and Goodman's (1940) "calculus of individuals". The idea is to add to the semantic model a new primitive relation between the individuals called overlap. This in turn gives rise to an inclusion relation between the individuals. Then the description the water that John drank is taken to denote the maximal element under this inclusion relation in the set denoted by water that John drank. What is needed to make this work, is some restrictions on the semantics that secures this set to contain a maximal element, and at the same time does not force the set denoted by gold that weighs 2 kilos to contain a maximal element. I have not yet seen the necessary details carried out.

The other approach to definite descriptions has been to keep the model standard and then paraphrase the gold in Smith's ring as something like the gold constituting Smith's ring (Montague (1973 b)) which intuitively works in this example since there is only one gold quantity that constitutes Smith's ring. But it is problematic to paraphrase some definite descriptions and not all if one wants a semantic analysis that match the syntactic analysis as in Montague (1973 a) grammar. It is also a problem to find the right way of paraphrasing; e.g. it may seem odd to use constitute in connection with the water that John drank. The main objection is, however, that this is not a satisfactory formal semantic explanation. Even though it is intuitively clear that there is only one gold quantity that constitutes Smith's ring we cannot be sure that gold constituting Smith's ring will denote a unit set in every model, at least not if constitute is treated as a normal transitive verb and assigned as denotation a (non-specific) set of ordered pairs. In that case, the cardinality of the set denoted by constituting Smith's ring will vary from model to model, and the same will happen with the intersection of this set and the set denoted by gold, i.e. the set denoted by gold constituting Smith's ring. The result will be models where the definite description is undefined.

Then there is the question how mass noun phrases with non-logical determiners like much and most or amount terms can be handled when water is interpreted as quantities of water.

Much water does not mean the same as many quantities of water, if the latter one has any meaning at all. We can say that one glass contains more water than another glass, but we will not say that it contains more quantities of water. One proposal for handling much (Montague (1973 b)) is to paraphrase much water F as some large quantity of water F. But this does not generalize to little or to amount terms. John drank less than one liter of milk does not mean the same as John drank a quantity of milk with volume less than one liter. The latter sentence is true even if John drank 5 liters of milk. Furthermore it is hard to see that such rewritings is compatible with a strategy that aims at a semantic analysis which is parallel to the syntactic analysis.

A related problem has to do with negation. If some water boiled and some water did not boil, there were quantities of water that partly boiled and partly did not boil. More precisely, there would have been at least as many such quantities as there were quantities that boiled. If now negation is interpreted as set-theoretic complement in the set of quantities then the set denoted by water that did not boil cannot have smaller cardinality than the set denoted by water that boiled. It follows that the sentence More than half (of the) water boiled only can be true if All (the) water boiled is true. The sentence

(1) Much water did not boil

has according to my intuitions two readings - one with wide scope negation and the other with narrow scope negation:

(1a) It is not the case that much water boiled

(1b) It was much water that did not boil,

respectively. If much water boiled, and some water did not boil, (1b) will become true with these interpretations. So it seems that interpreting water as quantity of water is not compatible with normal interpretation of negation.

The goal of this paper is to present a new theory for mass term quantification which takes these problems into consideration. At the same time we must be able to give an account for the valid inferences of previous theories, in particular, of Bunt's (1979) examples

(2) (a) All water is water

(b) All blue water is water

(c) All blue water is blue

We will also consider it as a goal to show how it is possible to get a valid form to the argument

(3) Much water evaporated

All that evaporated, disappeared

∴ Much water disappeared

and at the same time not to get a valid argument if much is exchanged with little. The final goal is to show that the theory naturally extends to treat amount terms and to give the connections between the sentences

(4) Less than two kilos of cheese disappeared

(5) The cheese that disappeared weighed less than two kilos

Before continuing let me add some general remarks. Our aim is to give in a uniform way a semantics of natural languages in modeltheoretic terms. Thus we aim to give the semantics uniformly (or mechanically) from the syntactic analysis as in Montague (1973 a) grammar. This excludes, in particular, any kind of rewriting which is not uniform.

Further, our interest lies in the natural language itself, not in the world it describes. This means that the models we build are not necessarily "true" models of the physical world, but means to understand the language, in particular, to give valid forms to intuitively true sentences and inferences.

Finally, natural languages have priority not only over the physical world, but also over all kinds of formal languages and logical systems. What is of primary interest is the natural language and its modeltheoretic denotations, not any intermediate formal language. But an intermediate formal language may be a valuable tool in studying the semantics of our natural language fragment, in particular, in gaining some insight into the complexity of the structure and the inferences in the fragment we are studying. It is to this end that we introduce the system LM in section 2.

The treatment here will be purely extensional. This is not because every mass noun phrase can be taken to be extensional. But we need first to get a correct account of the difference between mass noun phrases and count noun phrases. And that is independent of intensionality.

The rest of the paper is organized as follows. In section one we discuss the idea of homogeneous references and show how this gives a clue to a "Boolean" semantics for mass noun phrases. In section two we construct a simple formal system LM and show in section three how a fragment of English can be interpreted via it. In section four we discuss the metamathematics of LM, this section can be omitted. In section five we discuss a combined system for mass terms and count terms, and in section six we extend the fragment to include amount terms.

1. Homogeneous reference and boolean algebras.

Two proposed universals for the semantics of mass tems, cummulative and distributive reference, have been widely discussed. On cummulative reference we read in Quine (1960): "So-called mass terms like 'water', 'footwear', and 'red' have the semantical property of referring cummulatively: any sum of parts which are water is water" (p 91). This may, however, also be fulfilled by terms which are not naturally classified as mass terms such as plurals of nouns, and - if one, like Quine, extends the category mass terms to some occurences of adjectives - an adjective like heavy.

A furhter criterion has therefore been proposed, namely distributive reference (Cheng (1973)): "any parts of something which is water are water"[2]. This has been more disputed than the first criterion for the following reason (Quine (1960)): "... there are parts of water, sugar, and furniture too small to count as water, sugar, furniture. Moreover, what is too small to count as furniture is not too small to count as water or sugar; ..." (p 99).

Bunt (1979) has pointed out that this can be read in two different ways. Either as a statement about the world around us, and in that case it is true, but irrelevant. Or as a statement about our use of the language. In that case it is necessary to point out where in our use of the mass terms, the existence of minimal parts is reflected. But Bunt finds no such evidence forthcoming.

Quine's goals in "Word & Object" are different from the goals of this paper. He is interested in the relation between the language and the outer world, how the language is used for referring. This is reflected not only in his objections to the distributive reference criterion but also in the formulation of the cummulative reference criterion. For our purposes it is not critical if mass terms really refers homogeneously, that is both cummulatively and distributively. Rather what is of importance is whether they behave as if they did and what it means to behave in such a way.

From this discussion and from the fact that we have defined mass terms as certain specific occurrences of nouns in sentences, it is natural to take as point of departure some examples of sentences containing mass terms:

(6) (a) Much water boiled

(b) John drank much water

(c) *Much water weighed two grams

(d) *Much water contained ten grams of salt.

Rather than focusing on the mass terms alone, let us shift attention to the rest of the sentences and ask which semantic property is shared by the phrases boiled and John drank in opposition to weighed two grams and contained ten grams of salt? It is, of course, the property of homogeneous reference. If John drank two parts, he drank the sum of them and any smaller parts contained in them. On the other hand, the sum of two (different) parts, each weighing two grams is more than two grams and parts of them weigh less than two grams. An analysis of more examples gives the same result, so I propose the following as a linguistical universal[3]:

Homogeneous Constraint:

Mass noun phrases combine only with homogeneous expressions[4] to form sentences.

Intuitively, a homogeneous expression is an expression which refers homogeneously. But to avoid undesirable references to the outer world and speculation on whether an expression refers homogeneously in all possible worlds we will redefine homogeneity without mentioning reference:

Let α be a noun which can occur as a mass term, and let β be a noun modifier which can combine with α to form a mass term. The expression δ is said to be homogeneous if and only if the following two inferences are valid:

CUM The $\beta\alpha$ δ
The not-β α δ
∴ The α δ

DISTR The α δ
There is some $\beta\alpha$
∴ The $\beta\alpha$ δ

To give an example on how this definition works, say we want to test disappeared : Then we can choose gold for α, white for β and substitute disappeared for δ. The two inferences then become

(7) The white gold disappeared
The not-white gold disappeared
∴ The gold disappeared

(8) The gold disappeared
There was some white gold
∴ The white gold disappeared

which are valid. We can test any expression which can combine with the white gold to form a sentence, with this test. For a phrase like John drank the word order has to be changed, of course, so the first premiss in CUM reads John drank the white gold.

It is our opinion that the phrases disappeared, boiled and John drank pass this test, while contained ten grams of salt and weighed two kilos fail.

Is it possible to restate the Homogenous Constraint in purely syntactic terms? We think it can be done at least to some extent. First, nearly all intransitive verbs seem to be homogeneous and can combine with mass noun phrases. But some exceptions exist, including meet and gather.

For transitive verbs the picture is more complex. We have seen that John drank can combine with mass noun phrases. That is also true for the phrase Many men drank, even though this does not pass the homogeneity test. It is clearly enough that what each man drank was homogeneous, so a slight modification either in the Homogeneous Constraint or in the test is necessary to cover this. It is more doubtful whether mass noun phrases can be found in direct object position with any transitive verb. It is at least difficult to get an extensional reading of love much water.

Sentence (6d) indicates that a mass noun phrase cannot stand in subject position to a transitive verb when the direct object is quantified, and that seems to be the case for all transitive verbs. On the other hand, when the direct object is not quantified a combination is possible, e.g. much water contained salt.

Similar observations as for transitive verbs may be done for 3-place verbs. Finally, sentence (6c) exemplifies that mass noun phrases cannot have subject position to measure verbs.

Observe that if much water is exchanged with the water that John drank in any of the example sentences in (6), the result is a well-formed sentence. Definite descriptions are not subject to the Homogeneous Constraint, and that is one reason why they should be kept separate from mass noun phrases.

Let us turn to the formal semantic treatment of mass terms, and start with the observation that countability pervades current formal semantics. This means that the denotation types of all types of terms is determined by the fact that they occur in sentences where the noun terms are count terms. Thus a verb phrase like disappeared is taken to denote a set, and we get a natural semantics of the sentence a man disappeared, viz. it is true if and only if the intersection between the set denoted by man and the set denoted by disappeared is non-empty, i.e. contains an individual. But if we look at the sentence much water disappeared it may be unnatural to let disappear denote a set, and we may run into problems if disappear is presupposed to denote a set and the denotations of water and much water are forced to submit to this. We must thus be willing to rethink the semantics and not only ask for the right types of denotations for the mass terms, but also for the co-occurring terms. And it is here that we shall take our cue from the Homogeneous Constraint. And we shall also have to account for the semantic difference between the verb phrases in (6a) and (6d), which are given the same type of denotation in the count case.

The proposal for a semantic model for mass terms is based on the concept of a boolean algebra. A boolean algebra is a set A with at least two elements named 0 and 1, where there is defined a binary function +, sum, which to any two elements a and b gives a new element $a+b$, another binary function $\times$, product, which to two elements gives the product, $a\times b$, and a unary function $-$, complement, which to the element a gives its complement $\bar{a}$. (Other names and symbols such as $\vee$ and joint for $+$ and $\wedge$, $\cdot$, meet and intersection for $\times$ are frequently used). The last part of the definition of a boolean algebra is certain postulates or laws the functions and special elements have to consider, such as $a+b = b+a$ (commutativity), $a+a = a$ (idempotens), $a+\bar{a} = 1$ and $a\times\bar{a} = 0$. A typical example of a boolean algebra is the set of all subsets of a nonempty set X, where 0 and 1 are interpreted as $\emptyset$ and X and $+$, $\times$ and $-$ are interpreted as $\cup$ (union), $\cap$ (intersection) and $^{-X}$ (complement with respect to X) respectively. What will be of interest to us later is that it can be proved that every boolean algebra is isomorphic to a subset of the set of all subsets of a nonempty set X with these interpretations.

It is not necessary with so many primitives to define the boolean algebras. It is, for example sufficient with $\times$, $-$, and the following set of postulates (Hughes and Cresswell (1972)):

BA1 A contains at least two elements

BA2 If $a,b\in A$ then $\bar{a}\in A$ and $a\times b\in A$

BA3 If $a,b\in A$ then $a\times b = b\times a$

BA4 If $a,b,c\in A$ then $a\times(b\times c) = (a\times b)\times c$

BA5 For all $a,b\in A$, if there is some c such that $a\times\bar{b} = c\times\bar{c}$, then $a\times b = a$

BA6 For all $a,b,c\in A$ if $a\times b = a$, then $a\times\bar{b} = c\times\bar{c}$.

From this 0, 1 and $+$ can be defined:

$$0 =_{df} a\times\bar{a}$$

$$1 =_{df} \bar{0}$$

$$a+b =_{df} \overline{(\bar{a}\times\bar{b})}$$ [5]

The intended use of a boolean algebra as a model for sentences containing mass noun phrases is then the following: Homogeneous expressions like boiled and John drank shall denote elements in the algebra. Mass noun phrases like much water shall denote subsets of the algebra and the sentence much water boiled will be true if the element denoted by boiled is a member of the set denoted by much water. Mass terms shall also denote elements of the algebra, and determiners shall denote relations between elements of the algebra, or, equivalently, functions which to elements of the algebra assigns subsets of it. This resembles the generalized quantifier approach to count terms, where determiners are taken to be functions which to subsets of the individual domain assign sets of such subsets (Barwise and Cooper (1981)). The determiners all and some will have a fixed (or logical or modelindependent) interpretation. The interprettion of some shall be the function which to each element α in the algebra assigns $\|\text{some}\|(\alpha) = \{\beta \mid \alpha\times\beta\neq 0\}$, and $\|\text{all}\|$ shall be such that $\|\text{all}\|(\alpha) = \{\beta \mid \alpha\leqslant\beta\}$ where $\alpha\leqslant\beta$ is defined to be $\alpha\times\beta = \alpha$.

This model is a formal model for the language - not a model of the world. Still it may be of help to think that the elements of the algebra are the quantities (or portions of matter or bits of matter) in the actual world. 1 is the quantity which is the sum of all the other quantities and 0 is the empty quantity. Water refers to the totality of the worlds water, boiled to the totality of what boiled at the time interval involved and some water boiled is true if and only if the quantity which is the product of these quantities or what is refered to by the water that boiled is different from the empty quantity. But since the model works as well for abstract mass terms as for concrete mass terms, it is necessary to think of the elements of the boolean algebra as more general than physical quantities.

Let us see how this model can be used to solve some of the problems presented in the introduction, and let us start with the question on how adjectives can be treated. Only predicative adjectives will be considered. An adjective α is predicative if x is an $\alpha\beta$ if and only if x is α and x is a β for every noun β. For count terms this is the same as saying that α can be considered as a general term and be treated as a predicate in an analysis in a first order language. Typical examples are red and square. We observe however, like Quine (1960) and Bunt (1981), that only a smaller class of these can stand attributively to mass terms

(9) (a) Much red water boiled

(b) *Much square water boiled

It is the same class which can stand predicatively to mass noun phrases

(10) (a) Much of the water was red

*Much of the water was square

Of course the semantic difference between red and square lies in the fact that red refers homogeneously and square does not. We can use the previously established criteria CUM and DISTR to the expression was β to decide whether a predicative adjective β is homogeneous or not. And our observation is that of the predicative adjectives only the homogeneous ones stand attributively to mass terms and predicatively to mass noun phrases[6].

Homogeneous predicative adjectives shall denote elements in the boolean algebra. A mass term like red water shall denote the product of the denotations of red and water. The verb phrase was red shall denote the same as red in this simplified model where only one time interval is considered. The reader may check that the sentences of example (2) become valid with these interpretations.

Adjectival clauses behave similar to adjectives. We observe that only clauses constructed from homogeneous expressions combine with mass terms as these examples show:

(11) (a) Much (of the) water that John drank, ...

(b) *Much (of the) water that weighed two grams, ...

In the model that John drank can be given the same denotation as John drank, an element of the boolean algebra. What the denotations of John and drank shall be, will be discussed in part 5.

In the introduction we discussed the possible readings of the negation in the sentence much water did not boil. The wide scope reading (1a) will be captured in the propositional part of the model. To capture the narrow scope negation we let (1b) did not boil denote the (boolean) complement of the denotation of boil. This will also give a sound denotation to water that did not boil.

Definite descriptions will be given a very simple analysis in this model. The water that John drank shall denote the same element of the algebra as water that John drank. This is well-defined since water that John drank denotes one and only one element. The difference between water and the water is intuitively that water refers to the totality of water in the world while the water refers to a definite subquantity. That is also why it is more natural to say much of the water that John drank than much water that John drank. In a more fine-grained analysis where place-references and tense, aspect, and time-references are taken into consideration, this difference has to be made clear. At this stage however, we shall not differentiate semantically between water and the water nor between much water and much of the water.

The sentence (12a) can then be paraphrased as (12b).

(12) (a) The water boiled

(b) All (the) water boiled

Definite descriptions can also combine with inhomogeneous expressions as we saw in

(5) The cheese that disappeared weighed less than two kilos

Here the inhomogeneous verb phrase weighed less than two kilos shall denote a subset of the boolean algebra. The sentence is true if the element denoted by the cheese that disappeared is a member of this set. It is easy to see the problems which would arise if inhomogeneous verb phrases were to denote elements of the algebra. Thus in our proposal verb phrases get denotations on different levels depending on whether they are homogeneous or not. This will give a natural account for the lack of grammaticality of the examples (6c) and (6d). It will also bring about that the rewriting in (12) is uniform, i.e. whenever the verb phrase is homogeneous rewrite the α as all (the) α.

2. LM - A logic for mass noun phrases.

A LM language consists of:

2.1. Logical symbols.

a) propositional connectives: $\vee$, $\neg$

b) parantheses: (,)

c) two operator symbols: - (unary), • (binary)

d) two logical determiners: All, Some

2.2. Non-logical symbols.

a) a nonempty set of constant symbols, e.g. a, b, c, ...

b) a (possibly empty) set of nonlogical determiners, e.g. D_1, D_2 ...

2.3. Formation rules.

F1-Terms

a) constant symbols are terms

b) if t is a term, then (-t) is a term

c) if s and t are terms then (t•s) is a term

F2-Quantifiers

if D is a determiner and t is a term, then D(t) is a quantifier

F3- Formulas

a) if Q is a quantifier and t is a term then Q(t) is a formula

b) If p and q are formulas, then $(p \vee q)$ and $\neg p$ are formulas

The well-formed expressions of the language is defined to be anything that can be built up by using the rules F1-F3 a finite number of times.

Other propositional connectives can be defined in the usual way:

$$p \wedge q =_{df} \neg(\neg p \vee \neg q)$$

$$p \rightarrow q =_{df} \neg p \vee q$$

Parantheses will be omitted when this does not give rise to ambiguities.

2.4. Semantics

A model for a LM-language L consists of a boolean algebra $\langle A,+,\times,\bar{\ },0,1\rangle$ and an interpretation function $\|\cdot\|$ defined on the non-logical symbols of L such that

S1 a) $\|\underline{a}\| \in A$ for every constant symbol $\underline{a}$

b) For each non-logical determiner symbol $\underline{D}$, $\|\underline{D}\|$ shall be a function which to each element $a \in A$ gives a subset $\|\underline{D}\|(a)$ of A such that

$b \in \|\underline{D}\|(a)$ if and only if $a \times b \in \|\underline{D}\|(a)$

($\|\underline{D}\|(a)$ "lives on" a).

The interpretations of the expressions of the language can then be given by extending the function $\|\cdot\|$.

S2 a) $\|\underline{\text{All}}\|$ is the function which to each $a \in A$ gives

$\|\underline{\text{All}}\|(a) = \{b \in A : a \leq b\}$.

($\leq$ is the boolean ordering relation which can be defined by: $a \leq b$ if and only if $a \times b = a$)

b) $\|\underline{\text{Some}}\|$ is the function which to each $a \in A$ gives

$\|\underline{\text{Some}}\|(a) = \{b \in A : a \times b \neq 0\}$.

S3 a) $\|(-\underline{t})\| = \overline{\|\underline{t}\|}$ for every term $\underline{t}$

b) $\|(\underline{t} \cdot \underline{s})\| = \|\underline{t}\| \times \|\underline{s}\|$ for all terms $\underline{t}$ and $\underline{s}$.

S4 $\|\underline{D}(\underline{t})\| = \|\underline{D}\|(\|\underline{t}\|)$ for every term $\underline{t}$ and and non-logical determiner $\underline{D}$.

S5 a) If $\underline{Q}$ is a quantifier and $\underline{t}$ a term, then

$$\|\underline{Q}(\underline{t})\| = \begin{cases} 1 & \text{if and only if } \|\underline{t}\| \in \|\underline{Q}\| \\ 0 & \text{otherwise} \end{cases}$$

b) If ϕ and ψ are formulas, then

$$\|\neg\phi\| = \begin{cases} 1 & \text{if } \|\phi\| = 0 \\ 0 & \text{if } \|\phi\| = 1 \end{cases}$$

$$\|\phi \vee \psi\| = \begin{cases} 1 & \text{if } \|\phi\| = 1 \text{ or } \|\psi\| = 1 \\ 0 & \text{if } \|\phi\| = 0 \text{ and } \|\psi\| = 0 \end{cases}$$

If $\mathfrak{A} = \langle\langle A, +, \times, ^{-}, 0, 1\rangle, \|\cdot\|\rangle$ is a model and ϕ a formula such that $\|\phi\| = 1$, ϕ is said to be <u>true in $\mathfrak{A}$</u> or equivalently <u>$\mathfrak{A}$ is a model for ϕ</u> (in symbols $\mathfrak{A} \models \phi$). If all the formulas in a set of formulas Γ are true in $\mathfrak{A}$, $\mathfrak{A}$ is said to be <u>a model for Γ</u> ($\mathfrak{A} \models \Gamma$). A formula ϕ which is true in every model is <u>valid</u> ($\models \phi$).

In the rule S1 b) it was presupposed that every determiner $\underline{D}$ forms a quantifier $\|\underline{D}\|(a)$ which "<u>lives on</u>" a. This is an empirical fact taken over from Barwise and Cooper's (1981) treatment of the count determiners. It seems to be true of mass determiners as well and is as easy to formulate in this semantic frame as in the set-theoretical one.

An important subclassification of the determiners in the count case can also be formulated in this frame. A quantifier $\underline{Q}$ is said to be <u>monotone (increasing)</u> in a model if for all $a, b \in A$ the following holds: $a \in \|\underline{Q}\|$ and $a \leq b$ implies $b \in \|\underline{Q}\|$. A determiner $\underline{D}$ is said to be monotone if it always gives rise to monotone quantifiers. <u>Monotone decreasing</u> quantifiers and determiners can be defined in a similar way.

3. Application of LM to a small fragment of English.

In this section we present a small fragment of English containing mass noun phrases together with a translation procedure into the system LM, thus giving the fragment a modeltheoretic semantics.

3.1. Lexicon.

N - {water, gold, salt}

AD - {blue, hot}

VP - {boiled, evaporated, disappeared}

Det - {all, some, much, little, one kilo of, less than two kilos of}

3.2. Syntactic rules.

We define the set of <u>structural descriptions</u>.

<u>SR0</u> - Lexical insertion:

If α is a word listed in the lexicon under A then $[_A\alpha]$ is a SD.

The phrase structure rules will have the form A → BC which is to be read: if α,β are SD's of forms $[_B\gamma],[_C\delta]$ respectively then $[_A\alpha\beta]$ is a SD.

<u>SR1</u> - S → { NP VP
NP negVP - if the VP is not already negated
S <u>and</u> S
S <u>or</u> S }

<u>SR2</u> - NP → { Det N
Det R
Det <u>of</u> DD
DD }

<u>SR3</u> - DD → <u>the</u> N

<u>SR4</u> - N → $\begin{cases} \text{AD} \quad \text{N} \\ \text{N} \quad \text{R} \end{cases}$

<u>SR5</u> - VP → $\begin{cases} \underline{\text{was}} \quad \text{AD} \\ \underline{\text{was}} \quad \text{N} \\ \text{negVP - if the VP is not already negated} \end{cases}$

<u>SR6</u> - R → <u>that</u> VP

negVP is the negation of the VP. A VP which is not negated has either the form <u>was</u> α or it is a lexical expression of the form α-<u>ed</u>. In the first case negVP is <u>was not</u> α, in the second case negVP is <u>did not</u> α.

3.3. Translation into LM

We define the translations into LM of the expressions in the English fragment inductively on the structural descriptions. A translation α' of an expression α is uniquely determined by the derivation of α, but not necessarily by α due to the ambiguity of the negation.

TO - a) Each word listed in the lexicon under N, AD or VP shall be translated to a constant symbol.

b) The determiners <u>all</u> and <u>some</u> shall be translated to the logical determiners <u>All</u> and <u>Some</u> respectively.

c) Other words listed in the lexicon under <u>Det</u> shall be translated to non-logical determiner symbols.

d) If α is a SD of the form $[_X \beta]$ where β is in the lexicon, then the translation of α, α' is β' given in a)-c).

T1 -	$[_S[_{NP}\alpha][_{VP}\beta]]$	translates	as	$\alpha'(\beta')$
	$[_S[_{NP}\alpha]neg[_{VP}\beta]$	"	"	$\neg\alpha'(\beta')$
	$[_S[_S\phi]\underline{and}[_S\psi]]$	"	"	$\phi'\wedge\psi'$
	$[_S[_S\phi]\underline{or}[_S\psi]]$	"	"	$\phi'\vee\psi'$
T2 -	$[_{NP}[_{Det}\delta][_N\alpha]]$	"	"	$\delta'(\alpha')$
	$[_{NP}[_{Det}\delta][_R\rho]]$	"	"	$\delta'(\rho')$
	$[_{NP}[_{Det}\delta]\underline{of}[_{DD}\gamma]]$	"	"	$\delta'(\gamma')$
	$[_{NP}[_{DD}\gamma]]$	"	"	$\underline{All}(\gamma')$
T3 -	$[_{DD}\underline{the}[_N\alpha]]$	"	"	α'
T4 -	$[_N[_{AD}\varepsilon][_N\alpha]]$	"	"	$(\varepsilon'\bullet\alpha')$
	$[_N[_N\alpha][_R\rho]]$	"	"	$(\alpha'\bullet\rho')$
T5 -	$[_{VP}\underline{was}[_{AD}\varepsilon]]$	"	"	ε'
	$[_{VP}\underline{was}[_N\alpha]]$	"	"	α'
	$[_{VP}neg[_{VP}\beta]]$	"	"	$(-\beta')$
T6 -	$[_R\underline{that}[_{VP}\beta]]$	"	"	β'

3.4. Examples.

The sentence

(1) Much water did not boil

can be given two different syntactical derivations:

(1a)' $[_S[_{NP}[_{Det}\text{Much}][_N\text{water}]][_{neg}\text{did not}][_{VP}\text{boil}]]$

(1b)' $[_S[_{NP}[_{Det}\text{Much}][_N\text{water}]][_{VP}[_{neg}\text{did not}][_{VP}\text{boil}]]]$

with corresponding LM-translations.

(1a)" $\neg\underline{\text{Much}}\ (\underline{\text{water}})\ (\underline{\text{boil}})$

(1b)" $\underline{\text{Much}}\ (\underline{\text{water}})\ (\underline{-\text{boil}})$

which reflects the two possible readings of the sentence.

The sentence (2c) gets the derivation (2c)' and the valid LM-translation (2c)".

(2c) All blue water was blue[7]

(2c)' $[_S[_{NP}[_{Det}\text{All}][_N[_{AD}\text{blue}][_N\text{water}]]][_{VP}\underline{\text{was}}[_{AD}\text{blue}]]]$

(2c)" All(blue•water)(blue)

The inference (3) gets the translation

(3)' Much(water)(evaporate)

All(evaporate)(disappear)

∴ Much(water)(disappear)

This inference is valid if Much is demanded to be monotone increasing. Since it is not reasonable that little or less than two kilos of are monotone increasing, the inference will cease to be valid if much is exchanged with any of these. But if on the other hand less than two kilos of is demanded to be monotone decreasing then the following inference gets a valid translation

(13) Less than two kilos of water disappeared

All that evaporated, disappeared

∴ Less than two kilos of water evaporated

4. Formal aspects of the logic LM.

We now turn to a study of the model theory of the language LM. A reader who does not want to go into the mathematical details may without loss of continuity proceed to section 5.

We start with the simpler system LA. A LA-language is a LM-language with no non-logical determiners. The determiner <u>Some</u> can be defined from <u>All</u> by the schema:

$$\underline{\text{Some}}(t)(s) \leftrightarrow \neg\underline{\text{All}}(t)(-s)$$

LA can then be axiomatized as follows:

<u>Axiom Schemata</u>[8]:

for all terms t, s and r:

A1 $\underline{\text{All}}(t)(t)$

A2 $\underline{\text{All}}(t)(s) \wedge \underline{\text{All}}(s)(r) \rightarrow \underline{\text{All}}(t)(r)$

A3 $\underline{\text{All}}(t \cdot s)(s \cdot t)$

A4 $\underline{\text{All}}(t \cdot s)(t)$

A5 $\underline{\text{All}}(t)(s) \wedge \underline{\text{All}}(t)(r) \rightarrow \underline{\text{All}}(t)(s \cdot r)$

A6 $\underline{\text{All}}(t \cdot (-s))(r \cdot (-r)) \rightarrow \underline{\text{All}}(t)(s)$

A7 $\underline{\text{All}}(t)(s) \rightarrow \underline{\text{All}}(t \cdot (-s))(r \cdot (-r))$

A8 $\underline{\text{All}}(t)(r \cdot (-r)) \rightarrow \neg\underline{\text{All}}(-t)(r \cdot (-r))$

<u>Rule of inference</u>:

If ϕ is a tautological consequence of $\phi_1, \phi_2, \ldots, \phi_n$, then from $\phi_1, \phi_2, \ldots, \phi_n$ to infer ϕ.

A <u>proof</u> is defined (in the usual way) as a sequence of formulas where each formula is either an axiom or follows from the earlier formulas in the sequence by the rule of inference. The last formula in such a sequence is said to be <u>provable</u> or to be a <u>theorem</u>. As usual we write $\vdash \phi$ to assert that ϕ is a theorem.

If we allow formulas from some set of formulas Σ in the sequence, ϕ is said to be provable from Σ ($\Sigma \vdash \phi$). A set of formulas is said to be consistent if there is a formula ϕ such that not $\Sigma \vdash \phi$.

Theorem 1. Every theorem is valid.

Proof. Remembering that the interpretation of All is nothing but the boolean ordering relation $\leq$, it is straightforward to check that all the axioms are true in every boolean algebra. And the rule of inference of course preserves validity since the interpretation of the propositional calculus is "normal".

q.e.d.

Theorem 2.

a) Every consistent set of formulas has a model

b) ϕ is provable from Σ if and only if ϕ is true in every model for Σ.

c) ϕ is provable if and only if it is valid.

Proof. Since the tautology rule is the only rule of inference, we have that $\Sigma \cup \{\psi\} \vdash \phi$ if and only if $\Sigma \vdash \psi \rightarrow \phi$ and $\Sigma \vdash \phi$ if and only if $\Sigma \cup \{\neg\phi\}$ is inconsistent. Hence b) follows from a) and c) is just a special case of b). To prove a) suppose Σ is a consistent set. Then Σ can be extended to a maximal consistent set Γ, that is a consistent set which is not properly contained in any larger consistent set. For every formula $\phi, \Gamma \vdash \phi$ if and only if $\phi \in \Gamma$, and either $\phi \in \Gamma$ or $\neg\phi \in \Gamma$ but not both.

A model for Γ, and hence for Σ, can then be constructed as follows. Define the relation $\approx$ on the terms by

$$t \simeq s \quad \text{if and only if} \quad \Gamma \vdash \underline{\mathrm{All}}(t)(s) \wedge \underline{\mathrm{All}}(s)(t)$$

This relation is clearly symmetric and from the axioms A1 and A2 it follows that it is reflexive and transitive, so it is an equivalence relation. We can write $[t]$ for the equivalence class of t, A for the set of equivalence classes and define the unary operation $-$ and binary operation $\times$ on A by

$$\overline{[t]} = [(-t)]$$
$$[t]\times[s] = [(t\cdot s)]$$

It is necessary to check that these are well-defined. We leave the first one to the reader and verify the second one. Suppose $t \simeq r$ and $s \simeq p$. We must show that $(t\cdot s)\simeq(r\cdot p)$.

I	$\Gamma\vdash\underline{\mathrm{All}}(t)(r)$	- from the definition of $t\simeq r$ by the tautology rule
II	$\Gamma\vdash\underline{\mathrm{All}}(t\cdot s)(t)$	- instance of A4
III	$\Gamma\vdash\underline{\mathrm{All}}(t\cdot s)(r)$	- from I and II with the aid of A2
IV	$\Gamma\vdash\underline{\mathrm{All}}(s)(p)$	- from the definition of $s\simeq p$
V	$\Gamma\vdash\underline{\mathrm{All}}(t\cdot s)(s)$	- from the axioms A4, A3 and A2
VI	$\Gamma\vdash\underline{\mathrm{All}}(t\cdot s)(p)$	- from IV and V with the aid of A2
VII	$\Gamma\vdash\underline{\mathrm{All}}(t\cdot s)(r\cdot p)$	- from III, VI and A5
VIII	$\Gamma\vdash\underline{\mathrm{All}}(r\cdot p)(t\cdot s)$	- by an argument symmetric to I-VII
IX	$(t\cdot s)\simeq(r\cdot p)$	- from VII and VIII.

The next step is to prove that the structure $\langle A,-,\times\rangle$ is a boolean algebra. One way to do this is to check that the six principles BA1-BA6 listed in part 1 are satisfied. We give one example, BA5. Suppose $a,b,c\in A$ and $(a\times\bar{b}) = (c\times\bar{c})$. We shall show that $a\times b = a$. There must be terms t,s,r in the language of Γ such that $[t] = a$, $[s] = b$, $[r] = c$. It is sufficient to show that $(t \cdot s) \approx t$.

I	$\Gamma \vdash \underline{All}(t\cdot(-s))(r\cdot(-r))$	- from the supposition
II	$\Gamma \vdash \underline{All}(t)(s)$	- from I and A6
III	$\Gamma \vdash \underline{All}(t)(t)$	- instance of A1
IV	$\Gamma \vdash \underline{All}(t)(t\cdot s)$	- from II, III and A5
V	$\Gamma \vdash \underline{All}(t\cdot s)(t)$	- instance of A4
VI	$(t\cdot s) \approx t$	- from IV and V.

The last step is to define the interpretation function $\| \cdot \|$ on the constant symbols by $\| a \| = [a]$, and to check that $\mathfrak{A} = \langle\langle A, {}^{-}, \times\rangle, \| \ \| \rangle$ is a model for Γ or in other words that $\Gamma \vdash \phi$ if and only if $\mathfrak{A} \models \phi$ for every formula ϕ. This must be checked first for formulas of the form $\underline{All}(t)(s)$ and then by induction for propositional combinations of such formulas.

q.e.d.

Theorem 3. LA has a decision procedure.

Proof. Theorem 2c gives a characterization of the valid formulas, but not a way to decide if a formula is valid or not. To see that there is a method for doing this we start with an example. A formula with only two constant symbols $\underline{a}$ and $\underline{b}$ will have its truth-value in a model totally determined if we know which of the elements $\|a\| \times \|b\|$, $\|\bar{a}\| \times \|b\|$, $\|a\| \times \|\bar{b}\|$, $\|\bar{a}\| \times \|\bar{b}\|$ that equals 0 and which does not. There is only 2^4 different answers to this question and since a model has at least two elements it will be sufficient to check 2^4-1 different models. This line of argumentation may be repeated in the more general case to state that for a formula with n different constant symbols it is sufficient to inspect $2^{(2^n)}-1$ different models.

Implicit in the proof of theorem 3 is a normal form theorem for the system LA. Let ϕ be a formula in the constant symbols $a_1, a_2, \ldots, a_n$. Then it is not difficult to prove that ϕ is provably equivalent in LA to a formula ϕ' which is a propositional combination of formulas of the form

<u>Some</u>$(b_1 \cdot b_2 \cdot \ldots \cdot b_n)(b_1 \cdot b_2 \cdot \ldots \cdot b_n)$ where each b_i is either a_i or $(-a_i)$. The reason for this is that <u>Some</u>$(b_1 \cdot b_2 \cdot \ldots \cdot b_n)(b_1 \cdot \ldots \cdot b_n)$ is true in a model $\mathfrak{A} = \langle A, \| \ \| \rangle$ if and only if $\|b_1\| \times \|b_2\| \times \ldots \times \|b_n\| \neq 0$. From the normal form we can derive a proof procedure for the system LA.

q.e.d.

Before taking the whole LM into consideration we shall make some more comments on the system LA. First observe that A1-A8 can be taken as axioms for the first order theory of boolean algebras letting t, r and s be variables, $\cdot$ and $-$ function symbols and <u>All</u> a relation symbol. So this is one possible answer to the question (raised by Parsons (1970)) whether mass terms shall be translated into names or predicates in a first order analysis. For the other possible answer we start with a lemma.

<u>Lemma</u>. A set of LA formulas Γ which has a model has an atomic model, that is, a model where the boolean algebra is atomic.

<u>Proof</u>. A boolean algebra $\langle A, \times, +, ^-, 0, 1 \rangle$ is isomorphic to an algebra $\langle B, \cap, \cup, ^{-X}, \emptyset, X \rangle$ for a non-empty set X where $B \subset \mathcal{P}(X)$ (see e.g. Halmos (1963) for a proof). So if $\langle\langle A, \times, +, ^-, 0, 1 \rangle, \| \cdot \| \rangle$ is a model then $\langle\langle B, \cap, \cup, ^{-X}, \emptyset, X \rangle, f \circ \| \cdot \| \rangle$ is a model where f is an isomorphism from A to B. It is easy to see that the completed model $\langle\langle \mathcal{P}(X), \cap, \cup, ^{-X}, \emptyset, X \rangle, f \circ \| \cdot \| \rangle$ verifies the same formulas and hence is a model too.

q.e.d.

One consequence of this is that is is impossible to formulate in LA that a model is non-atomic, even if one uses an infinite set of formulas[9]. So the semantic aspects of mass terms that can be expressed in LA, and as we shall see later in the whole LM, is independent of whether mass term references have smallest parts or not.

LA does not only have a direct relationship to the first order theory of boolean algebras but also to first order languages, in particular in the form given in Barwise and Cooper (1981). To make this explicit, given a LA-language L, the L(GQ)-language L' which has a set of unary predicate symbols corresponding to the set of constant symbols in L as the only non-logical symbols, will be called the corresponding L(GQ)-language. A translation f from L into L' can be defined inductively as follows:

i) If $\underline{a}$ is a constant symbol then $f(\underline{a}) = \underline{Pa}$, the corresponding predicate symbol

ii) If t is a term of form (-s) and f(s) is defined, then $f(t) = \hat{x}[\neg f(s)(x)]$

iii) If t is a term of form $(s \cdot r)$ and f(s) and f(r) are defined then $f(t) = \hat{x}[f(s)(x) \wedge f(r)(x)]$

iv) $f(\underline{All}) = \underline{Every}$

v) If $Q = D(t)$ is a quantifier and f(D) and f(t) are defined then $f(Q) = f(D)(f(t))$

vi) If $\phi = Q(t)$ is a formula and f(Q) and f(t) are defined then $f(\phi) = f(Q)(f(t))$

vii) If $\phi = \neg\psi$ and $f(\psi)$ is defined then $f(\phi) = \neg f(\psi)$

viii) If $\phi = \psi \vee \eta$ and $f(\psi)$ and $f(\eta)$ are defined then $f(\phi) = f(\psi) \vee f(\eta)$

<u>Theorem 4</u>. A set for formulas Γ in a LA-language has a model if and only if the translation $f[\Gamma]$ into the corresponding L(GQ)-language has a model.

<u>Proof</u>. It is straightforward to check (by induction on the definition of f) that $\langle \langle \wp(X), \cap, \cup, {}^{-x}, 0, 1 \rangle, \| \cdot \| \rangle$ is an LA-model for Γ if and only if $\langle X, \| \cdot \| \rangle$ is an L(GQ)-model for $f[\Gamma]$. Then use the lemma.

q.e.d.

So this is the other possible answer to the question whether mass terms shall be translated into names or predicates in a first order analysis. It says also that LA is a reformulation of monadic first order logic with no reference to individuals.[10] So first order language is an extension of LA where more structure is imposed on the models and the language is equipped with tools to profit from these extensions, i.e. terms which make it possible to talk about individuals. One word of caution is appropriate. The translation of <u>All</u> into <u>Every</u> does not mean that the LA (and LM)-analysis of mass terms is equivalent to a L(GQ)-analysis which reads <u>all water</u> as <u>every quantity of water</u>, since the translation does not give rise to quantification over the set of quantities but rather over the set of atoms.

Let us now turn to the general LM-language. It can be axiomatized if we add the following schemata to A1-A8 of LA:

for all terms t, s, r and non-logical determiners D:

MQ $\quad$ <u>All</u>(t)(s)∧<u>All</u>(s)(t) → (D(r)(s) → D(r)(t))

MD1 $\quad$ <u>All</u>(t)(s)∧<u>All</u>(s)(t) → (D(s)(r) → D(t)(r))

MD2 $\quad$ D(t)(s) <-> D(t)(t·s)

The first two schemata give the extensionality of the non-logical determiners. The third one states that $\|D(t)\|$ lives on $\|t\|$ and should of course not have been added if we had not imposed this restriction on the semantics. Proof, provable (from Σ), theorem and consistent are defined as for the system LA (of course with this larger set of axioms).

Theorem 5.

a) Every consistent set of formulas has a model.

b) ϕ is provable from Σ if and only if ϕ is true in every model for Σ.

c) ϕ is provable if and only if it is valid.

Proof. One part of b) and c) is proved in the same way as theorem 1. What is new is to check that the new axioms are valid. a) and thereby the second part of b) and c) are proved in the same way as theorem 2a). What is new here is that we must define the interpretation function $\|\cdot\|$ on the non-logical determiners, and that in a way which secures that $\Gamma \vdash \phi$ if and only if $\langle\langle A, {}^{-}, \times\rangle, \|\cdot\|\rangle \models \phi$ for every formula ϕ. So if D is a non-logical determiner, let $\|D\|$ be the function which to each $a \in A$ gives $\|D\|(a) = \{b \in A \mid \text{there exists terms } t \text{ and } s \text{ such that } [t] = a,$ $[s] = a \times b \text{ and } \Gamma \vdash D(t)(s)\}$. This definition is clearly in accordance with the semantic rule for determiners. Moreover, with this definition it is possible to show that $\Gamma \vdash D(t)(s)$ if and only if $\langle\langle A, {}^{-}, \times\rangle, \|\cdot\|\rangle \models D(t)(s)$ for all terms t and s. Notice that all the three new schemata are used in this proof. Then the conclusion follows as in theorem 2.

q.e.d.

We shall not try to copy theorem 3 and show how to construct a decision- and proof procedure for LM. But this can easily be done along the folowing lines: We saw that for checking LA-validity it was only necessary to check a finite number of models where each such model could be assumed to be finite. Now, in such a finite model, it is only a finite number of ways to introduce an interpretation of a non-logical determiner. And this means that there will only be a finite number of non-isomorphic LM-models which must be considered to check the validity of a LM-formula ϕ. How many depends on the number of non-logical symbols in ϕ, i.e. constant symbols and non-logical determiners. This number will, however, be very large if more than very few symbols are involved. So other procedures must be developed to practical purposes.

If LA is considered as the first order theory of boolean algebras, LM is this theory extended with more relation symbols, all restricted by MD2. The lemma is valid for LM as well as for LA. And if a corresponding L(GQ)-language is defined as above with the addition that the L(GQ)-language shall have the same set of non-logical determiners as the LM-language and that the definition of f is extended with

o) f(D) = D for each non-logical determiner symbol,

then theorem 4 can be extended to LM-languages too.

We conclude this section with showing that the semantic important notion of monotonicity is axiomatizable. A determiner D is <u>consistent with</u> the schema

MM <u>All</u>(t)(s) → (D(r)(t) → D(r)(s)) if all instances of this, (i.e. for all t, r and s), can be added without causing inconsistencies.

<u>Theorem 6</u>. A determiner D is monotone (increasing) if and only if it is consistent with the schema MM.

<u>Proof</u>. It is immediate to see that every monotone determiner has to be consistent with MM. The other part is parallel to theorem 5a). The only difference is that $\| \; \|$ will have to be defined in a slightly modified way on the non-logical determiners consisent with MM. If D is such a determiner then $\|D\|$ shall be the function which to each $a \in A$ gives $\|D\|(a) = \{b \in A$: there exist terms s and t such that $[t] = a$, $[s] \leq a \times b$ and $\Gamma \vdash D(t)(s)\}$. This is clearly monotone and MM will secure that $\Gamma \vdash D(t)(s)$ if and only if $\langle\langle A, ^{-}, \times\rangle, \| \cdot \| \rangle \models D(t)(s)$, with this interpretation.

q.e.d.

5. Mass terms and count terms.

In our proposal for a semantics for mass terms, verb phrases, adjectives etc. have got denotations different from the usual ones. Hence it is not at all clear that count noun phrases can be treated in this frame. Moreover, examples like

(14) Much gold and many coins disappeared

show that mass noun phrases and count noun phrases occur together and that an analysis is needed of a verb phrase such as <u>disappear</u> which can work both in combination with mass noun phrases and count noun phrases. Transitive verbs also often occur with both count noun phrases and mass noun phrases, like

(6b) John drank much water

and this must also be accounted for in our proposal.

In this section we present in outline how a combined system of mass terms and count terms can be constructed. The starting point is that we have two models at hand: One based on a boolean algebra $\langle A,\times,+,^{-},0,1\rangle$ and one based on a non-empty set I[11]. The first can be used for the pure mass term part, the second for the pure count term part. Our first idea is to take the union of the two models in some direct way - either the union of A and I (where A and I do not have to be disjoint, $A \subset I$ is one possible proposal), or the union of A and $\mathcal{P}(I)$. This is problematic for two reasons.

In connection with mass terms homogeneous verb phrases and inhomogeneous verb phrases got denotations at different levels, in combination with count noun phrases they get denotations at the same level. So if we take the union in either of the two ways the result is that one of the two classes of verb phrases will get denotations at two levels at the same time.

The second reason is that taking e.g. A as a subset of I may be interpreted as identifying quantities with individuals which is questionable, as already Parsons (1970) pointed out, because the identity conditions seem to differ for the two classes. We may tolerate many changes in the material basis of an individual and still claim that it is the same individual, but we are not so willing to tolerate changes in quantities. Of course this objection turns on the relationship between the language and the outer world and it is therefore not crucial according to the principles stated in the introduction. But it may be important if we want to exted our analysis to more complex sentences which take time references into consideration.

We therefore propose to take as basis for a model both a boolean algebra $\langle A,+,\times,^{-},0,1\rangle$ and a set I. A homogeneous verb phrase, like disappear, shall then denote an ordered pair $\langle d_1,d_2\rangle$ where $d_1 \in A$ (an element in A) and $d_2 \subset I$ (a subset of I). A mass noun phrase like much gold shall, as before, denote a subset of A and a count noun phrase like many coins shall denote a subset of $\mathcal{P}(I)$, as usual. The conjunction much gold and many coins shall have the denotation $\{\langle x,y\rangle \in A\times\mathcal{P}(I): x \in \|\text{much gold}\|$[12] and $y \in \|\text{many coins}\|\}$. The sentence

(15) Much gold disappeared

is then true just in the case $d_1 \in \|\text{much gold}\|$, while the sentence (14) is true if $\langle d_1,d_2\rangle \in \|\text{much gold and many coins}\|$ which is the same as $d_1 \in \|\text{much gold}\|$ and $d_2 \in \|\text{many coins}\|$.

Inhomogeneous verb phrases, such as weighed less than three kilos, shall denote an ordered pair where the first element is a subset of A and the other element is a subset of I. The first element is needed for sentences like:

(5) The cheese that disappeared weighed less than two kilos

where we still want the cheese that disappeared to denote the element $d_1 \times c$ in A (where c is the element that cheese denotes).

In the introduction we proposed to regard the distinction between mass terms and count terms as a distinction between different kinds of noun occurences. For a noun having both types of occurences, like water, this can be captured as follows: The noun water is given a denotation of the same type as disappear, i.e. an ordered pair $\langle w_1, w_2 \rangle$ where $w_1 \in A$ and $w_2 \subset I$. Much is a mass determiner, i.e. much water is a mass noun phrase, so much shall denote a function which assigns a subset of A to w_1, ignoring w_2. Many, on the other hand, makes count noun phrases like many waters, so the denotation of many will take w_2 as argument and produce a set of subsets of I.

Homogeneous predicative adjectives like red shall have denotations of the same type as disappear and water, ordered pairs where the first element is an element in A, the second element a subset of I. While inhomogeneous predicative adjectives shall denote subsets of I only.

John drank has already been proposed to denote an element in A. Since John naturally denotes a member of I it seems appropriate to let drink denote a function which to elements in I yields members of A. A sentence like John drank many beers shows that John drank must also denote a subset of I. So it seems most natural to let drink denote a function which to elements in I yields ordered pairs where the first element is a member of A, the second element a subset of I. Notice that restricted to the countable domain this gives the normal interpretation since $(2^I)^I$ is isomorphic to $2^{I\times I}$. If drink denotes $\|drink\|$ and $\|drink\|(i) = \langle x,y\rangle$ we can call x for $(\|drink\|(i))_1$ and y for $(\|drink\|(i))_2$. Then the sentence

(16) Many men drank much water

is true if and only if $\{i\in I:(\|drink\|(i))_1 \in \|much\ water\|\}$ is a member of the set $\|many\ men\|$. (Recall that we have restricted attention to the distributive reading of count noun phrases.)

The question remains whether all transitive verbs shall be given this type of denotation. For a given transitive verb β we can check whether the expression John β passes the homogeneity test. If the answer is yes, this treatment should be appropriate. If the answer is no, the sentence John β much water will not be wellformed according to the Homogeneous Constraint, and β is better taken to denote a function which to elements in I yields subsets of I, or maybe subsets of I and subsets of A. The example:

(17) The water that John drank contained 10 grams of salt

opens for transitive verb denotations which are functions defined on A and I, but we do not yet see how to come from this to the homogeneous contained salt in

(18) Much water contained salt.

Definite descriptions are harder to capture than other noun phrases, because it is not possible on syntactic grounds to classify the noun as a mass term or as a count term. We are thus forced to treat the water and even the water that John drank as ambiguous between two readings. Recall that water has a denotation $\langle w_1, w_2 \rangle$ where $w_1 \in A$ and $w_2 \subset I$. Then in the first reading, the "mass" reading, the water shall denote w_1. In the second reading, the "count" reading, the water will be defined only if w_2 is a singleton set, and in that case it shall denote the unique member of w_2. So with this reading the water that John drank shall have a denotation if and only if there was a water that John drank and only one such water.

When this model is applied to a fragment of a natural language it is necessary to refine the syntactic rules to take care of the differences between mass noun phrases and count noun phrases, which expressions they can occur together with and so on. The details are lengthy and will not be given here[13], we only give some examples. The simple rule

$$N \rightarrow AD\ N$$

needs at least the following refinements. N must be separated in three subcategories $N_{[M]}$-mass nouns, $N_{[C]}$-count nouns, and $N_{[U]}$-the nouns that can have both types of occurences (U for undetermined)[14]. The predicative adjectives must be separated in two subcategories $AD_{[H]}$-homogeneous adjectives, and $AD_{[I]}$-inhomogeneous adjectives. The reformulated rule is then

$$N_{[U]} \rightarrow AD_{[H]} N_{[U]}$$
$$N_{[M]} \rightarrow AD_{[H]} N_{[M]}$$

$$N_{[C]} \rightarrow \begin{cases} AD_{[H]} N_{[C]} \\ AD_{[I]} N_{[U]} \\ AD_{[I]} N_{[C]} \end{cases}$$

The reader may like to formulate the corresponding semantic rules from what is said earlier about the denotation types of the different categories.

Determiners must also be separated in two subcategories: $Det_{[M]}$-mass determiners, and $Det_{[C]}$-count determiners. The rule

$$NP \rightarrow Det\ N$$

is then refined to

$$NP_{[M]} \rightarrow \begin{cases} Det_{[M]} N_{[U]} \\ Det_{[M]} N_{[M]} \end{cases}$$

$$NP_{[C]} \rightarrow \begin{cases} Det_{[C]} N_{[U]} \\ Det_{[C]} N_{[C]} \end{cases}$$

Further refinements taking number into consideration to get the right analysis of the two noun phrases <u>some water</u> and <u>some waters</u> may also be necessary.

Rules like

$$S \rightarrow NP\ VP$$

where the category NP occur to the right must be refined to take care of the Homogeneous Constraint. Notice that this is extended to cover noun phrases which are a conjunction of a mass noun phrase and a count noun phrase as this example illustrate:

(19) *Much gold and many coins weighed two grams.

There are some connections between mass terms and count terms which we have not commented on yet. Parsons (1970) starts his treatment of mass terms with the example sentence

(20) My ring is gold

I have problems with ascribing meaning to this sentence because the direct translation into Norwegian, my native tongue, ringen min er gull is not well-formed. One would have to say: ringen min er av gull (is of gold) or ringen min er laget av gull (is made of gold). That (20) may be odd in english too is indicated by a remark in Quine (1960): "... things are red, stuff alone is water" (p 92). But since (20) is proposed it must be well-formed for some speakers of English and hence demand an explanation. We offer two. The first is simply to give gold the same type of denotation as water, both an element in A and a subset of I and say that the sentence is true if the element in I denoted by my ring is a member of this set.

The other approach is an attempt to explain the relationship between the things that are gold and the quantities that are gold, and it resembles a proposal by Link (1982). We introduce a function f_t[15] mapping each element in I to a member of A, intuitively mapping each object to the quantity that constitutes it. If now $\|gold\|$ is a member of A, $\|is\ gold\|$ can be given the denotation $\langle \|gold\|, \{i \in I : f_t(i) \leq \|gold\|\} \rangle$. The awkwardness of this approach is that my ring has to be wholly made of gold, it cannot e.g. possess a pearl, if the sentence (20) shall be true. However, with some ingenuity it should be possible to get around this particular problem.

We conclude this section with one further semantic observation. Count noun phrases can occur together with homogeneous expressions and in combination with inhomogeneous expressions. But there is a semantic difference between the two types of occurrences, at least concerning occurences exemplified by John and Harry, the men, and three men.

(21) John and Harry weighed 100 kilo

can either mean that each of them weighed 100 kilo or that their combined weight was 100 kilo. Now this ambiguity is only possible if the verb phrase is inhomogeneous. If the verb phrase is homogeneous, the two readings are equivalent. Thus, if John disappeared and Harry disappeared then John and Harry disappeared, and the other way round. The possible ambiguity between the two boys disappearing at the same time or at different times is irrelevent for the simple sentence John and Harry disappeared. It simply does not carry that much information.

In connection with mass terms we established a test for checking homogeneity (CUM and DISTR). With minor changes this can be used with count terms to. What is needed is to exchange mass with count everywhere in the introduction to the test and allow plural forms in the inferences, e.g.

CUM' The β α/α-s δ

The not-β α/α-s δ

∴ The α-s δ

If we want to test disappeared for CUM' then this can be done with this inference, e.g.

The red car/cars disappeared

The not-red car/cars disappeared

∴ The cars disappeared

This test does not explain why some expressions give rise to ambiguities while others do not. It only gives us a tool to check if an expression give rise to ambiguities. Moreover we can use this tool if we want to try to give a syntactic classification of the homogeneous expressions like we indicated for mass terms. The general scheme seems very similar in connection with count terms. Nearly all intransitive verbs seem to be homogeneous. And transitive verbs seem to be homogeneous in relationship to the direct object, but inhomogeneous in relationship to the subject when the object is quantified. Note that the verb love which was problematic in connection with mass noun phrases seem to be homogeneous with respect to the direct object now, e.g. John loved two girls.

6. Amount terms.

An amount term is composed of a numeral plus a unit of measure (or denomination) (e.g. two kilos) or is an expression of the form less than, more than, at least etc. plus a numeral and a unit of measure (e.g. less than two kilos).

In connection with a mass term it can occur as a part of a determiner (sentence (4)) or as a part of an (inhomogeneous) verb phrase (sentence (5)).

(4) Less than two kilos of cheese disappeared.

(5) The cheese that disappeared weighed less than two kilos.

In this section we extend our semantic model to cover these occurences of amount terms and, in particular, we will show that the sentences (4) and (5) are equivalent. That will be attained if the subset of the boolean algebra denoted by weighed less than two kilos is related to the denotation of the determiner less than two kilos of according to the following formula

(22) $b \in \|$Less than two kilos of$\|(a)$ if and only if
$(a \times b) \in \|$weighed less than two kilos$\|$

Before we go into the details we add the following comment: Parsons (1970) starts his study of amount terms with this example:

(23) Three teaspoons of gold weigh thirty ounces.

This sentence diverges from sentence (4) in at least two respects. While sentence (4) is about one specific quantity of cheese, sentence (23) is about (nearly) all or a typical quantity of gold of a certain size. And the verb phrase in (23) is not homogeneous. This means that the Homogeneous Constraint universal will have to be changed to allow sentence (23). Either this can be done by giving up the idea that mass noun phrases can be given a purely syntactic definition saying that the noun phrase in (23) is not a mass noun phrase, or by changing the constraint so that it says that mass noun phrases do only combine with homogeneous expressions to form sentences with existensial readings. Be that as it may, it would, however, be worthwhile further study to classify which of the earlier classified mass noun phrases which can occur in sentences like sentence (23).

Note that the contrast between the sentences (23) and (4) can also be observed in connection with some types of count noun phrases

(24) Two men disappeared

(25) Two apples weigh thirty ounces.

The verb phrase in (25) is inhomogeneous, and the sentence is not about two definite apples but about (nearly) every pair or one typical pair of apples. So this shows that the distinction is independent of the mass/count distinction. We feel therefore justified here to concentrate on sentence (4) and reserve for later treatment the problems raised by sentence (25).

Let us return to the sentences (4) and (5). Then some questions arise. First what is the meaning of each of the words in the amount term and the measure verb and what kind of denotations shall we give them. The second question arise because there are numerals involved in the amount terms and concerns how much of mathematics shall be regarded as a part of the semantics. I think these questions have in common that they do not have a right or even a best answer. So let us consider some possible answers starting with the second question.

In the boolean framework the concept of measure (or measure function) comes immediately to mind when thinking about amount terms and measure verbs. A (finitely) additive measure μ on the boolean algebra $\langle A,\ldots\rangle$ is a function from A into the nonnegative real numbers such that for any two elements $a,b \in A$

ADD: if $a \times b = 0$ then $\mu(a+b) = \mu(a)+\mu(b)$

(+ to the left here is boolean sum, the + to the right is the standard arithmetic operation of addition.) The application of this to our semantic model should then give

(26) $\|$weighed less than two kilos$\|$ $= \{a \in A : \mu(a) < k\}$

for some measure μ and constant k which are related to the words of the phrase in some well-determined way. It would then follow from this, formula (22) and ADD that $\|$Less than two kilos of$\|$ is monotone decreasing.

This reflects fairly well what is going on when we measure something. But it is somewhat questionable if all mathematical truths which can be derived from an act of simple measurement shall be regarded as semantic truths. To take one example, shall it be regarded as a semantic truth that 2^{10} kilos of salt is more than 1000 kilos of salt? If our goal is a semantics for a natural language, then I think the answer is no. So let a measure instead be a function from A to a primitive set R.[16] We can then study which constraints one should impose on R and the measure functions to validitate some typical inferences involving amount terms. First observe that the relationship in (22) and thereby the equivalence of the sentences (4) and (5) may be formulated independently of the structure of R and the measures.

The inference

(27) The water weighed 3 kilos

∴ The water weighed more than 2 kilos

becomes valid if R is taken to be (at least) a (pre-)ordered set where $\|3\| \geq \|2\|$ (in this ordering). No restrictions on the measures are necessary. But as mentioned earlier we are not sure that (27) shall be taken as valid if 3 and 2 are exchanged with more complex number terms. One way out of this is still to demand that R is ordered but not to assume that $\|2^{10}\| \geq \|1000\|$ or even $\|3\| \geq \|2\|$. Then 3 is greater than 2 or 3 kilos are more than 2 kilos can be taken as an additional premiss in (27) to get a valid inference.

There is probably agreement on the following inference

(28) Two liters of water disappeared

∴ Some water disappeared

It follows from ADD if we demand $\|\text{Two}\| \neq 0$. But the only fact from ADD which is needed here is that there is a special element 0 in R such that $\mu(0) = 0$ (where 0 to the left is the null-element of A). The validity of (28) with two exchanged with more than two will also follow from this if we in addition demand $\|\text{two}\| > 0$.

One of the most striking properties of determiners constructed from amount terms is the monotonicity:

(29) More than five liters of water evaporated

All that evaporated disappeared

∴ More than five liters of water disappeared

As mentioned, this follows from ADD, but it is sufficient that R is ordered and that the measure is monotone:

MON if $a \leq b$ then $\mu(a) \leq \mu(b)$

which of course is a consequence of ADD. MON also implies that quantifiers constructed from amount terms starting with less than (see example (13)) are monotone decreasing. And MON in conjunction with (22) yields the left monotonicity which validitates inferences of the following form:

(30) More than two liters of red water disappeared

∴ More than two liters of water disappeared

The inferences studied until now depend only on the ordering structure of R and the monotonicity (MON) of the measures. Inferences that also need the arithmetic structure of R and (larger parts of) ADD may be constructed, but they seem a bit more "mathematical" in structure

(31) John drank 2 glasses of white wine and one glass of red wine

No white wine is red wine

∴ John drank (at least) 3 glasses of wine

This presupposes that an operation of addition is defined on R, ADD or at least one half of it where = is exchanged with $\geq$, and a "normal" interpretation of the numerals such that $\| \text{one} \| + \| 2 \| = \| 3 \|$. Similar examples can be constructed which needs the other half of

ADD, and it is also possible to construct examples which need formulation of multiplication on R.

We have touched upon some issues concerning natural language inferences involving amount terms and our discussion is far from complete. Another topic beyond the scope of this paper is to relate our treatment of amount terms to Kamp's (1975) approach to comparatives of adjectives.

Then to the question of what the different words shall denote. If we think about what we are doing when we measure or weigh something, or if we start with the sentence (5), the most natural thing seems to be to let weighed denote a measure, that is a function from A to R, and let two kilos denote a measuring result, i.e. an element in R. Then if a multiplication operation $\cdot$ on R is available, two and kilo can both be taken to denote elements in R, and two kilos can denote $\|two\| \cdot \|kilo\|$, less than two kilos denote the subset of R of elements r such that $r \leqslant \|two\ kilos\|$ (where $\leqslant$ is a relation on R) and weighed less than two kilos denote the set of elements in A mapped into this set by $\|weighed\|$.

But this approach leads to problems with sentence (4) because there is no measure verb available. The most natural analysis if one starts with sentence (4) is to let kilo denote the measure, less than two denote the set of elements r in R such that $r \leqslant \|two\|$, and less than two kilos denote the set of elements in A mapped into this set by the measure function $\|kilo\|$. This can also take care of sentence (5). All one has to do is to give weighed less than two kilos the same denotation as less than two kilos.

The weakness of the second approach is that weighed is given no meaning in sentence (5). This is a bit counterintuitive and it does not give any semantic explanation of what is wrong with an expression like weighed more than two liters. Another problem which the first approach manages better is to explain the closer connections that exist between e.g. kilo and ounce than between kilo and liter.

One could think of various compromises between the two approaches. One would be to hide the word weigh somewhere in the syntactic deep structure of sentence (4) and claim that it has been deleted in the derivation of the surface structure. Another way out is to push the problem to the semantics and give kilo a denotation consisting of an ordered pair of an element in R and a measure, and then call upon the right part of the pair in building more complex denotations. In particular, in giving a denotation to (5), we shall never call upon the measure part of kilo since that will be provided by the measure verb weigh.

But this is not the end of our problems. Nothing that we have said so far blocks a sentence such as

(32) ?The water weighed two liters.

One can argue whether it should be blocked on syntactic or semantic grounds.

Our proposal to get around these problems is as follows: Instead of letting measure verbs denote function into one set R, one can let them denote functions into different sets $R_1, R_2, \ldots,$[17] one for each measure verb. If then weigh denotes a function into R_1, kilo shall denote an element in R_1 while the denotation of liter is not in R_1, but in some other R_i.[18] To secure that the denotation of kilo is in the right R_i the syntactic category denomination is equipped with a feature: weight, volume, ... or another measure verb. It is then possible to reject sentence (32) syntactically if also amount terms get features - the same feature as the denomination it contains - and if it is required that the feature of an amount term has to correspond to the measure verb for the combination of them to be well-formed. Where the amount term is a part of a determiner the feature can be regarded as a syntactic counterpart to the stored measure in the semantics. This means that it is not necessary to let a denomination denote both an element in an R_i and a measure since the measure can be derived from the feature in a uniform way.

Instead of giving a set of rules for a fragment of English,[19] let us see how the example sentences (4) and (5) may be treated in a way consistent with what is said above. Observe that of is not given a denotation but is introduced syncategorematically when determiners are formed out of amount terms. That is partly because there are languages (e.g. Norwegian) where of is not realised. It is also assumed that there are only finitely many words except numerals which can be the head of an amount term, and that they can be given a special (logical) treatment.

The syntactic categories used are:

$Den_{[w]}$ - denomination; feature: weight

Num - numeral

$AT_{[w]}$ - amount term; feature: weight

$MV_{[w]}$ - measure verb; feature: weight

HVP - homogeneous verb phrase

IVP - inhomogeneous verb phrase

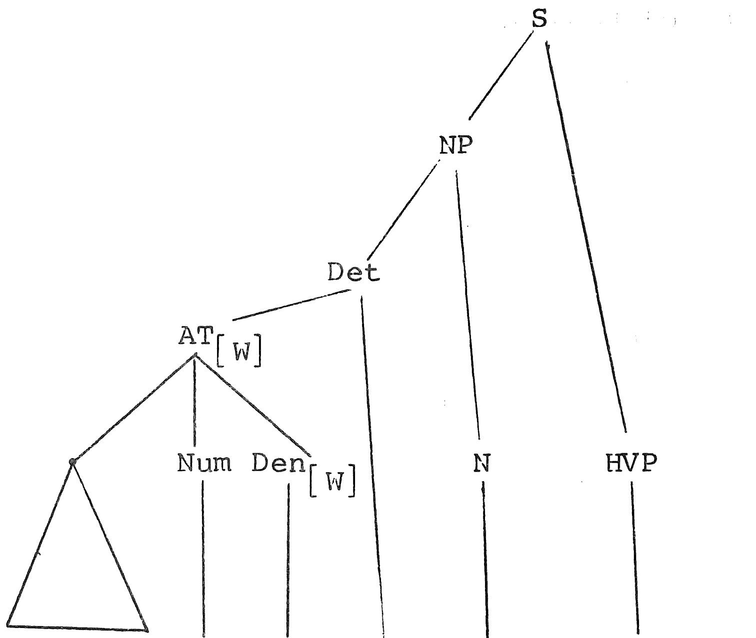
Det - determiner

N - (mass) noun

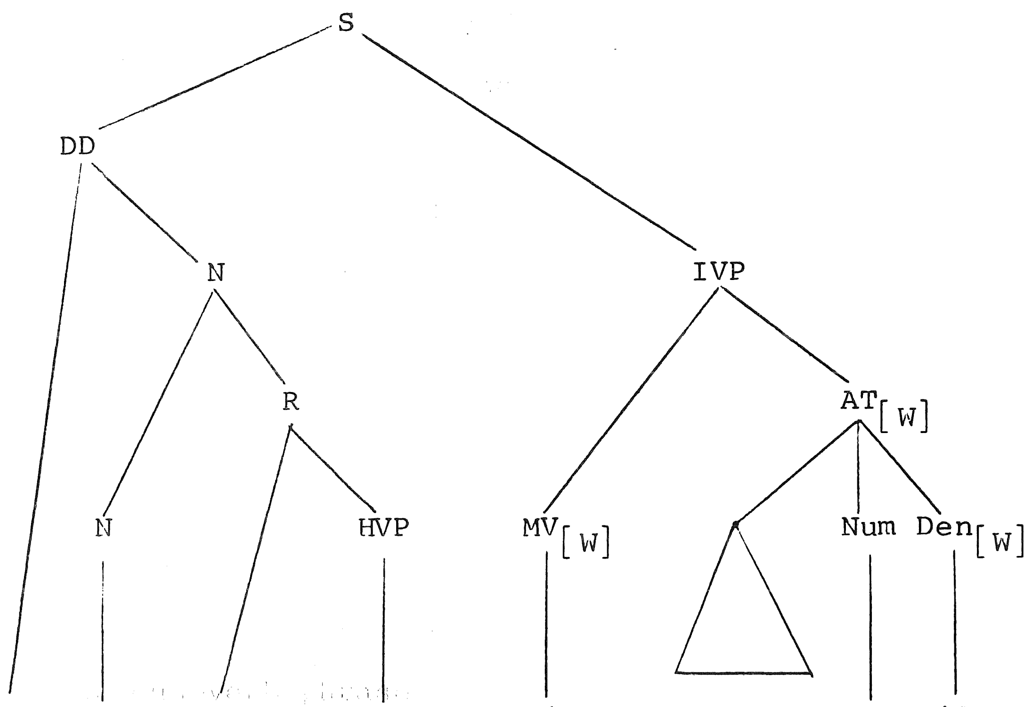
NP - (mass) noun phrase

DD - definite (mass noun) description

The syntax trees are then:



(4) Less than two kilos of cheese disappeared



(5) The cheese that disappeared weighed less than two kilos.

One model for this is based on a boolean algebra $\langle A, \ldots \rangle$. Since only one feature weight is involved, it is sufficient with one additional set R, with at least one binary relation $<$ and one binary function $\cdot$, and one function $\underline{w}$ from A to R. The denotations of cheese, disappeared, and the cheese that disappeared are exactly as in part 3. In addition

$\| [_{Den_{[w]}} \text{kilos}] \|$ is an element in R

$\| [_{Num} \text{two}] \|$ is an element in R

$\| [_{AT_{[w]}} \text{less than} [_{Num} \text{two}] [_{Den_{[w]}} \text{kilos}] \| =$

$\{ r \in R : r < \| [_{Num} \text{two}] \| \cdot \| [_{Den_{[w]}} \text{kilos}] \| \}$

Let us call this last set LTTK. Then

$$\| [_{Det} [_{AT_{[w]}} \text{less than} [_{Num} \text{two}] [_{Den_{[w]}} \text{kilos}]] \text{of}] \|$$

is the function f from A to the set of subsets of A such that

$$b \in f(a) \text{ if and only if } w(a \times b) \in LTTK.$$

Notice here that the measure $\underline{w}$ stems from the feature weight of the amount term (in a uniform way). The rest of the semantics of sentence (4) is now standard. For sentence (5) we require that

$$\| [_{IVP} [_{MV_{[w]}} \text{weighed}] [_{AT_{[w]}} \text{less than} [_{Num} \text{two}] [_{Den_{[w]}} \text{kilos}]]] \| = \{ a \in A : w(a) \in LTTK \}.$$

Since the denotation of the definite description which introduce this sentence is $\| [_N \text{cheese}] \| \times \| [_{HVP} \text{disappeared}] \|$, it is easy to see that the two sentences become semantically equivalent.

One special problem related to amount terms concerns noun phrases like _two cups of coffee_. Our proposal for an analysis is that this phrase is ambiguous between two readings. In the first reading _two cups_ functions as an amount term along the same lines as _two liters_. So _cup_ can be put in the syntactic category _denomination_ with feature _volume_. This reading is the most natural in

(33) John drank two cups of coffee

where it is not necessary that there were two cups present, not even one if John drank from a glass. The second reading is dominant in situations where the presence of two cups are necessary. One such example is

(34) Two cups of coffee are on the table.

In this case the noun phrase is a count noun phrase with _cups_ as its head noun, _of coffee_ functions as a noun modifier, and _two_ is the (count) determiner.

7. Further directions.

Throughout the paper we have listed some problems that need further study. To begin with we gave some reasons for taking mass terms with determiners as basic. But it is an open question if this approach can be extended to (some) occurrences with no determiners. We have not proposed to rewrite drank water as drank some water, because it is not clear how this rewriting can be done uniformly, e.g. it does not extend to water is wet. Another reason is that the two verb phrases drank water and drank some water react differently to time references. They combine with different members of the pair for an hour/in an hour.

This last point shows the importance of extending the analysis to cover tense and time references. Another reason to such an extension is that we want a better analysis of the differences between definite and indefinite mass terms, e.g. ice and the ice. In the sentence Much ice α, we evaluate ice at the time interval indicated by the verb phrase α. While in the sentence Much of the ice that John found α, ice is evaluated at the time indicated by found, and that does not have to be the same time as the action time of α. So what we are talking about in the last sentence need not be ice when it α.

Verb phrases, like disappear, which express a change of state may also get a more realistic analysis, i.e. an analysis which covers more of their semantics, if a more dynamical model is established.

The remarks at the end of part 5 suggest a further study of collective and distributive readings of count noun phrases. We need here an analysis which does not overgenerate i.e. which does only produce a conflict between a collective and distributive reading in the cases where the verb phrase is inhomogeneous. Closely connected to this is the problem of existensial and (nearly) universal collective readings mentioned in part 6, both when mass terms and count terms are considered.

Footnotes.

* This paper is an extract of my (unpublished) cand.real.thesis (Lønning (1982)). I wish to thank my advisor Professor Jens Erik Fenstad for valuable help and encouragements, both in the work of the thesis and this paper. I wish also to thank Helle Frisak Sem and the participants in the Groningen workshop on generalized quantifiers in July 1983 for comments on earlier versions of this paper.

1. What follows is not a satisfactory criticism of all contemporary theories of mass term quantification, but some problems which face everyone who want to study this phenomenon.
2. This is reformulated from Cheng to match better with the other criterion.
3. That this is a proposal for a universal means that is open for amplification and changes (in the sequel).
4. Expression means here the rest of the sentence. That is, it includes e.g. John drank which is not normally regarded as a unit.
5. For more information on boolean algebras see Halmos (1963).
6. Bunt (1981) finds it necessary to single out some exceptions to this. But none of this examples are predicative.

7. *is* is exchanged with *was* to get a sentence generated in the fragment.

8. This is not an independent set of axioms since e.g. A1 can be proved from the other ones.

9. Here one can see how the LA-language is weaker than the first order theory of boolean algebras. The latter one needs only to be extended with one sentence: $\forall x(x \neq 0 \rightarrow \exists y(y \leq x \wedge y \neq 0 \wedge y \neq x))$, to get the theory of non-atomic boolean algebras.

10. This follows because it can be proved that every formula in the corresponding L(GQ)-language with no determiners except *Every* and *Some* is equivalent to a translation of a L-formula. If one extend LA with modal operators $\Diamond$, $\Box$ and the semantics which is the natural combination of LA's semantics and S5's semantics, then the result is a decidable system. Modal monadic first order logic is not decidable on the other hand. One reason for this discrepancy is that there are modal monadic first order formulae which do not correspond to LA+S5 formulae, e.g. formulas of the form $\exists x \Box \phi$.

11. This is not a complete analysis of count noun phrases and their semantics. In particular so-called collective readings of count noun phrases will not be considered. My cand.real.-thesis (Lønning (1982)) includes a logic for treating collective readings of numerals (e.g. *two men*), which is formulated in a way which allows a complete axiomatization with respect to the semantics.

12. Here and for the rest of the paper $\| \cdot \|$ will be a function which to natural language expressions gives their formal denotations.

13. In Lønning (1982) this is carried out.

14. If it is true that every noun can have both types of occurences then the lexical entries of $N_{[C]}$ and $N_{[M]}$ will be empty and the following rule accordingly shorter.
15. The <u>t</u> in f_t indicates the dependence of time.
16. Parsons (1970) lets measures be relations. That is uneccesary since nothing can have more than one weight.
17. An interesting question is if it is sufficient with finetly many such R_i's.
18. A numeral like <u>two</u> must have denotations in each R_i.
19. In Lønning (1982), I gave such rules and formulated a formal language which had semantics in terms of $\langle A,+,\cdot,\ldots\rangle$ and R. I also showed there that if R is supposed to be an ordered field and the measures satisfies the ADD requirement, a complete axiomatization with respect to the semantics can be given.

References.

Barwise, J. and Cooper R. (1981) "Generalized quantifiers and natural language", Linguistics and Philosophy, Vol. 4, No 2, D. Reidel, Dordrecht, Holland.

Burge, T. (1972) "Truth and Mass Terms", The Journal of Philosophy, Vol. LXIX, No 10.

Bunt, H. (1979) "Ensembles and the Formal Semantic Properties of Mass Terms" in Pelletier, F. J. (ed): "Mass terms: Some philosophical problems", D. Reidel, Dordrecht, Holland.

Bunt, H. (1981) "On the why, the how, and the whether of a count/mass distinction among adjectives" in Groenendijk, J.A.G., Janssen, T.M.V., Stokhof, M.B.J. (eds.): "Formal methods in the study of language", Mathematical Centre Tracts 135, Amsterdam.

Cheng, C.-Y. (1973) "Comments on Moravcsik's Paper" in Hintikka, K.J.J., Moravcsik, J.M.E., Suppes, P. (eds): "Approaches to Natural Language", D. Reidel, Dordrecht, Holland.

Halmos (1963) "Lectures on Boolean Algebras", D. Van Nostrand, Princeton, New Jersey.

Huges, G.E. and Gresswell, M.J. (1972) "An Introduction to Modal Logic", Methuen and co, London.

Kamp, J.A.W. (1975) "Two theories about adjectives" in Keenan, E.L. (ed): "Formal semantics of natural language", Cambridge University Press, Cambridge.

Leonard, H.S. and Goodman, N. (1940) "The calculus of individuals and its uses", The Journal of Symbolic Logic, Vol. 5, No 2.

Link, G. (1982) "The logical analysis of plurals and mass terms: A lattice-theoretical approach", Unpublished manuscript.

Lønning, J.T. (1982) "Kvantiteter og kvantorer. Om formell semantikk for mengdetermer ("mass terms")". Unpublished cand.real.thesis, Dept. of mathematics, University of Oslo.

Montague, R. (1973a) "The Proper Treatment of Quantification in Ordinary English" in Hintikka, ... (see Cheng (1973)).

Montague, R. (1973b) "Comments on Moravcsik's Paper" in Hintikka, ... (see Cheng (1973)).

Parsons, T. (1970) "An analysis of mass terms and amount terms", Foundations of Language, Vol. 6, No 3.

Pelletier, F.J. (1975) "Non-Singular Reference: Some Preliminaries", Philosophia, Vol. 5, No 4

Quine, W.V.O. (1960) "Word and Object", M.I.T. Press, Cambridge, Massachusetts.

QUANTIFIER SCOPE AND COREFERENTIALITY *

Contents

Helle Frisak Sem

0. Preface.

In this article we discuss questions of quantifier scope and anaphoric relations between noun phrases. We approach the topic through a study of two different systems for representing language meaning, one the system using the "storage mechanism" developed in R. Cooper: "Quantification and Syntactic Theory" (Cooper 1983), the second the theory of Discourse Representation Systems as introduced in H. Kamp: "A Theory of Truth and Semantic Representation" (Kamp 1981).

We start out by giving a modified treatment of the Cooper systems, in particular, we eliminate the use of multivalued functions in the semantic interpretation. The storage mechanism is successful in treating many questions of quantifier scope in natural languages, but there are examples which the system does not seem able to handle. We conclude this part of our study with a systematic discussion of limitations of the storage method.

Many of these limitations can be overcome by using the theory of Discourse Representation Systems. We extend the analysis of Kamp's (Kamp 1981) to cover the syntactic fragment of Cooper (1983) and are thus able to obtain a satisfactory treatment of some of the examples that could not be handled by the storage method.

The way the Discourse Representations describe submodels of a model for Discourse Representation Systems bear clear resemblances to the way event-types describe partial models in the Situation Semantics presented by J. Barwise and J. Perry in "Situations and Attitudes" (Barwise and Perry 1983). We end the paper with an illustration of how the theory of Discourse Representation Systems can be applied in the construction of meaning relations for linguistic expressions in Situation Semantics.

We assume familiarity with (Cooper 1983) and (Kamp 1981) even if, in principle, this article is self-contained as the systems are presented in full details. For reasons of space few examples are discussed. A more comprehensive treatment is found in my cand.scient thesis "Kvantifikasjon, syntaks og semantikk" (Sem 1983) (in Norwegian) which was written under the supervision of Professor Jens Erik Fenstad, and I am indebted to him for encouragement and valuable advice during the work on my thesis.

1. The Cooper stores.

1.1.

The aim of this section is to study "the Cooper stores", their possibilities and limitations. As a starting point for our investigations we take Fragment 6 from Cooper (1983). This is the simplest non-transformational fragment rich enough to contain the full use of the storing mechanisms with relative clause formation and embedded questions.

For reasons of clarity and ease of readability some simplifications both of the phrase structure rules in the syntax and of the technical apparatus of the semantics will be made. Thus, a better basis for further extensions, especially with regards to interaction between the different storage applications, is obtained. Though the notation and to some extent the technical apparatus undergoes several changes, the main virtues of the storage technique remain unchanged. The limitations discovered for the present version will also apply to the original one, as far as the original contains the possibilities of the present version. That a fragment with the syntax of Fragment 6 and semantics simplified similarly as the present fragment is syntactically and semantically (weakly) equivalent to the present fragment, is proved in Sem (1983, I.4).

The syntactic simplifications are all motivated by the desire to avoid the following kind of syntactic ambiguity. I want to avoid any two SD's of the form $[\xi]_A, [\xi]_B$ where A,B are syntactic categories, $A \neq B$, and ξ is a string of SD's. The induction base of the syntax then defines the basic elements of the set SD, while the phrase structure rules defines a function from $SD \times SD$ into SD (provided we regard 'that' and $\emptyset$ as basic SD's in the case of $\bar{S} \rightarrow$ that S and NP $\rightarrow$ R).

The use of multivalued functions in Cooper (1983) is motivated partly by free pronouns, partly by semantic ambiguity of syntactically unambiguous phrases (Cooper 1983, Ch. II, 2.5). The full interpretation of a phrase is the set of all the different readings of the phrase. The point is to make one phrase correspond to exactly one interpretation. During the interpretation process, only one of the readings for each constituent phrase can be used to calculate one of the readings for a compound phrase. Thus, the process itself is obscured by a notation designed only for denoting the entire set of meanings for a fully interpreted phrase. The multi-valued notation can therefore be dispensed with altogether without any loss of compositionality. The set of obtainable readings will remain the same for every phrase. To further increase readability, the κ and ρ functions will not be used.

In addition to giving a simpler and more uniform treatment of the WH-clauses, the present syntax definition provides a fragment that follows a somewhat stricter principle of compositionality than the original fragment. A further discussion of the fragments with respect to the principle of compositionality and a definition of an equivalent fragment following a strict formulation of the principle is found in Sem (1983).

Som restrictions will then be removed from the semantics to get a fragment that allows more complicated use of the storage mechanisms. Although the storage technique turns out to meet most of the needs of this more extended fragment, one major class of problematic constructions emerges - a class comprising the constructions known in the literature as "the donkey sentences".

The storage technique is used to obatin the correct relationship between the referents of the various NP-phrases (including gaps and WH-terms). Different quantifier scopes, de re/de dicto readings, coreferentiality between antecedent and anaphoric pronouns and finally reference for gaps in relative clauses and questions are obtained by means of storage. The main idea behind the storage technique is to put the intension of a NP-term on ice for a while, in order to give it wider scope. When appropriate, the stored NP-intension is lifted out of the store to get scope over the phrase now composed. We will return to the actual definitions of these mechanisms later on.

There are essentially two kinds of storage: ordinary NP-storage and WH-storage. The two should not interfere. As a technical means to carry out the storage and retrieving processes, we let the intensions depend on certain terms of countable sequences of elements from the domain E, and distinguish between the two storage types by means of a standard trick: the one storage type use the odd terms of the sequence, the other the even terms. This approach differs from the approach in Cooper (1983) where the sequence is used exclusively for NP-storage and -retrieval while WH-storage makes use of a special element as a flag in the store (Cooper, Ch. IV, 3).

1.2. Introduction to Fragment 6^{II}.

The simplified version of Fragment 6, Fragment 6^{II}, consists of:

Syntax:
- Lexicon
- Induction base
- Recursion clauses (Phrase structure rules)
- (Exclusion clause)
- Morphology.

Semantics: - Denotations for each word in the lexicon

- Denotations and interpretations for the base
- Recursive rules for interpreting compound phrases
- Rules for the storage mechanisms: opening, storage, retrieval.

The basic building blocks of the semantics are as usual a set E (of entities), a set W (of possible worlds) and the set of truth values, $\{0,1\}$, from which we build a possible world semantics for the fragment.

<u>Notation 1.</u>

For each Structural Description (SD) α defined by the syntax[1]), we want to define interpretations of the form

$$\langle \|\alpha\|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} \rangle \qquad 0 \leq n,\ n \in \mathbb{N} \quad i_j \in \mathbb{N} \ \text{for} \ j = 1, \ldots, n$$

where

$\sigma \in E^{\omega}$, an infinite (countable) sequence of elements of E

$w \in W$

$\|\alpha\|_{\sigma,w} = \|\alpha\|(\sigma)(w)$ is a <u>denotation</u> for α with respect to the sequence σ and world w

$\|\alpha\|_{\sigma} = \|\alpha\|(\sigma)$ is an <u>intension</u> (function with domain W) for α wrt. the sequence σ. The intension $\|\alpha\|_{\sigma}$ may depend on σ_k either because the storing of a NP-intension in a binding operator has left an opening, or because there is an opening created from a pronoun or a gap on the k-th term.

$(NP_j)_{i_j}$ is a stored binding operator for the NP NP_j. We follow the convention of letting the category name denote a SD of the category in question, the j is just a convenient way to enumerate the stored NP's in a general case, the index i_j indicates which term in σ the binding operator is to operate on.

$(NP_1)_{i_1},\ldots,(NP_n)_{i_n}$ is called the store. If $n = 0$, the store is said to be empty.

For convenience, we will in the following use the terms storage, binding, binding operator, retrieval for the general NP-storage terms, and use the prefixes NP and WH for the ordinary NP-case and WH-case respectively.

$\mathcal{N}$ will be used to denote a variable over NP-intensions,

$$\mathcal{N}:W \rightarrow \{X|X \subseteq \wp(E)\} \qquad \text{NP-int} = \{X|X \subseteq \wp(E)\}^W$$

p will be used to denote a variable over propositions,

$$p:W \rightarrow \{0,1\} \qquad \text{Prop} = \{0,1\}^W$$

The following definitions will be useful:

Definition 1: An intension $\|\alpha\|_\sigma$ for a SD α is said to depend on the k-th term in the sequence σ if $\|\alpha\|_\sigma$ varies with the value assigned to σ_k. Similarly for denotations.

Definition 2: α is called a WH-term if α is a SD of category NP not on the form $[R]_{NP}$, and the leftmost word in α is a WH-word (who, what or which in Fragment 6).

Definition 3: α is called a WH-phrase if α is a SD on the form $[NP\ S]_X$, where NP is a WH-term and X is some syntactic category.

The rules for interpreting a compund phrase on the basis of its constituents will all be on the following form:

If α is a SD on the form $[AB]_C$, where

$$\langle \|A\|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} \rangle$$

$$\langle \|B\|_{\sigma,w}, (NP_{n+1})_{j_1}, \ldots, (NP_{n+m})_{j_m} \rangle \qquad 0 \leq n,m$$

are interpretations of A and B respectively, and

i) $i_1, \ldots, i_n, j_1, \ldots, j_m$ are even numbers (possibly i_ℓ an odd number).

ii) $\{i_1, \ldots, i_n\} \cap \{j_1, \ldots, j_m\} = \emptyset$

iii) If $\|A\|_{\sigma,w}$ depends on $\sigma_{k_1}, \ldots, \sigma_{k_\ell}$, then $\{k_1, \ldots, k_\ell\} \cap \{j_1, \ldots, j_m\} = \emptyset$

iv) If $k_{1_1}, \ldots, k_{a_h}$ and $k_{b_1}, \ldots, k_{b_\ell}$ are odd terms in σ that $\|A\|_{\sigma,w}, \|B\|_{\sigma,w}$ are dependent on, respectively, then $\{k_{a_1}, \ldots, k_{a_h}\} \cap \{k_{b_1}, \ldots, k_{b_\ell}\} = \emptyset$

v) If i_ℓ is an odd number for a ℓ, $0 \leq \ell \leq n$, then $\|B\|_{\sigma,w}$ depends on σ_{i_ℓ}

then

$$\langle \|[AB]_C\|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n}, (NP_{n+1})_{j_1}, \ldots, (NP_{n+m})_{j_m} \rangle$$

is an interpretation of α, where .....

($(NP_\ell)_{i_\ell}$ must be retrieved and removed from the store if i_ℓ is odd).

As mentioned above, WH-binding operators and NP-binding operators are distinguished by means of the difference between the odd and even numbers. The conditions i)...v) are safety mechanisms, where:

i) prevents further storage of WH-binding operators
ii) prevents more than one binding operator for each term in σ
iii) prevents backwards binding of pronouns
iv) prevents more than one gap to be bound by the same WH-binding operator
v) ensures that if the first constituent, A, is a WH-phrase, then there is a gap in the second constituent, B, that will be bound in the composing process.

All the safety mechanisms will not be required in every rule, and redundant mechanisms will be omitted.

1.2.1. <u>The storing mechanisms: some preliminary reflections</u>.

Before giving the exact definition of the simplified fragment, some questions concerning the storage mechanisms and their interaction must be considered. What restrictions should be made? How general should the mechanisms be - what restrictions should be inherited from the definition of the mechanisms and what restrictions should be imposed on the rules? The answers given here are by no means exhaustive. The guiding principle is to make the basic definitions and rules as general and independent as possible to get a base that allows for flexibility. Specific restrictions may then be imposed on the combination rules as demonstrated above, or given as constraints. These imposed restrictions and constraints can easily be changed to accommodate particularities proper to the different languages.

In addition to the restrictions concerning the manipulation of binding operators and openings mentioned in the general scheme for phrase-structure rules, a choice regarding the possibility of the following options must be made before defining the basic storage mechanisms:

I) Should more than one gap be accepted in one phrase?

II) In the case of NP-storage, should the NP-interpretations be allowed to

1) be an open pronoun exclusively?

2) be an open gap exclusively?

3) contain open pronouns (as proper parts)?

4) contain open gaps (as proper parts)?

5) have WH-binding operators in the store?

6) have NP-binding operators in the store?

Point II.4 is only relevant if the answer to I) is affirmative. In that case, a decision concerning phenomena like crossover dependencies must also be made. In the definition of Fragment 6, Cooper gives an affirmative answer to I)[2). Concerning NP-storage, Cooper requires the NP-interpretations to be stored to have empty store and no openings. Thus, storage of NP-terms with one of the features 1),...6) is prevented.

The aim of NP-storage is to change NP-scope by storing an NP-intension and thereby postponing the combination of the NP-intension[3) with the intension of the rest of the phrase of which it is a constituent until suitable. Consequently there should be a proper NP-intension to store. The openings serve to mark off where a possible stored intension is to have effect. When stored, a NP-intension will leave behind an opening exactly similar to a pronoun opening. At retrieval, these two kinds of openings will not be distinguished. The openings themselves are in a sense intensionless. At retrieval of a corresponding binding operator from the store, they are to be filled with the stored intension.

Nor have the WH-terms ordinary NP-intension, they are to be functions on propositions (S-intension) as well as indicating the possible NP-intensions for the gap.

The position taken here is therefore not to allow the NP-storage of pure openings and WH-terms, that is, a negative answer is given to point 1), 2) and 5). For the moment, nested storage, point 6), will also be prevented.

1.2.2. Storage, opening and retrieval.

Definition 4: An intension depending exclusively on the value assigned to a certain term of the sequence σ is called an opening at that term.

As opposed to Cooper (1983), pronouns and gaps will here be treated in a uniform manner. The two opening types will be distinguished by letting openings at odd terms of σ be gap openings while openings at even terms will be pronoun openings. In addition to giving a basis for uniform formulation of the two storage types, this approach enables us to avoid the special NP-intension reserved for marking off the gaps in the store (Cooper 1983, Ch. IV, 3).

The definitions of the storage machinery will also differ from the definitions in Cooper (1983), especially in the case of WH-storage (Cooper 1983, Ch. IV, 3.2). The NP-storage definition from Cooper (1983) will be used as a pattern, and we make the following general definitions:

Definition 5: A binding operator is a function constructed from a NP-intension, with the domain $(\{0\}^W)^{E^\omega}$, replacing a certain term, say σ_i, of each element from E^ω by a NP-denotation obtained from the NP-intension. The binding operator is said to be at the i-th term of σ. We denote such a binding operator by $(NP)_i$, where NP denotes the stored NP-intension.

Definition 6: Storage is the process of taking the intension from the denotation part of an NP-interpretation and use it to form a binding operator in the store at some term of σ, say σ_i. After storage, the denotation part of the NP-interpretation may be empty or an opening at i.

Definition 7: Retrieval is the process of applying a stored binding operator to the intension taken from the denotation part of the interpretation to give a new denotation part and remove the binding operator from the store.

With the formulation of the storage mechanisms to be presented below, the two storage types can be regarded as a weaker and a stronger kind, respectively. NP-storage, the weaker kind, leaves behind an opening, and allows freedom with respect to the moment of retrieval, while the stronger kind, WH-storage, leaves the denotation part empty, and retrieval is obligatory at fusion. In addition, the weaker one yields a sentence intension from a sentence intension while the stronger one yields a function from E to the set of sentence intensions.

To fulfil the conditions of the definitions and be in accordance with the principle of compositionality, the denotations for

WH-words must be chosen carefully. Intuition does not give any clear support to the choice of meanings for WH-words. To choose E or some subset of E as denotations does not seem too far fetched. In this context, no distinction between amimate/inamimate etc. is made, so all the WH-words will have E as denotation. In a more sophisticated fragment such a distinction could be implemented by choosing the appropriate subset of E as denotation for each WH-word. Note that the WH-word <u>which</u> will have only one denotation though it is listed under two categories in the lexicon. The two SD's $[\underline{\text{which}}]_{NP}$ and $[\underline{\text{which}}]_{Det}$ will get different interpretations by lexical insertion, corresponding to their different syntactical function.

I put forward two alternative proposals for building interpretations for NP-terms resulting from lexical insertion of WH-words:

a) <u>Guiding principle</u>: Every NP-denotation is a set of subsets of E.

This is the usual claim rising from the principle of compositionality. SD's of the same category should have set-theoretically the same kind of denotation in order to make the semantical machinery work in a uniform way. WH-terms stand in a class by themselves as NP-terms. During the interpretation process, they are always distinguished and treated differently from other NP-terms. A minor reduction in the guiding principle of alternative a) allows a somewhat simpler solution:

b) <u>Guiding principle</u>: An ordinary NP-denotation is a set of subsets of E, a WH-term denotation is a subset of E.

A final decision concerning the two alternatives will not be taken here. I will give rules for both alternatives. The semantics will not be affected by the choice other than in the rule for lexical insertion of WH-words and in the rule for WH-storage.

1.3. Fragment 6^{II}.

1.3.1. Syntax

Lexicon:

- NP: John, Mary, Bill, Sam, Leslie, Kim, Chris, ..., he, who, what, which
- N: woman, man, fish, unicorn, centaur, ...
- VP: runs, ...
- V: loves, admires, seeks, needs, believes, knows, says, thinks, wonders, ...
- Det: a, every, no, which

Induction base

0. Lexical insertion: If α is listed under X in the lexicon, then $[\alpha]_X \in SD$

   $[\ ]_{NP} \in SD$

Phrase structure rules.

1. $S \rightarrow NP\ VP$

2. $VP \rightarrow V \begin{Bmatrix} NP \\ \bar{S} \\ R \end{Bmatrix}$

3. $\bar{S} \rightarrow$ that S

4. $NP \rightarrow \begin{Bmatrix} NP & R \\ Det & N \\ R & \end{Bmatrix}$

5. $N \rightarrow N \begin{Bmatrix} R \\ \bar{S} \end{Bmatrix}$

6. $R \rightarrow$ NP S

In rules 1 and 2, NP may be headless relative, in rule 2, R functions as an embedded question, in rule 4, first alternative, R functions as a non-restrictive relative, and in rule 5, R and $\bar{S}$ function as restrictive relatives. Rule 3 creates both sentence objects and relative clauses starting with <u>that</u>, rule 4, third alternative creates a headless relative from a relative (WH-) clause, rule 6 creates all relative clauses starting with a WH-word.

<u>Exclusion clause</u>: Nothing is a member of SD unless resulting from applications of the rules 0-6.

<u>Morhology</u>: If α is a SD, then the result of applying rule 7 to α represents a syntactically well-formed expression of the fragment.

7. <u>Pronoun case-marking</u>: Change every occurrence of $[\underline{he}]_{NP}$, $[\underline{she}]_{NP}$ not immedately below S in a tree representation of α to $[\underline{him}]_{NP}$, $[\underline{her}]_{NP}$, respectively.

1.3.2. <u>Semantics</u>.

The semantics is based on a <u>model</u> consisting of:

A non-empty set E - the universe

A non-empty set W - the set of possible worlds

$\{0,1\}$ - the set of truth values

<u>Denotations</u> for the words in the <u>lexicon</u>:

NP: i) If α is <u>John</u>, <u>Mary</u>, ..., <u>Chris</u> there is an $e \in E$ for each α such that $\|\alpha\|_{\sigma,w} = e$

ii) For any $e \in E$, $\|he\|_{\sigma,w} = e$ is a denotation for <u>he</u>

iii) If α is <u>who</u>, <u>which</u>, <u>what</u>, $\|\alpha\|_{\sigma,w} = E$

N,VP: If α is listed under N or VP, then $\|\alpha\|_{\sigma,w} = f(w)$, where f is a given function for each α such that $f: W \to \{X \mid X \subseteq E\}$

V: i) If α is
a) <u>loves</u>, <u>admires</u>
b) <u>seeks</u>, <u>needs</u>
c) <u>believes</u>
d) <u>says</u>, <u>thinks</u>
e) <u>wonders</u>

then $\|\alpha\|_{\sigma,w} = f(w)$ where f is a given function for each α such that

a) $f: W \to \{X \mid X \subseteq E \times E\}$

b) $f: W \to \{X \mid X \subseteq E \times \text{NP-int}\}$

c) $f: W \to \{X \mid X \subseteq E \times (\text{Prop} \cup E)\}$

d) $f: W \to \{X \mid X \subseteq E \times \text{Prop}\}$

e) $f: W \to \{X \mid X \subseteq E \times \mathcal{P}(\{q \mid q: W \to \text{Prop}\})\}$ 4)

ii) $\|\underline{\text{know}}\|_{\sigma,w} = f(w)$, $g(w)$ or $h(w)$, where f,g,h are given functions such that

$f: W \to \{X \mid X \subseteq E \times E\}$

$g: W \to \{X \mid X \subseteq E \times \text{Prop}\}$

$h: W \to \{X \mid X \subseteq E \times \mathcal{P}(\{q \mid q: W \to \text{Prop}\})\}$

Det: For given $A \subseteq E$, then

$$\| \underline{a} \|_{\sigma,w}(A) = \{X \subseteq E \mid X \cap A \neq \emptyset\}$$

$$\| \underline{every} \|_{\sigma,w}(A) = \{X \subseteq E \mid A \subseteq X\}$$

$$\| \underline{no} \|_{\sigma,w}(A) = \{X \subseteq E \mid X \cap A = \emptyset\}$$

$$\| \underline{which} \|_{\sigma,w} = E$$

Induction base:

0. If α is listed under X in the lexicon

$$\| [\alpha]_X \|_{\sigma,w} = \| \alpha \|_{\sigma,w}$$

unless

i) α is listed under NP, $\alpha \neq$ who, what, which, then

$$\| [\alpha]_{NP} \|_{\sigma,w} = \{X \subseteq E \mid \| \alpha \|_{\sigma,w} \in X\}$$

For alternative a): If α = who, what or which, then

$$\| [\alpha]_{NP} \|_{\sigma,w} = \{\| \alpha \|_{\sigma,w}\}$$ [5)]

ii) α is love or admire, then

$$\| [\alpha]_V \|_{\sigma,w} = \{\langle x, \mathcal{N} \rangle \mid \{y \mid \langle x,y \rangle \in \| \alpha \|_{\sigma,w}\} \in \mathcal{N}(w)\}$$

iii) α is believe, then

$$\| [\alpha]_V \|_{\sigma,w} = \{\langle x, \mathcal{N} \rangle \mid \{y \mid \langle x,y \rangle \in \| \alpha \|_{\sigma,w}\} \in \mathcal{N}(w)\} \cup \{\langle x,p \rangle \mid \langle x,p \rangle \in \| \alpha \|_{\sigma,w}\}$$

iv) α is know, then

$$\| [\alpha]_V \|_{\sigma,w} = \{\langle x, \mathcal{N} \rangle \mid \{y \mid \langle x,y \rangle \in f(w)\} \in \mathcal{N}(w)\}$$

or

$$\| [\alpha]_V \|_{\sigma,w} = \{\langle x,p \rangle \mid \langle x,p \rangle \in g(w)\}$$

or

$\| [\alpha]_V \|_{\sigma,w} = \{ \langle x,Q \rangle \mid \langle x,Q \rangle \in h(w) \}$ where Q varies over subsets of $\{q \mid q:W \rightarrow \text{Prop}\}$

v) α is <u>which</u> and X is Det, then for given subset $A \subseteq E$, then

For alternative a):

$\| [\underline{\text{which}}]_{Det} \|_{\sigma,w}(A) = \{ X \subseteq E \mid X = A \cap \| \underline{\text{which}} \|_{\sigma,w} \}$

For alternative b):

$\| [\underline{\text{which}}]_{Det} \|_{\sigma,w}(A) = A \cap \| \underline{\text{which}} \|_{\sigma,w}$

When α is listed under X in the lexicon, and $[\alpha]_X$ is a SD with denotation $\| [\alpha]_X \|_{\sigma,w}$, then

$\langle \| [\alpha]_X \|_{\sigma,w} \rangle$ is an interpretation of $[\alpha]_X$.

<u>Phrase structure rules.</u>

1. If α is a SD on the form $[\text{NP VP}]_S$ where NP is not a WH-term, and

$\langle \| \text{NP} \|_{\sigma,w}, (\text{NP}_1)_{i_1}, \ldots, (\text{NP}_n)_{i_n} \rangle$

$\langle \| \text{VP} \|_{\sigma,w}, (\text{NP}_{n+1})_{j_1}, \ldots, (\text{NP}_{n+m})_{j_m} \rangle \quad 0 \leq n,m$

are interpretations of NP and VP respectively, where

i) $i_1, \ldots, i_n, j_1, \ldots, j_m$ are even numbers

ii) $\{i_1, \ldots, i_n\} \cap \{j_1, \ldots, j_m\} = \emptyset$

iii) If $\| \text{NP} \|_{\sigma,w}$ depends on $\sigma_{k_1}, \ldots, \sigma_{k_\ell}$, then $\{k_1, \ldots, k_\ell\} \cap \{j_1, \ldots, j_m\} = \emptyset$

iv) If $k_{a_1}, \ldots, k_{a_h}$ and $k_{b_1}, \ldots, k_{b_\ell}$ are the odd terms in σ that $\| \text{NP} \|_{\sigma,w}, \| \text{VP} \|_{\sigma,w}$ are dependent on respectively, then $\{k_{a_1}, \ldots, k_{a_h}\} \cap \{k_{b_1}, \ldots, k_{b_\ell}\} = \emptyset$

then

$$<\| [NP\ VP]_S \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n}, (NP_{n+1})_{j_1}, \ldots, (NP_{n+m})_{j_m}>$$

is an interpretation of α where

$$\| [NP\ VP]_S \|_{\sigma,w} = 1 \text{ iff } \| VP \|_{\sigma,w} \in \| NP \|_{\sigma,w}$$

2. If α is a SD on the form $[V\ A]_{VP}$ where A is a SD of category NP, $\bar{S}$ or R, A not a WH-term, and

$$<\| V \|_{\sigma,w}>$$
$$<\| A \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n}> \quad 0 \leq n$$

are interpretations of V and A respectively, where

i) $i_1, \ldots, i_n$ are even numbers

then let

$$\| A \|'_{\sigma,w} = \begin{cases} \| A \|_{\sigma,w} & \text{if A is of category NP or } \bar{S} \\ \{p \mid p(w)=1 \wedge \exists a \forall w' \in W (p(w')=1 \leftrightarrow a \in \| A \|_{\sigma,w})\} & \text{otherwise} \end{cases}$$

then

$$<\| [V\ A]_{VP} \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n}>$$

is an interpretation of α where $\| [V\ A]_{VP} \|_{\sigma,w} \subseteq E$ such that

$$b \in \| V\ A]_{VP} \|_{\sigma,w} \text{ iff } <b, \| A \|'_{\sigma}> \in \| V \|_{\sigma,w}$$

3. If α is a SD on the form $[\underline{that}\ S]_{\bar{S}}$ and

$$<\| S \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n}) \quad 0 \leq n$$

is an interpretation of S where

i) $i_1, \ldots, i_n$ are even numbers

then

$$<\| [\underline{that}\ S]_{S} \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n}>$$

is an interpretation of α where $\| [\underline{that}\ S]_{\bar{S}} \|_{\sigma,w} = \| S \|_{\sigma,w}$

4. a) If α is a SD on the form $[\text{NP R}]_{NP}$, NP is not a WH-term, R is on the form $[[\underline{\text{who}}]_{NP}S]_R$ or $[[\underline{\text{which}}]_{NP}S]_R$, and

$$\langle \| NP \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} \rangle$$

$$\langle \| R \|_{\sigma,w}, (NP_{n+1})_{j_1}, \ldots, (NP_{n+m})_{j_n} \rangle \qquad 0 \leq n, m$$

are interpretations of NP and R respectively, where

i) $i_1, \ldots, i_n$, $j_1, \ldots, j_m$ are even numbers

ii) $\{i_1, \ldots, i_n\} \cap \{j_1, \ldots, j_m\} = \emptyset$

iii) If $\| NP \|_{\sigma,w}$ depends on $\sigma_{k_1}, \ldots, \sigma_{k_\ell}$, then

$$\{k_1, \ldots, k_\ell\} \cap \{j_1, \ldots, j_m\} = \emptyset$$

iv) If $k_{a_1}, \ldots, k_{a_h}$ and $k_{b_1}, \ldots, k_{b_\ell}$ are the odd terms in σ that $\| NP \|_{\sigma,w}$, $\| R \|_{\sigma,w}$ are dependent or respectively, then $\{k_{a_1}, \ldots, k_{a_h}\} \cap \{k_{b_1}, \ldots, k_{b_\ell}\} = \emptyset$

then

$$\langle \| (\text{NP R}]_{NP} \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n}, (NP_{n+1})_{j_1}, \ldots, (NP_{n+m})_{j_m} \rangle$$

is an interpretation of α where

$$\| [\text{NP R}]_{NP} \|_{\sigma,w} = \{X \mid X \cap \| R \|_{\sigma,w} \in \| NP \|_{\sigma,w}\}$$

b) If α is a SD on the form $[\text{Det N}]_{NP}$, and

$$\langle \| Det \|_{\sigma,w} \rangle$$

$$\langle \| N \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} \rangle \qquad 0 \leq n$$

are interpretations of Det and N respectively, where

i) $i_1, \ldots, i_n$ are even numbers

then

$< \| [\text{Det N}]_{NP} \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} >$

is an interpretation of α where

$$\| [\text{Det N}]_{NP} \|_{\sigma,w} = \| \text{Det} \|_{\sigma,w}(\| N \|_{\sigma,w})$$

c) If α is a SD on the form $[R]_{NP}$, R is on the form $[[\underline{\text{what}}]_{NP} S]_R$, and

$< \| R \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} >$

is an interpretation of R where

i) $i_1, \ldots, i_n$ are even numbers

then

$< \| [R]_{NP} \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} >$

is an interpretation of α where

$$\| [R]_{NP} \|_{\sigma,w} = \{X \subseteq E \mid \| R \|_{\sigma,w} \subseteq X\}$$

5. If α is a SD on the form $[N\ A]_N$, A is a SD of category R or $\bar{S}$, if A is of type R then A is on the form $[[\underline{\text{who}}]_{NP} S]_R$ or $[[\underline{\text{which}}]_{NP} S]_R$, and

$< \| N \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} >$

$< \| A \|_{\sigma,w}, (NP_{n+1})_{j_1}, \ldots, (NP_{n+m})_{j_m} >$ $\quad 0 < n, m$

are interpretations of N and A respectively, where

i) $i_1, \ldots, i_n, j_1, \ldots, j_n$ are even numbers

ii) $\{i_1, \ldots, i_n\} \cap \{j_1, \ldots, j_n\} = \emptyset$

iii) If $\| N \|_{\sigma,w}$ depends on $\sigma_{k_1}, \ldots, \sigma_{k_\ell}$ then

$\{k_1, \ldots, k_\ell\} \cap \{j_1, \ldots, j_m\} = \emptyset$

iv) If $k_{a_1},\dots,k_{a_h}$ and $k_{b_1},\dots,k_{b_\ell}$ are the odd terms in σ that $\|N\|_{\sigma,w}, \|R\|_{\sigma,w}$ are dependent on respectively then $\{k_{a_1},\dots,k_{a_h}\}\cap\{k_{b_1},\dots,k_{b_\ell}\} = \emptyset$

v) If A is of type $\bar{S}$, then $\|A\|_{\sigma,w}$ depends on σ_j, for an odd number j

let $\|A\|'_{\sigma,w} = \begin{cases} \|A\|_{\sigma,w} & \text{if A is of type R} \\ \{a \mid \|A\|_{\sigma^j_a,w} = 1\} & \text{if A is of type } \bar{S} \end{cases}$

then

$$\langle \|[N\ A]_N\|_{\sigma,w}, (NP_1)_{i_1},\dots,(NP_n)_{i_n}, (NP_{n+1})_{j_1},\dots,(NP_{n+m})_{j_m}\rangle$$

is an interpretation of α where

$$\|[N\ A]_N\|_{\sigma,w} = \|N\|_{\sigma,w} \cap \|A\|'_{\sigma,w}$$

By σ^j_a we mean the sequence exactly like σ except possibly for the j-th term where σ^j_a has the value a.

6. If α is a SD on the form $[NP\ S]_R$, NP is a WH-term, and

$\langle \emptyset, (NP_0)_j\rangle$

$\langle \|S\|_{\sigma,w}, (NP_1)_{i_j},\dots,(NP_n)_{i_n}\rangle \quad 0 \leqslant n$

are interpretations of NP and S respectively where

i) j is an odd number, $i_1,\dots,i_n$ are even numbers

ii) $\|S\|_{\sigma,w}$ depends on σ_j

then

$$\langle (NP_0)_j(\|S\|_{\sigma,w}), (NP_1)_{i_1},\dots,(NP_n)_{i_n}\rangle$$

is an interpretation of α. (See the definition of WH-storage for the definition of $(NP_0)_j$.)

Note that the fragment does not include a special category for questions. If we wanted to single out the questions in the syntax, we could have chosen the approach used for headless relatives. That is, we could have added a syntactical rule $Q \to R$ and replaced R with Q in syntax rule 2. In the semantics we would then need a rule expressing

$$\|Q\|_{\sigma,w} = \{p \mid p(w) = 1 \wedge \exists a \forall w' \in W(p(w') = 1 \leftrightarrow a \in \|A\|_{\sigma,w})\},$$

eliminating the exception for embedded questions in rule 2. Such a fragment would be equivalent to the present one. Similarly, we could have chosen the approach used in the present fragment for embedded questions also in the case of headless relatives.

Rules for the storage mechanisms

For $i \in \mathbb{N}$ we define: $\|[i]_{NP}\|_{\sigma,w} = \{X \subseteq E \mid \sigma_i \in X\}$

Pronoun opening:

If α is $[he]_{NP}$ then for any even i, $i \in \mathbb{N}$

$\langle \|[i]_{NP}\|_{\sigma,w} \rangle$ is an interpretation of α

Gap opening:

If α is $[\,]_{NP}$ then for any odd i, $i \in \mathbb{N}$

$\langle \|[i]_{NP}\|_{\sigma,w} \rangle$ is an interpretation of α

NP-storage:

If α is a SD of category NP not a WH-term and

$\langle \|NP\|_{\sigma,w} \rangle$ is an interpretation of α

where $\|NP\|_{\sigma,w}$ does not depend on any term in σ, then for any even i, $i\in\mathbb{N}$

$$\langle \|[i]_{NP}\|_{\sigma,w}, (NP)_i\rangle \text{ is an interpretation of } \alpha$$

where

$(NP)_i$ is the function $(NP)_i : (\{0,1\}^W)^{E^\omega} \longrightarrow (\{0,1\}^W)^{E^\omega}$

such that if $\phi\in(\{0,1\}^W)^{E^\omega}$ then

$(NP_i(\phi)(\sigma)(w) = 1$ iff $\{a|\phi_{\sigma}{}^{i}_{a,w} = 1\}\in\|NP\|_{\sigma,w}$

<u>WH-storage</u>:

If α is a WH-term and

$$\langle \|NP\|_{\sigma,w}\rangle \text{ is an interpretation of } \alpha$$

where $\|NP\|_{\sigma,w}$ does not depend on any term in σ, then for any odd i, $i\in\mathbb{N}$

$$\langle \emptyset, (NP)_i\rangle \text{ is an interpretation of } \alpha$$

where

$(NP)_i$ is the function $(NP)_i : (\{0,1\}^W)^{E^\omega} \longrightarrow ((\{0,1\}^E)^W)^{E^\omega})$

such that if $\phi\in(\{0,1\}^W)^{E^\omega}$ then

$$(NP)_i(\phi)(\sigma)(w) = \begin{cases} \{a|\phi_{\sigma}{}^{i}_{a,w} = 1 \wedge \exists A(A\in\|NP\|_{\sigma,w} \wedge a\in A\} & \text{for alternative a)} \\ \{a|\phi_{\sigma}{}^{i}_{a,w} = 1 \wedge a\in\|NP\|_{\sigma,w}\} & \text{for alternative b)} \end{cases}$$

<u>NP-retrieval</u>:

If α is a SD of category S and

$$<\| S \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n}> \quad 1 \leq n$$

is an interpretation of α where $i_1, \ldots, i_n$ are even numbers and $1 \leq k \leq n$, then

$$<(NP_k)_{i_k}(\| S \|_{\sigma,w}), (NP_1)_{i_1}, \ldots, (NP_{k-1})_{i_{k-1}}$$
$$(NP_{k+1})_{i_{k+1}}, \ldots, (NP_n)_{i_n}>$$

is also an interpretation of α.

1.4. <u>Nested storage</u>.

In this paragraph, one of the most important restrictions on storage, the restriction against storage of two NP-terms, the one contained in the other, is removed to see whether the storage technique can accomplish the interpretations then required.

1.4.1. <u>What is nested storage</u>?

<u>Definition 8</u>: Let α be a SD of category NP where

$$<\| \alpha \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n}> \quad 0 \leq n$$

is an interpretation of α such that $\| \alpha \|_{\sigma,w}$ depends on σ_j, $j \in \mathbb{N}$. To apply storage on this interpretation is called <u>nested storage</u>.

In other words, nested storage is to store a NP-intension containing an opening. This is prevented in the earlier fragments. Note that an interpretation without openings must have empty store or empty denotation part. Whether the opening results from previous storage or pronoun/gap opening will not be distinguished. From a retrieval point of view, openings are recognized only by their correspondence to a term in σ, and will be treated in the same way no matter whether they result from the creation of a binding operator or from a pronoun/gap opening.

To have an interpretation satisfying the conditions of definition 8, a NP in Fragment 6^{II} must be on one of the following forms:

$[NP\ R]_{NP}$ $\quad [NP[NP[NP\ VP]_S]_R]_{NP}$ non-restrictive relative

$[Det[N[\begin{Bmatrix} NP \\ that \end{Bmatrix}[NP\ VP]_S]\begin{Bmatrix} R \\ \bar{S} \end{Bmatrix}]_N]_{NP}$ restrictive relative

$[NP[NP\ VP]_S]_R]_{NP}$ headless relative

In other fragments of English, simpler NP-constructions containing NP-terms would be generated, i.e. by a syntax rule like NP → NP' <u>and</u> NP". The present fragment is not particularly fit for the study of nested storage as the examples all will be unnecessarily complicated.

Nested storage is like simple storage motivated by wide scope readings of one of the following types:

1. Reversed quantifier scope readings
2. De re readings
3. Binding of pronouns.

This gives 6 types of nested storage. For each of the three kinds mentioned above the interpretation of the NP-term in question may contain

i) an even opening (with or without stored binding operator)

ii) a gap opening

In the present fragment, case ii) will only occur if the NP-term is a part of a larger relative clause (containing two gaps). This does not occur in English, but is grammatically acceptable in Scandinavian languages (see Engdahl 1980). The storage mechanisms should consequently give an appropriate account even for such constructions.

1.4.2. Fragment 6^{III}.

The fragment is presented as a series of alterations to Fragment 6^{II}.

Openings may now occur in the store, and consequently the safety mechanisms to prevent backwards binding and more than one gap opened on one term in σ must be changed accordingly. The phrase structure interpretation rules must all be modified corresponding to the following rewriting of the safety mechanisms iii) and iv) on page 7.

iii)' If $\|A\|_{\sigma,w},(NP_1)_{i_1},\ldots,(NP_n)_{i_n}$ depends on $\sigma_{k_1},\ldots,\sigma_{k_\ell}$ then $\{k_1,\ldots,k_\ell\}\cap\{j_1,\ldots,j_m\} = \emptyset$

iv)' If $k_{a_1},\ldots,k_{a_h}$ and $k_{b_1},\ldots,k_{b_\ell}$ are the odd terms in σ that $\|A\|_{\sigma,w},(NP_1)_{i_1},\ldots,(NP_n)_{i_n}$ and $\|B\|_{\sigma,w},(NP_{n+1})_{j_1},\ldots,(NP_{n+m})_{j_m}$ depends on respectively, then $\{k_{a_1},\ldots,k_{a_h}\}\cap\{k_{b_1},\ldots,k_{b_\ell}\} = \emptyset$.

To allow nested storage, the storage rules and the rule for NP-retrieval must be replaced by the following rules:

<u>NP-storage</u>:

If α is a SD of category NP not a WH-term,

$$\langle \| NP_0 \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} \rangle \quad 0 \leq n$$

is an interpretation of α, there is no $i \in \mathbb{N}$ such that $\| NP_0 \|_{\sigma,w} = \| [i]_{NP} \|_{\sigma,w}$ for all $\sigma \in E^{\omega}$, $w \in W$ and $i_1, \ldots, i_n$ are even numbers, then for any even $i \in \mathbb{N}$ such that $i \neq i_1, \ldots, i_n$ and $\| NP_0 \|_{\sigma,w}$ does not depend on σ_i,

$$\langle \| [i]_{NP} \|_{\sigma,w}, (NP_0)_i, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} \rangle$$

is an interpretation of α

where

$(NP_0)_i$ is the function $(NP_0)_i : (\{0,1\}^W)^{E^{\omega}} \to (\{0,1\}^W)^{E^{\omega}}$ such that if $\phi \in (\{0,1\}^W)^{E^{\omega}}$ then

$$(NP_0)_i(\phi)(\sigma)(w) = 1 \text{ iff } \{a \mid \phi_{\sigma_a^i, w} = 1\} \in \| NP_0 \|_{\sigma,w}$$

<u>WH-storage</u>:

If α is a WH-term and

$$\langle \| NP_0 \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} \rangle \quad 0 \leq n$$

is an interpretation of α where $i_1, \ldots, i_n$ are even numbers, then for any odd $i \in N$ such that $\| NP_0 \|_{\sigma,w}$ does not depend on σ_i

$$\langle \emptyset, (NP_0)_i, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} \rangle$$

is an interpretation of α

where

$(NP_0)_i$ is the function $(NP_0)_i:(\{0,1\}^W)^{E^\omega} \rightarrow ((\{0,1\}^E)^W)^{E^\omega}$

such that if $\phi \in (\{0,1\}^W)^{E^\omega}$ then

$$(NP_0)_i(\phi)(\sigma)(w) = \begin{cases} \{a \mid \phi_{\sigma^i_a,w} = 1 \wedge \exists A(A \in \| NP_0 \|_{\sigma,w} \wedge a \in A\} \\ \qquad \text{for alternative a)} \\ \{a \mid \phi_{\sigma^i_a,w} = 1 \wedge a \in \| NP_0 \|_{\sigma,w}\} \\ \qquad \text{for alternative b)} \end{cases}$$

NP-retrieval:

If α is a SD of category S and

$$< \| S \|_{\sigma,w}, (NP_1)_{i_1}, \ldots, (NP_n)_{i_n} \qquad 1 \leqslant n$$

is an interpretation of α where $i_1, \ldots, i_n$ are even numbers, and $k \in \mathbb{N}$ is a number such that $1 \leqslant k \leqslant n$ and neither of $(NP_1)_{i_1}, \ldots, (NP_n)_{i_n}$ depends on σ_{i_k}, then

$$<(NP_k)_{i_k}(\| S \|_{\sigma,w}), (NP_1)_{i_1}, \ldots, (NP_{k-1})_{i_{k-1}}, \\ (NP_{k+1})_{i_{k+1}}, \ldots, (NP_n)_{i_n}>$$

is also an interpretation of α.

The modifications are all only in the semantic machinery, and would apply equally well to a fragment with the syntax of Fragment 6 in Cooper (1983) with semantics modified similarly to the semantics of Fragment 6^{II}.

1.4.3. Examples.

The examples are phrases generated by the syntax of Fragment 6^{II} and, hence, of Fragment 6 in Cooper (1983) since the two fragments have the same generative power. The interpretations that require nested storage are not obtained in these fragments. To get examples that are more natural, we also use words that are not explicitely listed in the lexicon, and write the sentences morphologically correct. The binding relations are indicated with arrows. The example with two gaps in one relative clause is for obvious reasons in Norwegian. The interpretation processes are loosely described, exact calculations can be found in Sem (1983, I. 5.3)[6].

(1) Mary believes that a man who ___ loves her admires a woman who ___ hates him.

To get the NP a man who loves her to corefer with him at the same time as Mary corefers with her in a man who loves her, a man who loves her must be stored with her open, and retrieved to have scope over the embedded sentence or over the whole sentence before Mary is retrieved to bind her. The intended interpretation is obtainable in fragment 6^{III}, and by choosing a man who loves her to have scope over the embedded sentence, we get[7]:

$$<<\|Mary\|_{\sigma,w}, \|\bar{S}\|_{\sigma^{2}_{\|Mary\|_{\sigma,w}},w}> \in \|believe\|_{\sigma^{2}_{\|Mary\|_{\sigma,w}},w}>$$

where

$\|\bar{S}\|_{\sigma^{2}_{\|Mary\|_{\sigma,w}},w}$ is

$$\exists b \exists y(<b,\|Mary\|_{\sigma,w}> \in \|loves\|_{\sigma^{2\,3}_{\|Mary\|\,c},w} \wedge b \in \|man\|_{\sigma^{2}_{\|Mary\|},w}$$

$$\wedge <b,y> \in \|admires\|_{\sigma^{2\,4}_{\|Mary\|_{\sigma,w}\,b},w}$$

$$\wedge <y,b> \in \|hates\|_{\sigma^{2\,4\,5}_{\|Mary\|_{\sigma,w}\,b\,a},w} \wedge y \in \|woman\|_{\sigma^{2\,4}_{\|Mary\|_{\sigma,w}\,b},w}>$$

(2) Every man who ___ loves Mary

says that he needs her

In this example every man who loves Mary is to corefer with he at the same time as Mary is to corefer with her. To obtain this, both Mary and every man who loves Mary must be stored or, to be exact, Mary must be stored first, and then every man who loves i, with a binding operator on i for Mary in the store, must be stored. Both the stored NP's must have scope over the full sentence, and for technical reasons every man who loves i must be retrieved before Mary in order to have i bound by Mary. Just before retrieval, the situation can informaly be described like this:

$$\langle \| i \text{ says that } i \text{ needs } j \|_{\sigma,w}, (\text{Every man who loves } j)_i, (\text{Mary})_j \rangle$$

Since the first binding operator depends on the term bound by the second, the second binding operator cannot be retrieved before the first. Mary therefore gets scope over Every man who loves j. In this case however, this has no serious side effects. Formally, the obtained interpretation will be

$$\langle \forall a((\langle a, \|\text{Mary}\|_{\sigma,w}\rangle \in \|\text{loves}\|_{\sigma,w} \wedge a \in \|\text{man}\|_{\sigma,w}) \longrightarrow$$
$$\langle a, \|\bar{S}\|_{\sigma^{4}_{a}{}^{2}_{\|\text{Mary}\|_{\sigma,w}}} \rangle \in \|\text{says}\|_{\sigma,w})\rangle$$

where $\forall w \in W$,

$$\|\bar{S}\|_{\sigma^{4}_{a}{}^{2}_{\|\text{Mary}\|_{\sigma,w}}}(w) = (\langle a, \|\text{Mary}\|_{\sigma}\rangle \in \|\text{needs}\|_{\sigma^{4}_{a}{}^{2}_{\|\text{Mary}\|_{\sigma,w}}})$$

that can be paraphrased as

$$\forall a[(a \text{ loves Mary} \wedge a \text{ is a man}) \longrightarrow (a \text{ says that } a \text{ needs Mary})].$$

(3) Every man who ___ loves a woman

says that he needs her

This example is just like (2), exept that Mary is replaced by a woman. In this case, the fact that a woman must have scope over every man who loves i does have serious side effects. As in (2), we obtain the intended coreferentiality, but there is no way to obtain the main reading of this sentence, namely the reading with the indicated coreferentiality and with Every man who loves i having scope over a woman.

(4) 8) Enhver hest (som) Mari kjenner
Every horse that Mary knows

en mann som ___ eier ___, kaster ham av
a man that owns throws him off

To have en mann som ___ eier ___ bind ham, en mann som ___ eier ___ must have scope over the full sentence. This means we have to store en mann som ___ eier j with an open gap until the full sentence is constructed. Technically, this will not work, since WH-retrieval is obligatory at relative clause formation and retrieval cannot take place while a corresponding opening is in the store. To get an interpretation at all, en mann som ___ eier j must be retrieved before relative clause formation if stored at all, and we thereby lose the possibility of coreference with anything outside the relative clause. The interpretation we do obtain, may be paraphrased as

$\forall b \in \| hester \|_{\sigma,w} (\exists a \in \| mann \|_{\sigma,w}$ (Mari kjenner a $\wedge$ a eier b)
$\longrightarrow$ (b kaster σ_2 av)),

a solution varying with σ_2. This is not the intended interpretation since σ_2 is not bound to a.

1.5. Storage limitations.

We have that

Quantifier scope depends on the order of retrieval. The earlier the retrieval, the narrower the scope.

De re/de dicto De re readings are obtained by storage

Binding Coreferentiality is obtained by interpreting different openings in an S denotation by the same individual. This is implemented by storing and retrieval.

In fragement 6^{III} we also have that

At nested storage, an outer NP-term must be retrieved before inner NP-terms.

Obviously this may lead to difficulties, i.e. if we want an outer NP-term to have wider scope than an inner. To retrieve an outer NP-term before an inner is in any case a precarious principle. The least natural quantifier order is enforced, that is the order opposite to the order given from left to right in the phrase. The question is whether another strategy for reformulation of the storage and retrieval mechanisms would give a better result, or whether this is a problem implicit to the storage technique itself.

To obtain coreferentiality between openings (at the same term in σ), it is necessary to retrieve on a semantic unit, either on the denotation part or on one of the binding operators. Retrieval on the entire interpretation or parts of it not forming a semantic unit will not give coreferentiality. In addition to the retrieval definition given in Fragment 6^{III}, we have the following possibilities:

1) To retrieve on the denotation part of an interpretation of a SD of a category different from S.

2) To retrieve on a binding operator.

Alternative 1 gives the possibility of NP-scopes over phrases of other categories than S. This could eliminate some of the cases of conflict between de re/de dicto readings and narrow/wide scope, but cannot remove the problem entirely. A better solution with respect to quantifier order would not be obtained. For alternative 2), we could define the composition of two binding operators like this:

<u>Definition 9</u>: If $(NP)_i$ and $(NP')_j$ are two binding operators such that i,j are even numbers and $(NP')_j$ depends on σ_i, then $(NP)_i((NP')_j)$ is the function

$$(NP_i((NP'_j)):(\{0,1\}^W)^{E^\omega} \rightarrow (\{0,1\}^W)^{E^\omega}$$ such that if $\phi \in (\{0,1\}^W)^{E^\omega}$

then

$$(NP)_i((NP')_j)(\phi)(\sigma)(w) = 1 \text{ iff } \{a | \{b | \phi_{\sigma^{j\ \ i}_{b\ \ a},w} = 1\} \in \| NP' \|_{\sigma^i_{a,w}}\} \in \| NP \|_{\sigma,w}$$

However, this is even less satisfactory, forcing both NP-terms to have scope over the same phrase while maintaining to give the reversed quantifier order.

Our conclusion is therefore that extension of the storage mechanisms to NP-terms containing NP-terms is possible, but forces the reversed quantifier order at nested storage. This does not matter as long as we keep to binding of pronouns. Fragment 6^{III} works satisfactorily for (1) and (2). The problem appears in (3) and (4) where binding and scope relations interfere. The problem is essentially the same in these examples (though it manifests itself slightly differently): conflict between binding and scope. In example (3) the binding enforces the reversed quantifier order, and the left-to-right order would have blocked the intended binding relation. In (4), the intended binding relation is blocked by the obligatory scope order. The WH-word can be regarded as a scope marker for the WH-binding, and WH-retrieval is obligatory at formation of the relative clause. A stored opening must be retrieved before it can be bound. Thus if the NP A contains a NP B but not the NP C, that is, we have a SD on the form

$$\ldots[\ldots[B]_{NP}\ldots]_{NP}\ldots[C]_{NP}\ldots$$

$$(\text{or } \ldots[C]_{NP}\ldots[\ldots[B]_{NP}\ldots]_{NP}\ldots])$$

B cannot corefer with or have wider scope than C without also having wider scope than A.

The problem rises from the combination of manipulation of quantifier scope (or de re/de dicto reading) and binding through the same mechanism, the one enforcing a retrieval order conflicting with the retrieval order required by the other. This is not a consequence of the nested storage. The same problem occurs in the sentence

(5) Every man who ___ loves a woman needs her,

a variety of the so-called "donkey sentence" (after Geach, 1962). In the following sentence we have an example of conflict between quantifier scope and de re reading:

(6) The detective thinks a psychopath killed three freemasons.

on the reading that the detective is not yet aware of the fact that the three victims were freemasons, so that a psychopath is to be read de dicto while three freemasons is to be read de re without having wider scope han a psychopath. An example of conflict between binding relation and de dicto reading is

(7) John seeks a unicorn because he wants to ride on it

where a unicorn is to be read de dicto. To use the same mechanisms to obtain coreferentiality and de re/de dicto readings, and to manipulate quantifier scopes invariably leads to this kind of problems.

2. The Discourse Representation Systems

2.1. Introduction.

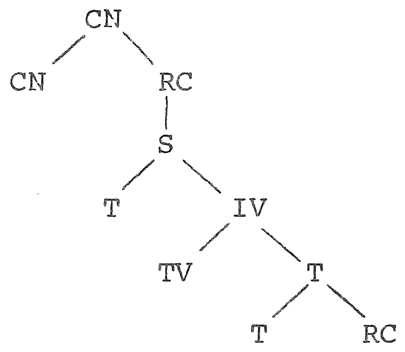
In this section we will study the approach to quantifier scope and anaphoric relation between noun phrases developed by Hans Kamp in "A Theory of Truth and Semantic Representation" Kamp (1981) as opposed to the Cooper approach as presented in the first section. The Discourse Representation Systems (DRS) of Kamp (1981) are designed to give an account of meaning both as "that which determines conditions of truth" and as "that which a language user grasps when he understands the words he hears or reads" (Kamp 1981, sect. 1). Kamp has two central concerns in his choice of fragment in (Kamp 1981) namely "(a) to study the anaphoric behaviour of personal pronouns; and (b) to formulate a plausible account of the truth conditions of the so-called 'donkey— sentences'" (Kamp 1981, sect. 1).

Cooper and Kamp both keep the syntax and semantics of their systems strictly apart and avoid using semantic motivation for the syntax definitions. This allows us in the following to interpret a slightly modified version of Fragment 6^{II} by a DRS-semantics. The semantics presented below will be a modified and extended version of the system presented in Kamp (1981). The modifications are mainly cosmetic and the main features of the system are preserved.

2.2. L_1 - Fragment 6^{II}-syntax and DRS-semantics.

We will follow Kamp in allowing relative clauses to combine only with basic terms or common nouns and not with derived phrases of category T or CN[9]. That is, the following constructions corresponding to a T or CN already specified or restricted by a relative clause being further specified or restricted, will be prevented:

On the other hand the following construction for terms containing more than one relative clause will be permitted:



That is, the fragment will generate relative clauses on the basis of terms with relative clause constituents. In paragraph 2.4 we will discuss some approaches allowing more than one non-restrictive relative to the same CN. Nevertheless, the fragment below will be sufficiently rich to illustrate the quantifier scope problems and anaphoric relations which we have focused on above (1.4.3, 1.5).

For the present, we will ignore the problem of intensionality and keep to a purely extensional fragment. The DRS-system is not fit for the intensionality treatement of possible world semantics. I believe that with the choice of another base structure, for instance the situation semantics of Barwise and Perry (1983), a fruitful account also for the phenomena we usually cathegorize as intensional can be given.

From these simplifications, it follows automatically that TV-expressions taking $\bar{S}$ objects and intensional TV's like <u>seek</u> will be omitted from L_1.

To avoid the overgenerating capacity resulting from the non-transformational construction of relative clauses on the basis of WH-words and gaps as words of category NP (Det), the relative clause construction will be as in L_0 in Kamp (1981). As far as only acceptable phrases are concerned, the DRS-construction is not affected by this choice. The determiners <u>a</u> and <u>every</u> are syncategorematically introduced, and the fragment contains conditionals. L_1 is to be an extension of L_0; the lexicon and names for the syntax categories are as in L_0. To emphasize the analogy between Fragment L_1 and 6^{II}, the syntactic rules are given in the same order as in Fragment 6^{II}.

2.3. <u>Fragment L_1</u>.

2.3.1. <u>Syntax</u>.

L_1 contains expressions of the following categories with the following basic members:

1) T: <u>Pedro</u>, <u>Chiquita</u>, ..., <u>he</u>, <u>she</u>, <u>it</u>
2) CN: <u>farmer</u>, <u>donkey</u>, <u>window</u>, <u>man</u>, <u>woman</u>, ...
3) IV: <u>thrives</u>, <u>runs</u>, ...
4) TV: <u>owns</u>, <u>beats</u>, <u>loves</u>, ...
5) S
6) RC
7 $\bar{S}$

The terms <u>he</u>, <u>she</u> and <u>it</u> are called <u>pronouns</u>.

<u>Formation rules</u>

FR1): If $\alpha \in IV$, $\beta \in T$ then $\beta\alpha \in S$

FR2): If $\alpha \in TV$, $\beta \in T$ then $\alpha\beta' \in IV$ where

$$\beta' = \begin{cases} \underline{\text{him}}\ \delta \text{ if } \beta = \underline{\text{he}}\ \delta \\ \underline{\text{her}}\ \delta \text{ if } \beta = \underline{\text{she}}\ \delta \\ \beta \text{ otherwise} \end{cases}$$

where δ is possibly the empty string.

FR3,k: If $\phi \in S$ and the k-th word in ϕ is a pronoun, then <u>that</u> $\phi' \in \bar{S}$ where ϕ' is the result of eliminating the k-th word from ϕ.

FR4): a) If $\alpha \in RC$ and β is a basic T or formed by FR4 b) then $\beta\alpha \in T$.

b) If $\alpha \in CN$ then $\left.\begin{array}{l} \text{i) } \underline{\text{a(n)}}\alpha \\ \text{ii) } \underline{\text{every}}\ \alpha \end{array}\right\} \in T$

c) If $\alpha \in RC$ and α starts with <u>what</u> or <u>whoever</u> then $\alpha \in T$

FR5): If α is a basic CN,

$$\beta \in \begin{cases} RC \text{ and } \beta \text{ starts with } \underline{\text{who}}/\underline{\text{whom}}/\underline{\text{which}} \\ \bar{S} \end{cases}$$

then $\alpha\beta \in CN$

FR6,k): If $\psi \in S$ and the k-th word of ψ is a pronoun, then $\beta\psi' \in RC$, where ψ' comes from ψ as in FR3k), and

β is { <u>who</u>/<u>whom</u>/<u>which</u> according as the pronoun is <u>he</u>, <u>she</u>/<u>him</u>, <u>her</u>/<u>it</u> respectively or <u>what</u> if the prounon is <u>it</u>, <u>whoever</u> otherwise. }

FR7): If $\phi, \psi \in S$ then <u>if</u> ϕ, ψ and <u>if</u> ϕ <u>then</u> ψ } $\in S$

Note: FR3,k) would not be sufficient for the construction of <u>that</u> S-phrases in a fragment containing <u>that</u> S-phrases other than relative clauses. L_1 is a purely extensional fragment without sentence embedding verbs. The problem with the two kinds of <u>that</u> S-phrases is thus avoided.

<u>Syntax terminology</u>

<u>Definition 10</u>: <u>A L_1-discourse</u> is an ordered n-tupple of expressions of category S in L_1

Let $D = \langle \Phi_1, \ldots, \Phi_n \rangle$ be a L_1-discourse, and $\langle \tau_1, \ldots, \tau_n \rangle$ a synatic analyses for D. The nodes in $\tau_1, \ldots, \tau_n$ can be uniquely numbered, e.g. by the leftmost branch principle.

<u>Definition 11</u>: <u>An occurrence</u> of an expression α in D (relative to a syntactic analysis) is an ordered couple $\langle \alpha, n \rangle$ where n is the index of the connection node for α in the syntactic analysis of D.

Let γ be a L_1-expression

<u>Definition 12</u>: If the rule applied last in the construction of γ is FR1 or FR2, then γ is formed by combining a term α with a IV or TV β.

Then

- α is the <u>main term</u> in γ
- α' is the <u>term with maximal scope</u> in γ,

  where α' =
  - If the rule applied last in the construction of α is FR4 a) then α is formed by combining a term with a RC ρ. Let α' be this term
  - α otherwise.

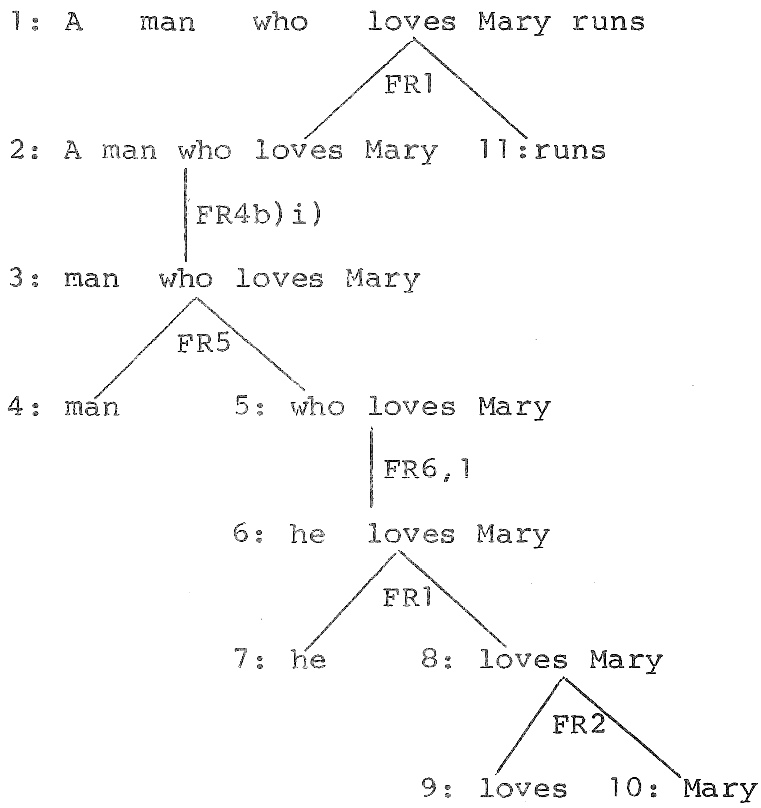
The term with maximal scope is the main term without non-restrictive relatives. The reason to single out non-restrictive relatives this way is the special position non-restrictive relatives hold in the semantics. They are intended to give supplementary information about the object(s) determined by the rest of the term, and do not give substantial contribution to the determination of the object(s).

We will follow Cooper (1983 Ch IV 2.3) in interpreting headless relatives as if they contained a hidden universal:

<u>Definition 13</u>: If the rule applied last in the construction of γ is FR1 or FR2 and the term with maximal scope starts with <u>every</u>, <u>whoever</u> or <u>what</u>, then γ is called a <u>universal sentence</u> or a universal <u>IV</u> respectively. If the rule applied last in the construction of γ is FR7, γ is called a <u>conditional</u>. The term with maximal scope in a universal sentence or IV is called a <u>universal term</u>

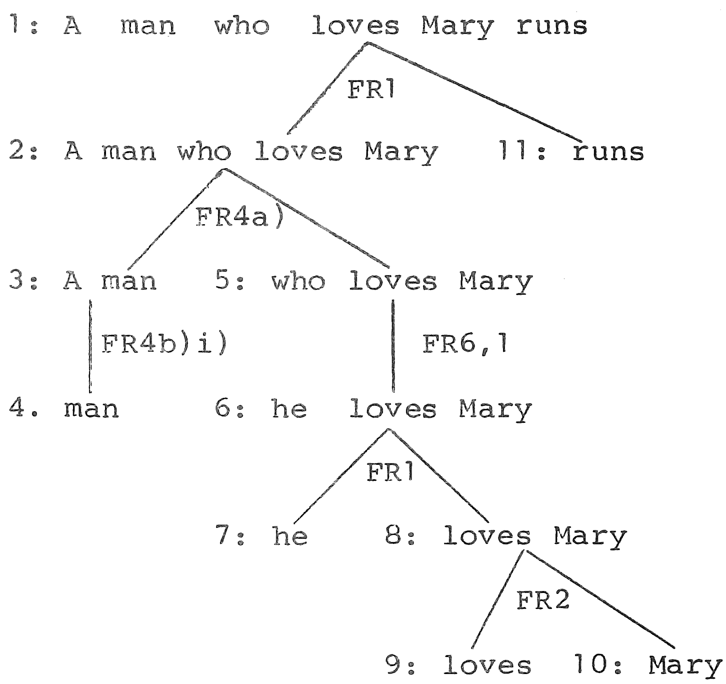
(8) γ = A man who loves Mary runs

Syntatic analysis 1:



Here, a man who loves Mary is the term with maximal scope in γ as well as the main term.

Syntactic analysis 2:



With this analysis, a man who loves Mary is still the main term while the term with maximal scope is simply a man.

2.3.2 Semantics

The semantics consists of a model, a system for how to describe submodels of this model from sentences in L_1, and a definition of truth in the model for such descriptions.

We use L_1 together with some auxiliary symbols and some variables called discourse referents to describe submodels. Given a L_1-discourse, we construct a "discourse representation system", an ordered system of "discourse representations" describing simple submodels on a small universe of discourse referents. The discourse will be true in the model (on the reading represented by the discourse representation system) if there is a mapping from the discourse referents involved to the model, satisfying the conditions set by the discourse representation system.

Definition 14. A model for L_1 is a structure on the form $M = \langle U,F \rangle$ where

i) U is a non-empty set

ii) F is an interpretation function mapping each

- proper name in L_1 to an element of U
- basic CN or IV in L_1 to a subset of U
- basic TV in L_1 to a set of ordered pairs of members of U.

The pronouns do not get any interpretation by F. This corresponds to the opening - interpretation of pronouns in the fragements in part 1 (and to the interpretation of gaps, as L_1 uses pronouns instead of gaps in the construction of relative clauses.)

Definition 15: Let V be a countable set of entities none of which is a basic expression in L_1 or a string of such expressions. V is called the set of discourse referents.

Definition 16: For any subset $X \subseteq V$ of V, let $L_1(X)$ be the result of adding the members of X as basic terms (expressions of category T) to L_1. Let $L_1'(X)$ be $L_1(X)$ extended with the symbols =, (,), and the syntactical rules:

i) If $u \in X$, α a proper name, then $u = \alpha \in S$

ii) If $u \in X$, $\alpha \in CN$ then $\alpha(u) \in S$

In $L_1'(X)$ the following is a derived rule:

iii) If $u \in X$, $\delta_1 \alpha \delta_2 \in S$, $\alpha \in T$, then $\delta_1 u \delta_2 \in S$

Definition 17: Let γ be an expression in $L_1'(X)$, $\gamma \in S$. γ may then be of one of the following three kinds:

i) γ is on the form If δ_1, δ_2 or If δ_1 then δ_2, $\delta_1, \delta_2 \in S$ in $L_1(X)$, and γ is called a conditional.

ii) γ is on the form $\alpha\beta$, $\alpha \in T$, $\beta \in IV$ in $L_1(X)$, and γ is called a simple sentence

iii) γ is an expression of $L_1'(X) \setminus L_1(X)$, and is called a qualifying sentence (qualifier)[10].

Definition 18: If γ is a $L_1'(X)$-expression, $\gamma \in S$, and all the well-formed subexpressions of γ (including γ itself) that are also expressions in L_1 are basic expressions of some syntactic category in L_1, then γ is called an atomic sentence in $L_1'(X)$.

The following table gives examples of the kind of sentences we have defined in definition 17 and 18.

Table: Examples of atomary and non-atomic sentences.

	non-atomic	atomic
conditional	If Pedro owns a donkey, he beats it	
qualifying sentence	man who loves Mary(u)	man (u) u=Pedro
simple sentence	Pedro owns a donkey	u runs u loves v

u,v are discourse referents.

If γ is a non-atomic simple sentence in $L_1(X)$, let γ' denote the longest well-formed subexpresson of γ in $L_1(X)$ not on the form $u\delta$ or δu, where δ is of any syntax category, and u is a discourse referent.

Proposition: If γ is a non-atomic simple sentence in $L_1(X)$, γ' is on the form $\alpha\beta$ where $\alpha \in T$, $\beta \in IV$, or γ' is on the form $\beta\alpha$ where $\alpha \in T$, $\beta \in TV$.

Proof: Suppose this is not the case. Since γ is a simple sentence, γ must be on the form $\alpha\beta$, $\alpha \in T$, $\beta \in IV$. γ' is not on this form, so $\gamma' \neq \gamma$, and α must be a discourse referent. γ is non-atomic, so β can not be a basic IV and is thus on the form $\beta'\alpha'$, $\beta' \in TV$, $\alpha' \in T$. Again, γ' is not on this form, so $\gamma' \neq \beta$ and α' must be a discourse referent. But all the expressions of category TV are basic, so in this case γ is atomary, contradicting the premises.

Definition 19: If γ is a non-atomic simple sentence in $L_1(x)$, the main term and the term with maximal scope in γ is then as in γ', and γ is said to be universal if γ' is a universal sentence or IV.

Definition 20: If γ_1 and γ_2 are $L_1'(x)$-expressions, and γ_2 comes from γ_1 exclusively by the use of the rules i)-iii) under definition 16, then γ_2 is a descendant of γ_1.

Definition 21: If D is a L_1-discourse, $\langle\delta,k\rangle$ an occurrence in L_1 of an expression in D, then $\langle\delta',k\rangle$ is an occurrence in $L_1'(x)$ of δ in D if $\delta'=\delta$ or δ' is a descendant of δ. If δ' is a descendant of δ, we say that $\langle\delta',k\rangle$ is a descendant of $\langle\delta,k\rangle$.

Discourse Representation

Notation 2: Indexed or primed V's are used for denoting sets of discourse referents (subsets of V), while subsets of the universe U are denoted by indexed or primed U's.

Definition 22: A possible DR (Discourse Representation) of the L_1-discourse D is an ordered pair $\langle V_m, Con_m\rangle$ where:

i) $V_m \subseteq V$

ii) Con_m consists of occurrences in L_1 of expressions in D or descendants in $L_1'(X)$ of such occurrences.

<u>Definition 23</u>: If m and m' are possible DR's for D, m' <u>extends</u> m if $V_m \subseteq V_{m'}$ and $Con_m \subseteq Con_{m'}$. Moreover, <u>m"=m+m'</u> (or m' added to m) is also a possible DR for D where

$$m+m' = \langle V_m \cup V_{m'}, Con_m \cup Con_{m'} \rangle$$

(9) D = <Mary loves a man who owns a donkey>

Syntactic analysis:

1: Mary loves a man who owns a donkey

2: Mary   3: loves a man who owns a donkey

4: loves   5: a man who owns a donkey

6: man who owns a donkey

7: man   8: who owns a donkey

9: he owns a donkey

10: he   11: owns a donkey

12: owns   13: a donkey

14: donkey

m:

u,v,w

1: <Mary loves a man who owns a donkey, 1>

2: <u=Mary, 2>

3: <u loves a man who owns a donkey, 1>

4: <man who owns a donkey (v), 6>

5: <u loves v, 1>

6: <man (v), 7>

7: <v owns a donkey, 9>

8: <donkey (w), 14>

9: <v owns w, 9>

m is a possible DR for D. The occurrences in lines 1, 3, 5, 7 and 9 are occurrences of simple sentences, while the occurrences in lines 2, 4, 6 and 8 are occurrences of qualifiers, the occurrence in line 4 being non-atomic. 1 is an occurrence in L_1, while 2 is a decendant in $L_1'-(x)$ of the occurrence $\langle$Mary,2$\rangle$ in L_1.

Definition 24: Let m be a possible DR for D. An occurrence $\langle\delta,n\rangle$ in Con_m is called unreduced if Con_m contains no descendant of subexpressions of $\langle\delta,n\rangle$.
m is called maximal if every unreduced element in Con_m is an occurrence of either of the following kinds:

i) atomic sentence
ii) conditional
iii) universal sentence

Definition 25: A partial DRS (Discourse Representation Structure) is an ordered triple $\langle K,A,E\rangle$ where:

I K is a set of possible DR's such that
$\forall m_1,m_2 \in K(Con_{m_1} \cap Con_{m_2} = \emptyset)$

II $A,E \subseteq K \times K$ such that $A \cap E = \emptyset$ and $A \cup E$ is a partial function from K to K that can be extended to a partial ordering $<$ on K with the same domain as $A \cup E$.

III Unique origin of the occurrences of a DR:
$\forall m \in K[\exists m_1 \in K\ (m<m_1) \longrightarrow \exists m_2 \in K(m<m_2 \wedge \exists\theta_2 \in Con_{m_2} \forall\theta \in Con_m$ (θ a descendant of a subexpression of θ_2))].

IV The ordering preserves descendance for occurrences:

If $\bigcup_{m \in K} Con_m$ contains two occurrences θ_1 and θ_2, $\theta_1 \in Con_{m_1}$, $\theta_2 \in Con_2$ for m_1, $m_2 \in K$ such that θ_2 is a decendant of a subexpression of θ_1, then $m_1 \geq m_2$.

(10) D = <Every man loves a woman>

DRS: $\mathcal{D} = \langle K,A,E \rangle$ where $K = \{m_0, m_1, m_2\}$, $A = \{\langle m_1, m_0 \rangle\}$,
$E = \{\langle m_2, m_1 \rangle\}$

m_0:

<Every man loves a woman, 1>

m_1:

v
<man(v),3<

m_2:

u
<v loves a woman, 1>
<woman (u), 7>
<v loves u, 1>

When a DR $m \in Dom(A)$, <u>all</u> possible mappings from the discourse referents in m compatible with mappings for the DRS that make m true in the model will be necessary in the evaluation of truth for the discourse. On the other hand, when a DR $m \in Dom(E)$, the <u>exi</u>stence of such a mapping is sufficient. So in (10) we want the discourse to be true in a model if whenever v is mapped to a man in the model, there is a mapping for u such that u is a woman and v loves u.

<u>Definition 26</u>: Let m_1, $m_2 \in K$ be possible DRs in the DRS $\langle K,A,E \rangle$

<u>m_1 is immediately subordinate to m_2</u>, or $m_1 \leq_{A \cup E} m_2$

if

$\langle m_1, m_2 \rangle \in A \cup E$

If $\langle m_1, m_2 \rangle \in A$ we write $m_1 <_A m_2$ and
if $\langle m_1, m_2 \rangle \in E$ we write $m_1 <_E m_2$.

<u>m_1 is subordinate to m_2</u>, or $m_1 \leq m_2$

if

there is a chain of immediate subordinate DRs from m_1 to m_2.

If K contains a unique element m_0 such that $m_0 < m$ for every $m \in K$, then m_0 is the <u>principal element</u> of K.

<u>Notation 3</u>: $K^{\geq}(m) = \{m' \in K \mid m' \geq m\}$

$V_K = \cup_{m \in K} V_m$

$V_k^{\geq}(m) = \cup \{V_{m'} \mid m' \in K^{\geq}(m)\}$

$Con_K = \cup_{m \in K} Con_m$

<u>Definition 27</u>: Let $\theta \in Con_K$ be an occurrence in the DRS $D = \langle K,A,E \rangle$

θ is <u>unreduced</u> in D

if

Con_K contains no descendant of a subexpression of θ and θ is not atomic.

<u>Definition 28</u>: We say that a partial DRS $\langle K,A,E \rangle$ is <u>complete</u> if

i) Every element in K is maximal

ii) No occurrence in Con_K is unreduced.

Note that as a consequence of definition 25 and 28 i), any partial DRS cannot be extended (by adding occurrences and possibly some new DR's) to a complete DRS. If a subexpression of an occurrence θ has a descendant that is not also a descendant of a subexpression of a conditional or universal subexpression of θ, or of a subexpression on the form $\alpha\beta$, $\alpha \in T$, $\beta \in RC$, in some DR other than the DR containing θ, the DRs in the partial DRS cannot be made maximal without violating condition I in definition 25.

DRS construction rules.

Notation 4: Let $\gamma[v], v \in V$ denote

- the result of inserting v to the k-th position in ζ, if $\gamma \in RC \cup \bar{S}$ and γ comes from ζ by FR3,k or FR6,k.
- the result of replacing the main term in γ with v if γ is a simple non-atomic sentence.

Let $\mathfrak{D} = \langle K, A, E \rangle$ be an incomplete DRS, $m \in K$, m_0 the principal element in the greatest subset of K containing m having a principal element, $\theta \in Con_m$, θ unreduced in $\mathfrak{D}$, $\theta = \langle \gamma, k \rangle$.

C1. γ is a conditional, where $\langle \psi, r \rangle, \langle \chi, s \rangle$ are the occurrences of the antecedent and consesequent respectively.

Add to K: $m_1 = \langle \emptyset, \{\langle \psi, r \rangle\} \rangle$

$m_2 = \langle \emptyset, \{\langle \chi, s \rangle\} \rangle$

A: $\langle m_1, m \rangle$

E: $\langle m_2, m_1 \rangle$

C2. <u>γ is a simple sentence.</u>

Let α denote the term with maximal scope in γ, and r the index of the occurrence of α as the term with maximal scope in γ.

$$\text{Choose a } v \in \begin{cases} V_K^{\geq}(m) & \text{such that } v \text{ is suitable if } \alpha \text{ is a pronoun}^{11)} \\ V \setminus V_K & \text{otherwise} \end{cases}$$

α must be either of the following:

a) a pronoun

b) a proper name

c) an indefinite term

d) a universal term.

Let ξ denote $\alpha = 2$ in case b. In case c)-d), α is on the form <u>a</u> β, <u>every</u> β or <u>what</u> δ, <u>whoever</u> δ, let ξ denote $\beta(v)$ or $v\delta$ respectively.

According to whether α is of type a), b), c) or d):

a)			let m_1 denote m
b)	add $\langle\{v\},\{\langle\xi,r\rangle\}\rangle$	to m_0,	let m_1 denote m
c)	add $\langle\{v\},\{\langle\xi,r\rangle\}\rangle$	to m,	let m_1 denote m
d)	add $m_1 = \langle\{v\}, \langle\xi,r\rangle\}\rangle$	to K,	add $\langle m_1,m\rangle$ to A

If α is not the main term of γ, then the main term is on the form $\alpha\rho, \rho \in RC$. Let s denote the index of ρ in the main term of γ, and

$$\text{add } m' = \langle\emptyset,\{\langle\rho[v],s\rangle\}\ \text{ to } K, \text{ and } \langle m',m_1\rangle \text{ to } E$$

In case:

a)-c): add $<\emptyset,\{<\gamma[v],k>\}>$ to m_1

d) add $m_2 = <\emptyset,\{<\gamma[v],k>\}>$ to K and $<m_2,m_1>$ to E

C3. <u>γ is a qualifier</u>, γ is on the form $\beta\rho(u)$, $\beta\in CN$, $\rho\in RC$. Let r,s denote the indices of the occurrences of β,ρ in γ respectively.

add $<\emptyset,\{<\beta(u),r>,<\rho[u],s>\}>$ to m

C4. <u>γ is a sequence of sentences</u>, $\gamma = \gamma_1,\ldots,\gamma_n$ and $k_1,\ldots,k_n$ is the indices of the occurrences of $\gamma_1,\ldots,\gamma_n$ in γ respectively.

add $<\phi, \{<\gamma_1,k_1>,\ldots,<\gamma_n,k_n>\}>$ to m.

<u>Definition 29</u>: Any complete DRS obtained from a partial DRS $<K,A,E>$ by a sequence of applications of the rules C1-C4 is called a <u>completion</u> of $<K,A,E>$

<u>Definition 30</u>: A DRS $\mathcal{D} = <K,A,E>$ is a <u>DRS for the L_1-discourse $D=\gamma$</u> (relative to a syntax analyses for D>

if

$\mathcal{D}$ is obtained from $<\{<\emptyset,\{<\gamma,k>\}>\},\emptyset,\emptyset>$, where k is the index of γ in the syntax analysis for D, by a sequence of applications of the rules C1-C4.

<u>Well-definedness of the DRS-construction in L_1</u>

We want to show that starting with the partial DRS $\mathcal{D}=<\{<\emptyset,\{<\gamma,k>\}>\},\emptyset,\emptyset>$ for the L_1-discourse $D=\gamma$, a completion of $\mathcal{D}$ will give us a complete DRS for D. Moreover, we want this DRS to reflect the universal sentences/IVs, the conditionals and the non-restrictive relatives.

The proof goes by induction on the construction of DRSs as is sketched below:

Let $D=\gamma$ be a L_1-discourse.

<u>Induction hypothesis:</u> $\mathfrak{D} = \langle K,A,E \rangle$ satisfies the definitions 25 and 30, and if $\theta \in Con_m$, $m \in K$ and θ is unreduced in m, then

a) θ is unreduced in Con_K

or b) θ is reduced in Con_K, and θ is a $\left\{\begin{matrix}\text{universal}\\ \text{conditional}\end{matrix}\right\}$ and there is a pair of DRs $m_1, m_2 \in K$ such that $m_1 <_A m$, $m_2 <_E m_1$ and m_1 contains an occurrence of a descendant of the $\left\{\begin{matrix}\text{universal term}\\ \text{antecedent}\end{matrix}\right\}$, m_2 contains an occurance of a descendant of $\left\{\begin{matrix}\theta\\ \text{the consequent}\end{matrix}\right\}$

1. $\mathfrak{D} = \langle \{\langle \emptyset, \{\langle \gamma, k \rangle\} \rangle\}, \emptyset, \emptyset \rangle$ is a parial DRS for D as $\mathfrak{D}$ trivially satisfies the induction hypothesis.
2. If $D = \langle K,A,E \rangle$ satisfies the induction hypothesis, the application of one of the rules C1-C4 for DRS-construction preserves the conditions of the induction hypothesis.
3. Each of the rules C1-C4 for DRS-construction leads to the reduction of one occurrence in Con_K together with the addition of some new occurence(s) to Con_K. The occurrences added to Con_K will have smaller parts of the longest subexpression also an expression in L_1 of the reduced occurence, as their longest subexpression also an expression in L_1. As γ is finite, having a finite number of subexpressions, this process must stop because no occurrence

in Con_K is unreduced in $\mathcal{D}$. When no occurrence in Con_K is unreduced in $\mathcal{D}$, every DR in K must be maximal since $\mathcal{D}$ satisfies the induction hypothesis, and $\mathcal{D}$ satisfies definition 28, and is thus a complete DRS for the L_1-discourse D.

We have now shown that a completion of $D = \langle\{\langle\emptyset,\{\langle\gamma,k\rangle\}\rangle\},\emptyset,\emptyset\rangle$ will give a complete DRS for the discourse $D = \langle\gamma\rangle$. Note that the construction rules C1-C4 preserves the property that no conditional or universal occurrence θ is reduced in the DR that contains it. In a complete DRS for D, there will thus be a pair of DRs $m_1, m_2 \in K$ such that $m_2 <_E m_1$, $m_1 <_A m$, $m \in K$ and $\theta \in Con_m$, and $\langle m_1, m_2\rangle$ represents θ in the terminology of Kamp (1981). Moreover, if $\theta \in Con_m$, $\theta = \langle\alpha\beta\gamma, j\rangle$, β a non-restrictive relative to the term α, then in a complete DRS for D, there is a $m' \in K$, $m' <_E m$ such that Con_{m}' contains θ', $\theta' = \langle\beta[v], k\rangle$ and Con_m contains $\langle\gamma[v], j\rangle$, v a discourse referent.

Note also that L_1 is an extension of L_0 in Kamp (1981) in the sense that if γ is a L_0-expression then γ is also a L_1-expression, and if $\mathcal{D} = \langle K_1, A, E\rangle$ is a complete DRS in L_1 for the L_0-discourse D, then K_0 will be a L_0-DRS for D, where K_0 is exactly like K_1 except for the deletion of all occurrences of non-atomic qualifiers and sequences of sentences.

2.3.3 Truth

Let $M = \langle U_M, F\rangle$ be a L_1-model and let $\mathcal{D} = \langle K, A, E\rangle$ be a complete DRS for the L_1-discourse D. Λ denotes as usual the empty function

D1: $\underline{D \text{ is true in } M \text{ on the reading } \mathcal{D}}$ (relative to the function $c: X \to U_M$, $X \subseteq V$) if

$$\Lambda(\cup c) \ ver_E \ m_0 \quad \text{in} \quad M$$ 12)

D2: If $m \in K$, $f: X \to U_M$ and $V_K^{\geq}(m) \subseteq X \subseteq V$,

$$\underline{f \text{ verifies } m \text{ in } M, \ f \text{ ver } m}$$

if

i) $f \ ver_A m$ if $m \in dom(A)$

ii) $f \ ver_E m$ if $m \in dom(E)$

D3: $f \ ver_A m$ iff $\forall g \in U_M^{V_m} (f \cup g \ ver_e m \to \forall m' (m' <_{A \cup E} m \to f \cup g \ ver \ m'))$

$f \ ver_E m$ iff $\exists g \in V_M^{V_m} (f \cup g \ ver_e m \wedge \forall m' (m' <_{A \cup E} m \to f \cup g \ ver \ m'))$

This is well-defined since $A \cup E$ is a total function on $K \setminus \{m_0\}$, $A \cap E = \emptyset$ and all chains of subordiniate DRs end in an atomic DR (in the domain of E).

D4: If m is a DR, $X \subseteq V$ such that every atomic expression in Con_m is a $L_1'(X)$-expression, and f is a function $f: X \to U_M$,

$\underline{m \text{ is elementary verfied in } M \text{ by } f, \ f \ ver_e m}$

if

every atomic occurrence in Con_m is true in M by f.

D5: If $X \subseteq V$, $\theta = \langle \phi, k \rangle$ is an occurrence of an atomic sentence in $L_1'(X)$, then ϕ is on one of the following forms:

a) $u = \alpha$ where $u \in V$, α a proper name

b) $\alpha(u)$ where $u \in V$, α a basic CN

c) $u\alpha$ where $u \in V$, α a basic IV

d) $u\alpha v$ where $u, v \in V$, α a basic TV

If f is a function, $f{:}X \rightarrow U_M$,

<u>θ is true in M by f</u>

if

a) $f(u) = F(\alpha)$

b,c) $f(u) \in F(\alpha)$

d) $\langle f(u), f(v)\rangle \in F(\alpha)$

(11) D = <Enhver hest som Mari kjenner en mann som eier, kaster ham av>

DRS: = $\langle K,A,E\rangle$ where

$K=\{m_0,m_1,m_2\}$ $A = \{\langle m_0,m_1\rangle\}$ $E = \{\langle m_2,m_1\rangle\}$

m_0:

u

<Enhver hest som Mari kjenner en mann som eier, kaster ham av, 1> C2d)

<u=Mari, 7>

m_1:

w, v

<hest som Mari kjenner en mann som eier(w), 3> C3

<hest (w), 4>

<Mari kjenner en mann som eier w, 5> C2b)

<u kjenner en mann som eier w, 5> C2c)

<mann som eier w(v), 11> C3

<u kjenner v, 5>

<mann (v), 12>

<v eier w, 13>

m_2:

<w kaster ham av, 1> C2a)

<w kaster v av, 1>

which gives D the truth conditions:

$\exists$ x(x=MARI $\wedge$ $\forall$y,z[(HEST(z) $\wedge$ MARI KJENNER y $\wedge$ MANN(y) $\wedge$ y EIER z)$\rightarrow$ z KASTER y AV])

with the coreferentiality as indicated in the first section.

At first sight it may be a surprise that an existential term like <u>en mann</u> is represented by a $\forall$-quantifier. We recoginze, however, the following equivalence from first order logic:

When x is not free in ψ:

$$\vdash_{\text{1.O.Logic}} \forall x(\phi \rightarrow \psi) \leftrightarrow (\exists x\phi \rightarrow \psi)$$

An existential term in the antecedent being bound by a $\forall$-quantifier outside the implication is thus in complete accordance with first order logic. At coreferentiality between the antecedent and the consequent, the variable is free in the consequent, and the quantifier movement is therefore not allowed.

2.4 <u>Some remarks on the Fragment L_1</u>

Relatively to a syntactic analysis for a discourse in L_1, each occurrence in the discourse can be reduced in only one way, except possibly for the choice of discourse referent. Thus any two complete DRSs for a discourse D based on the same syntactic analysis of D will be alphabetic variants of each other except possibly for the choice of discourse referents for pronomina. So far, we have ignored the consequences of the reduction order to the coreferentiality. To prevent backwards binding we might want the reduction order to be from left to right. This can be implemented by requiring the nodes in the syntactic analyses to be numbered according to the reading direction, and the occurrences to be reduced in increasing order.

Kamp does not give any definition for deictic use of pronomina. One possibility is to split off a subset of the set of discourse referents V, say $V' \subseteq V$ (or to choose some other set V'). Then given an assignement $c:V' \rightarrow U_M$, we could choose discourse referents for the deictic pronomina from V', for other terms from $V \setminus V'$. This strategy is indicated in the truth definitions.

For a fragment permitting more than one non-restrictive relative to the same term, the following modifications can be made to the syntax definition of L_1:

Replace FR6,k by

FR6'n,$k_1,\ldots,k_n$: If $\phi_1,\ldots,\phi_n \in S$ and the k_i-th word in ϕ_i is α_i where α_i is either he/him she/her or it for $i=1,\ldots,n$, then $\beta_1\phi_1',\ldots,\beta_n\phi_n' \in RC$ (the last comma may be replaced by 'and'), where ϕ_i' comes from ϕ_i as in FR3,k, and β_i is

$$\left\{\begin{array}{l}\text{who/whom/which according to}\\ \text{whether the } k_i\text{-th word in}\\ \phi_i \text{ is he,she/him,her/it}\\ \text{or}\\ \text{what (only if } n=1)\end{array}\right\}$$

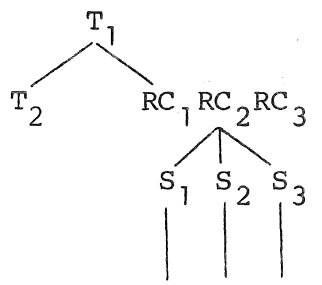
The rules making other use of RCs than as non-restrictive relatives needs a restriction against the RC being made up from more than one S. In addition, the first part of notation 4 must be replaced by:

Let $\gamma[v]$, $v \in V$ denote

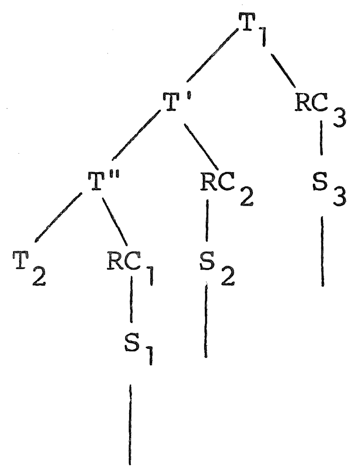
- the result of inserting v to the k_i-th position in ζ_i for each $i=1,\ldots,n$, where $\gamma \in RC \cup \bar{S}$ and γ comes from $\zeta_1,\ldots,\zeta_n$ by FR3,k and $n=1$ or FR6'$n,k_1,\ldots,k_n$

The alterations given above express a non-recursive analysis of terms with more than one non-restrictive relative. The non-restrictive relatives are understood as a sequence of co-ordinate clauses rather than as a hierarchy of terms with one non-restrictive relative to each term. Some alternative strategies for the analysis of non-restrictive relatives can be illustrated as follows:

1) co-ordinate non-restrictive relatives subordinate to only one term



2) A hierarchy of terms with one non-restrictive to each term.



DRS_1: $\langle m_2, m_1 \rangle \in E$

DRS_2: a) $\{\langle m_i, m_1 \rangle \mid i=2,\ldots,\} \in E$

b) $\{\langle m_i, m_{i-1} \rangle \mid i=2,\ldots,4\} \in E$

m_1: [T_2] m_2: [$S_1 S_2 S_3$]

m_1: [T' / T'' / T_2] m_2: [S_1] m_3: [S_2] m_4: [S_3]

which will give truth conditions according to the following patterns, where x,y,z are variables corresponding to discourse referents for terms that have their first occurrence in S_1, S_2, S_3 respectively:

for DRS_1: $\exists xyz(\phi_1 \wedge \phi_2 \wedge \phi_3)$ (if S_i is ϕ_i for i=1,2,3)

$DRS_{2a)}$: $\exists x(\phi_1) \wedge \exists y(\phi_2) \wedge \exists x(\phi_3)$

$DRS_{2b)}$: $\exists x(\phi_1 \wedge \exists y(\phi_2 \wedge \exists z\phi_3))$

If x is not free in ϕ_2, ϕ_3, y not free in ϕ_1, ϕ_3 and z not free in ϕ_1, ϕ_2, the three are equivalent. If y is not free in ϕ_1, and z is not free in ϕ_1, ϕ_2, 1) and 2b) are equivalent.

I have chosen alternative 1) for the following reasons:

i) Intuitively I understand non-restrictive relatives as a string of additional informations to one and the same term, and not as isolated additional information to a hierarchy of terms.

ii) Alternative 1) (and 2b)) gives the possibility of coreferentiality between coordinate clauses, as we would need for instance in:

(12) You know Bill, the one who met a girl at this party last year, who took her to Hawaii spending all his savings, who was then left by his wife and who has been a heavy drinker ever since,....

iii) If the reduction order of the clauses are from left to right, 1) and 2b) always give equivalent truth conditions. 1) is then preferable as it gives a simpler DRS.

No restrictions has so far been made with respect to the kind of sentences relative clauses can be made from. It seems quite clear that only simple sentences can be used, but there might also be other restrictions regarding the complexity of the sentence. Other restrictions, like island constraints, restrictions against crossing depencies etc. may also be imposed on the relative clause formation.

One of the main virtues of the "Cooper stores" is the great flexibility with respect to quantifier scope (and order) and coreferentiality. The only limitation to this flexibility is that "donkey-structures" (including the structures in (3) and (4)) are not feasible.

The Cooper interpretation process builds interpretations for compound phrases from the interpretations already calculated for the constituents. The interpretation process is difficult to follow, and even my intuition regarding the construction of sentences being vague, I do not think it looks much like what goes on at human sentence construction. As an illustration of what a language user grasps at hearing/reading a text, I think the DRS-system gets far better off.

The DRS-system as it stands puts obvious constraints on the quantifier scope and order, and with that on the coreferentiality. The quantitier scope and order is determined by the left to right direction, and the possibility of choice for discourse referents for pronouns is determined by reduction order and the subordinate relation. It is, however, possible to modify the rules for DRS-construction to obtain flexibility in quantifier scope and order:

Replace the first period in rule C2 (the one beginning with "Let" and ending with "γ") with:

If $\gamma \in S$ and α' is a term in γ, then

$$\text{let } \alpha \text{ denote } \begin{cases} \beta & \text{if } \alpha' \text{ is on the form } \beta\rho, \beta \in T, \rho \in RC \\ \alpha' & \text{otherwise} \end{cases}$$

and let r denote the index of the occurrence in question of α.

With this modification, we can alter the quantifier scope in L_1. Note that two complete DRSs for the same discourse D need at this not be alphabetic variants of each other except possibly for the choice of coreferentiality for pronouns. Even with a restriction on the reduction order, a term can now corefer with any suitable pronoun in its scope, also pronouns occuring to the left of it in the sentence. This is far too much flexibility and one can easily imagine constraints that one would want to impose on this modification: constraints against quantifying out of sentences (including relative clauses), island constraints etc.

Some unanswered questions

The preferable reading of a sentence is normally the reading with the quantifier scope and order as given from left to right in a written representation of the sentence. Several examples are put forward to support the need of reordering the scope and order, among which are the following:

(13) Every Englishman admires a woman
a) namely his mother
b) namely the Queen
(Engdahl 1980)

(14) Guinevere has a bone in every corner of the house
(Rodman 1976. I got it from Cooper 1983, Ch V, 2.5)

A reversed quantifier order is not necessary for the readings we want for (13)b, as we can obtain the correct interpretation also with the left-to-right order. (14) however gives evidence for the need of wide-scope mechanisms, unless there is some other way of interpreting indefinite descriptions (i.e. something like value-free and value-loaded interpretations of Barwise and Perry (1980) or the approach taken in situation semantics(Barwise and Perry 1983)). Such techniques may also be adequate in the interpretation of sentences like

(15) As you go north through the valley, the towns get smaller
(Hellan 1980)

A strong candidate against the DRS-treatment of "donkey-sentences" is the following example:

(16) Every man who has a daughter thinks she is the most beautiful girl in the world. (Cooper 1979)

Cooper asserts that the DRS-treatment would "commit any father of more than one daughter to the contradictory belief that each of his daughters is the most beautiful girl in the world" (Cooper 1979).

3. Application of the DRS-analysis to Situation Semantics.

3.1. Introduction.

In part 1 and 2 we discussed questions of quantifier scope and anaphoric relations between noun phrases through a study of the storage mechanisms developed by R. Cooper (Cooper 1983) and the theory of Discourse Representation Systems (Kamp 1981).

The storage systems turned out to be successful in treating many problems of quantifier scope in natural languages, but seem unable to handle sentences where there is a conflict with respect to the need of wide-scope mechanisms. Extension of the analysis of Kamp (1981) to cover the syntactic fragment of Cooper (1983) enabled us to obtain a satisfactory treatment of some of the examples that could not be handled by the storage method.

The way the DRS describe simple parts of the world by each DR, and systematic relations between these simple parts also gives a picture of what a language user 'grasps' at hearing/reading a text. This feature is strongly reinforced by the correspondence between the theory of DRS and situation semantics (Barwise and Perry 1983), a promising alternative to the traditional possible world semantics.

Situation semantics is based on a set of individuals A, a set of n-ary relations over A for each $n \in \mathbb{N}$ and a set L of space-time locations. The system also provides a rich number of concepts to describe parts of the model (reality), events and courses of events, together with a theory of meaning for the model.

How does language get into this machinery so powerful with respect to describing reality? To have the means to describe reality is one thing, to know what an expression in a (human) language means is another. Barwise and Perry write:

> "The linguistic meanings of expressions in a language are conventional constraints on utterances. To study semantics is to attempt to spell out these constraints, to spell out what it is the native speaker knows in knowing what utterances of his language mean.
>
> .....
>
> Given an indicative sentence $\underline{\phi}$, we think of the meaning of $\underline{\phi}$ as a relation $u[\![\phi]\!]e$ between situations u in which ϕ is uttered and situations e described by such utterances. The relation constrains both u and e."
>
> (Barwise and Perry 1983, Ch. 6)

The theory of DRS gives the means for finding the relations that have to hold between the individuals assigned to the discourse referents to satisfy the truth conditions of the sentence. Can this be used as a means for evaluating the same aspect of the meaning relation in situation semantics? My answer to the question is yes, and I will in the following indicate how this can be done. The event-type in situation semantics describe elements in a situation structure in very much the same way as the DR describes (sub)models of a model for DRS-theory. I will not give any complete system or exact specification of the correspondence rules in this paper, but merely illustrate the use of DRS-theory in construction of the meaning relation $[\![\phi]\!]$ by means of two examples. Familiarity with situation semantics is assumed, and concepts and rules from "Situations and Attitudes" will therefore be used without further comments. The definition of the most central concepts will, however, be included.

3.2. <u>The meaning relation $u[\![\phi]\!]e$.</u>

Recall the following concepts from "Situations and Attitudes":

<u>Definition 31</u>: Let <u>DU</u> be the event-type

$$\begin{aligned} DU: = \text{at } \ell:\ & \text{speaking, } \underline{a};\ \text{yes} \\ & \text{addressing, } \underline{a},\underline{b};\ \text{yes} \\ & \text{saying, } \underline{a},\underline{\alpha};\ \text{yes} \end{aligned}$$

If there is one and only one anchor f for DU such that a c.o.e. or state of affairs (situation) d is of type DU, then we call d a <u>discourse situation</u>.

<u>Definition 32</u>: An <u>utterance situation</u> u consists of a discourse situation d and possibly of a situation c called the speakers connections consisting of situations of the type

$$\begin{aligned} CU_{\underline{\beta}}: = \text{at } \underline{\ell}_{\underline{\beta}}:\ & \text{saying, } \underline{a},\underline{\beta};\ \text{yes} \\ & \text{referring to, } \underline{a},\underline{\beta},\underline{\beta'};\ \text{yes}^{13)} \end{aligned}$$

where $\underline{\ell}_{\underline{\beta}} \subseteq \underline{\ell}$, for one or more of the subexpressions $\underline{\beta}$ in α.

<u>Notation 5</u>: If U is an utterance situation-type for ϕ, and β is a subexpression of ϕ such that $CU_\beta \subseteq U$, we let $CU(\beta) = \langle \underline{\beta}', CU_\beta \rangle$, that is, the role of the referent of β in CU_β. If $u = dc$ is an utterance situation of ϕ, we write $c(\beta)$ for the referent of $\underline{\beta}$ in c, and let ℓ_d, a_d, b_d, α_d denote the values of $\underline{\ell},\underline{a},\underline{b},\underline{\alpha}$ in d respectively.

<u>Definition 33</u>: A <u>statement</u> Φ of an expression ϕ is an ordered triple $\Phi = \langle d,c,\phi \rangle$ where $u = d,c$ is an utterance situation of ϕ.

An expression ϕ expresses a meaning relation, a conventional constraint between utterance situations and described situations. We adopt the notation from Barwise and Perry (1983) of writing $[\![\phi]\!]$ for $MO_{C(\phi)}$, and write $MF_{[\![\phi]\!]}$ for $MF_{C(\phi)}$. $[\![\phi]\!]$ will be on the following form:

<u>Notation 6</u>: $[\![\phi]\!] := \text{at } \ell_u\text{: involves, U,E; yes}$

$$\vdots$$

where - $U \supseteq DU'[\langle \underline{\alpha}, \underline{\phi} \rangle] \cup CU_\phi$, CU_ϕ posibly empty,
DU' is DU with the addition of uniqueness reguirements for the linguistic roles.
- E is specified by the analysis of ϕ as is to be illustrated below.

$[\![\phi]\!]$ may contain more than a simple constraint.

Note that this means that

(17) $u \in MF_{[\![\phi]\!]}$ <u>iff</u> u is of type $U = DU' \cup CU_{\underline{\phi}}$ and $\underline{\alpha}_d = \underline{\phi}$ and .....
<u>iff</u> u is an utterance situation of type $DU \cup CU_\phi$ and $\underline{\alpha}_d = \phi$ and .....

The simple constraint $C := \text{at } \ell_u\text{: involves U,E; yes}$ in $[\![\phi]\!]$ holds a unique position for $[\![\phi]\!]$ as a meaning relation $u[\![\phi]\!]e$. That is, if $[\![\phi]\!]$ contains other requirements, additions[14)] or constraints, these are subordinate to C. No definition of how non-simple constraints work is given in "Situations and Attitudes", and we will propose the following defective definitions to comply with meaning constraints $[\![\phi]\!]$:

Definition 34: If $[\![\phi]\!]$ is a constraint on the form as in notation 6, then

i) $u \in MF_{[\![\phi]\!]}$ if u is of type U

ii) If $u \in MF_{[\![\phi]\!]}$, then $u[\![\phi]\!]e$ if

$$\forall f(U[f] \subseteq u \Rightarrow \exists g(E[f \cup g] \subseteq e \wedge \ldots .))$$

where f, g are anchors with domains $\subseteq$ the set of indeterminates for U, E respectively, and the additional conditions denoted by .... in $[\![\phi]\!]$ in notation 6 will specify .....

This definition is equivalent to the corresponding definitions for simple constraints in "Situations and Attitudes".
We also define

Definition 35: The interpretation of a statement $\Phi = \langle d, c, \phi \rangle$ is

$$[\![\Phi]\!] = \{e \mid d, c[\![\phi]\!]e\}$$

If $u \in MF_{[\![\phi]\!]}$, we have

(18) $$[\![\Phi]\!] = \{e \mid d, c[\![\phi]\!]e\} = \{e \mid uMO_{[\![\phi]\!]}e\}$$
$$= \{e \mid U[f] \text{ is part of } u \text{ and } e \text{ is of type } E[f] \text{ for a total anchor } f \text{ for } U\}$$

Note that Barwise and Perry seems to give a different definition of $[\![\phi]\!]$ in Ch. 6 and in the definition of ALIASS (Barwise and Perry 1983). The form given here is more general, and in accordance with the outlines and explanations of "Situations and Attitudes".

In part 2, we defined the conditions for a discourse $D = \langle\phi\rangle$ to be true in a total model M on a reading $\mathcal{D} = \langle K,A,E\rangle$ (relative to a function c). The conditions require that the total model M contains a specified submodel, or that it systematically contains several specified submodels, depending on the type of the sentence. In situation semantics, the underlying structure is more sophisticated:

<u>Definition 36:</u> A structure $\mathfrak{M}$ of situations consist of a collection M of c.o.e's (the factual c.o.e's) with a non-empty subcollection M_0 (the actual c.o.e's) satisfying:

i) Every $e \in M_0$ is coherent

ii) $e \in M \wedge e_0 \subseteq e \Rightarrow e_0 \in M$

iii) X is a subset of $M \Rightarrow \exists e \in M_0 (\forall e' \in X (e' \subseteq e))$

iv) If C is any constraint in M, then M respects C.

Our program is now to characterize those situations in a situation structure $\mathfrak{M}$ that are described by the expression ϕ on the basis of the utterance situation of ϕ. More precisely: we want to specify the meaning relation $[\![\phi]\!]$ expressed by ϕ on the reading $\mathcal{D} = \langle K,A,E\rangle$ of $D = \langle\phi\rangle$, depending systematically on the utterance situation of ϕ. $[\![\phi]\!]$ is to be specified so that given a statement of ϕ, $\Phi = \langle d,c,\phi\rangle$, if a c.o.e $e \in [\![\Phi]\!]$, that is, if e is a model for ϕ in the sense of situation semantics, and ϕ is a simple indicative sentence without non-restrictive relatives, then e will be a model verifying ϕ with respect to $\mathcal{D}$ in the DRS-theory. If Φ is true in $\mathfrak{M}$, then a model verifying ϕ with respect to $\mathcal{D}$ in the DRS-theory can be constructed by union of situations in $\mathfrak{M}$, ignoring locations, if the difference of locations does not contribute substantially to the truth conditions of ϕ in $\mathfrak{M}$.

Given an utterance situation u such that $u \in MF_{[\![\phi]\!]}$, $e \in [\![\Phi_u]\!]$ <u>iff</u> $u[\![\phi]\!]e$. ϕ will be absolutely true in $\mathfrak{M}$ if $\mathfrak{M}$ respects $[\![\phi]\!]$, and for simple indicative sentences, ϕ is true uttered in u if $[\![\Phi_u]\!] \cap M_0 \neq \emptyset$. For sentences that require systematic patterns of simple situations, like universals or conditions, this may be a too strong requirement. We may want a universal statement to be true for a structure if the structure respects the conditions set by the statement without the conditions themselves being part of the structure. In this paper, however, we will not go further into problems concerning universals and conditions, but concentrate on simple indicative sentences without universal terms.

3.3. <u>Specification of $u[\![\phi]\!]e$ by means of DRS-theory. Two examples</u>.

The meaning relations in these examples are not complete, and important aspects of the meaning of expressions, like tense, will be ignored or treated in an ad hoc manner. The focus is on the relations between the individuals referred to by the noun phrases, and we make use of the following notation:

<u>Notation 7</u>: For every discourse referent $v \in V$ in the theory of DRS, we let $\underline{v}$ denote a special individual indeterminant in the situation semantics.

L_1 has unique syntax representations (up to the numbering of the nodes in the syntax tree) for the two examples we present, and syntax analyses will therefore be omitted. The numbering of occurrences corresponds to a leftmost branch numbering of the syntax trees. We also make use of the following notation:

Notation 8: Syntactic expressions are underlined. Capital letters are used in denoting real individuals or relations over A, apart from the special system relations.

Example 1: ϕ_1 = Jackie bites Molly

$\mathcal{D}_1 = \langle\{m_0\}, \emptyset, \emptyset\rangle$

m_0:

v,w
<Jackie bites Molly, 1>
<v = Jackie, 2>
<v bites Molly, 1>
<w = Molly, 5>
<v bites w, 1>

$\mathcal{D}_1$ gives the meaning relation:

$[\![\phi_1]\!]$: = at ℓ_n: involves, u,E_0; yes

E_0: at ℓ_0: same, v, CU_1(Jackie); yes
same, w, CU_1(Molly); yes
BITES, v,w; yes

E_0 corresponds closely to the DR m_0, with one line in E_0 corresponding to each atomary occurrence in m_0. Note that E_0 depends on the utterance situation for the referents for Jackie and Molly. A specification of U in $[\![\phi_1]\!]$ (which we omit here, as we are mainly interested in the right part of the relations $[\![\phi]\!]$ rising from expressions ϕ) should therefore include $CU_{\underline{Jackie}}$ and $CU_{\underline{Molly}}$-situation types as part of the speakers connections description (see Definition 32). E_0 should also depend on the utterance situation in the choice of referent for $\underline{\ell}_0$, according to some rule for present tense.

Notation 9: If r is a n-ary relation, e a c.o.e, we write

$$r_{e,\ell}, x_1, \ldots, x_n \quad \text{for} \quad \text{in } e: = \text{at } \ell : r, x_1, \ldots, x_n; \text{ yes}$$

If $n = 1$ or $n = 2$, we may also write

$$r_{e,\ell}(x_1) \quad \text{for} \quad r_{e,\ell}, x_1 \quad \text{and} \quad x_1 r_{e,\ell} x_2 \quad \text{for} \quad r_{e,\ell}, x_1, x_2$$

respectively. We may denote $same_{e,\ell}$ by $=_{e,\ell}$.

Now, given a statement of ϕ_1, $\Phi_1 = \langle d_1, c_1, \phi_1 \rangle$, where $c_1(\underline{Jackie}) = \text{JACKIE}$ and $c_1(\underline{Molly}) = \text{MOLLY}$, (and such that $u \in MF_{[\![\phi_1]\!]}$) we get

$e \in [\![\Phi]\!]$ iff $u \in MF_{[\![\phi_1]\!]}$ and $u[\![\phi_1]\!]e$

iff u is of type $U[c] \Rightarrow e$ is of type $E_0[c]$ for every total anchor c for U.

iff $(c_1(\underline{Jackie}) = \text{JACKIE} \wedge c_1(\underline{Molly}) = \text{MOLLY})$

$$\Downarrow$$

$$\exists f \in A^{\{\underline{v},\underline{w}\}}(f(\underline{v}) =_{e,\ell_0} c_1(\underline{Jackie}) \wedge f(\underline{w}) =_{e,\ell_0} c_1(\underline{Molly}) \wedge f(\underline{v}) \text{ BITES}_{e,\ell_0} f(\underline{w}))$$

iff $\exists f \in A^{\{\underline{v},\underline{w}\}}(f(\underline{v}) =_{e,\ell_0} \text{JACKIE} \wedge f(\underline{w}) =_{e,\ell_0} \text{MOLLY} \wedge f(\underline{v}) \text{ BITES}_{e,\ell_0} f(w))$

This is exactly the condition for e to verify ϕ_1 on the reading $\mathcal{D}_1$ in the theory of DRS.

Example 2: ϕ_2 = Pedro, who owns a donkey, beats it.

In ϕ_2, who owns a donkey is a non-restrictive relative. A donkey is therefore not to corefer with it (unless by accident), and the use of it is deictic.

$\mathcal{D}_2 = \langle\{m_0, m_1\}, \emptyset, \{\langle m_1, m_0\rangle\}\rangle$ $\qquad$ $V' = \{u\}$

$c(v)$ = CHIQUITA

m_0:

w
<Pedro, who owns a donkey, beats it, 1>
<w = Pedro, 3>
<w beats it, 1>
<w beats v, 1>

m_1:

y
<w owns a donkey, 4>
<donkey (y), 10>
<w owns y, 4>

$\mathcal{D}_2$ gives the meaning relation:

$[\![\phi]\!]_2$: = at ℓ_u: involves, U, E_0; yes

at $\underline{\ell}^2$: besides, E_0, E_1; yes

E_0: at $\underline{\ell}_0$: same, $\underline{w}$, CU_ϕ(Pedro); yes

BEATS, $\underline{w}$, CU_ϕ(it); yes

E_1: at $\underline{\ell}_1$: DONKEY, $\underline{y}$; yes

OWNS, $\underline{w}, \underline{y}$; yes

This specification of $[\![\phi_2]\!]$ corresponds closely to the DRS $\mathcal{D}_2$. Each DR, E_0 and E_1 corresponds to the DRs m_0 and m_1 in the same way as E_0 to m_0 in example 1. The relation $<_E$ between

m_1 and m_0 is implemented by use of a new concept called addition. The addition is based on a primitive relation 'besides' in very much the same way as the constraint is based on 'involves':

<u>Definition 37</u>: An <u>addition</u> B is a state of affairs on the form

$$B: = \text{at } \underline{\ell}\text{: besides, } S_0, S_1\text{; yes}$$

where S_0, S_1 are uniform schemata.

<u>Definition 38</u>: e_0 is <u>meaningful with respect to B</u>, $e_0 \in MF_B$ if e_0 is of type S_0.

<u>Definition 39</u>: If $e_0 \in MF_B$, e_1 is a <u>meaningful addition to e_0 with respect to B</u>, $e_0 MA_B e_1$, if

$$e_0 \text{ is of type } S_0[f] \text{ and } e_1 \text{ is of type } S_1[f]$$

for a total anchor f for S_0.

<u>Definition 40</u>: A structure of situations $\mathfrak{M}$ <u>respects B</u> if

$$\forall e_0 \in M_0 (e_0 \in MF_B \rightarrow \exists e_1 \in M(e_0 MA_B e_1)).$$

We also imagine a 5th requirement added to the definition of a situation structure to respect every factual addition.

Now, assuming that $\underline{\ell}^2$ has to be anchored in some systematic way according to $\underline{\ell}_d$ and some rule for the present tense, we get the following result given a statement of ϕ_2, $\Phi_2 = \langle d_2, c_2\phi_2\rangle$, where $c_2(\underline{Pedro})$ = PEDRO and $c_2(\underline{it})$ = CHIQUITA (and $u \in MF_{[\![\phi_2]\!]}$):

$e \in [\![\Phi_2]\!]$ <u>iff</u> $uMF_{[\![\phi_2]\!]}$ and $u[\![\phi_2]\!]e$

<u>iff</u> $U[c] \subseteq u \Rightarrow \exists g \in A^{\{\underline{w}\}}(E_0[g \cup c] \subseteq e \wedge \exists e'(eMA_B e'))$
for every c with domain $\subseteq$ the set of indeterminates used in U.

<u>iff</u> $c_2(\underline{Pedro}) = \text{PEDRO} \wedge c_2(\underline{it}) = \text{CHIQUITA}$
$\wedge\ \exists g \in A^{\{\underline{w}\}}(E_0[g \cup c] \subseteq e \wedge \exists e' \exists g' \in A^{\{\underline{y}\}}(E_1[c \cup g \cup g'] \subseteq e'))$

<u>iff</u> $c_2(\underline{Pedro}) = \text{PEDRO} \wedge c_2(\underline{it}) = \text{CHIQUITA}$
$\wedge\ \exists g \in A^{\{\underline{w}\}}(g(\underline{w}) =_{e,\ell_0} c_2(\underline{Pedro}) \wedge g(\underline{w})\ \text{BEATS}_{e,\ell_0}\ c_2(\underline{it})$
$\wedge\ \exists e' \exists g' \in A^{\{\underline{y}\}}(\text{DONKEY}_{e',\ell_1}(g'(\underline{y}))$
$\wedge\ g(\underline{w})\ \text{OWNS}_{e',\ell_1}\ g'(\underline{y})))$

<u>iff</u> $\exists g \in A^{\{w\}}(g(\underline{w}) =_{e,\ell_0} \text{PEDRO} \wedge g(\underline{w})\ \text{BEATS}_{e,\ell_0}\ \text{CHIQUITA}$
$\wedge\ \exists e' \exists g' \in A^{\{\underline{y}\}}(\text{DONKEY}_{e',\ell_1}(g'(\underline{y}))$
$\wedge\ g(\underline{w})\text{OWNS}_{e',\ell_1}\ g'(\underline{y})))$

In the DRS-theory we have:

$D = \langle\phi_2\rangle$ is true in M on $\mathcal{D}_2$ relative to c

<u>iff</u>

$$\exists f \in U^{\{w\}}(f(w) = \text{PEDRO} \wedge f(w) \text{ BEATS } c(v)$$
$$\wedge \exists f' U^{\{y\}}(\text{DONKEY}(f'(y)) \wedge f(w) \text{ OWNS } f'(y))).$$

In this example, $[\![\phi_2]\!]$ is not a simple constraint. It constrains the situation structure in such a way that it requires a certain pattern of situations. The interpretation, however, is not a set of such patterns, but of situations e satisfying the conditions of the main clause without the non-restrictive relative - if the required pattern of situations can be found in the situation structure for e, that is.

The truth condition $[\![\Phi_2]\!] \cap M_0 \neq \emptyset$ will not be affected by this choice of what situations are described, as $e \in [\![\Phi_2]\!]$ only provided the existence of e'. Models for $\mathcal{D}_2$ in the sense of DRS-theory will constitute the "best approximations" to situation structures $\mathfrak{M}$ such that $e, e' \in M$, if locations are ignored.

There may be reasons for preferring other ways to implement $<_E$ when not preceeded by $<_A$, like using schemata in the main constraint in $[\![\phi]\!]$ or choosing to let the addition be the situation-type of the described situation. The reason for the choice made here is that $[\![\Phi]\!]$ remains a set of situations that can be regarded as described by the utterance of ϕ, and does not become a set of sets of situations nor a set of situations that are additions. ϕ_2 talks about the beating that Pedro does, and not about the relation between the non-restrictive relative an the main clause.

3.4. <u>Conclusion</u>.

We have now given some ilustrations of the connection between the theory of Discourse Representation Structures (Kamp 1981) as set out in part 2, and situations semantics (Barwise and Perry 1983). By this we have tried to show that a complete DRS for an expression ϕ generates substantial parts of the meaning relation $[\![\phi]\!]$ in situation semantics. Focus is on the meaning relation $[\![\phi]\!]$, and not on the interpretation $[\![\Phi]\!]$ of statements of ϕ, though we have also carried out calculations for the elements e of $[\![\Phi]\!]$ in order to study the truth conditions of the statements obtained by using $[\![\phi]\!]$.

Many other interesting aspects of the meaning of expressions may be worked into this frame, not only regarding the technical specification of the relation $[\![\phi]\!]$, but also regarding the difference between what a statement means for the sender and what it means for the receiver, the latter having other (usually far less) information about the utterance situation than the former.

Notes.

* This paper is extracts and refinements of my (unpublished) cand.scient. thesis (Sem 1983). I wish to thank my advisor, Professor Jens Erik Fenstad for valuable advice and encouragement both in the work of my thesis and in the work of this paper. I wish also to thank both him and Jan Tore Lønning for comments on earlier versions of this paper.

1) Apart from the ones we want to rule out by various semantic filters.

2) That is, he opens for the possibility of more than one gap, while giving a single gap constraint in Ch. V, 2.3.

3) Remember that the intension is a function that yields a denotation for each world, and that the intension can be constructed from the denotations for each world. To put the denotation rather than the intension as the first term in the interpretations is just a matter of notational choice. The calculation of the denotation of some types of compound expressions, like e.g. in semantic rule 2, may require the intension and not only the denotation of one or both of the constituents.

4) See notation 1 for a description of NP-int and Prop.

5) For alternative b), $[\underline{\text{who}}]_{NP}$, $[\underline{\text{what}}]_{NP}$, $[\underline{\text{which}}]_{NP}$ will get their correct denotations by the standard rule $\|[\alpha]_X\|_{\alpha,w} = \|\alpha\|_{\sigma,w}$. An exception rule is therefore not required in this case.

6) The examples are there calculated in a fragment with the syntax of Fragment 6 in Cooper (1983), but with semantics modified similarly to Fragment 6^{II} and 6^{III}. Some of the

words in example 4 are exchanged to get a more natural sentence. The structure and the calculation process remain the same except for the substitution of words.

7) The numbers used in the storage mechanisms are arbitrarily chosen between the odd or even numbers, and will not be further commented.

8) This sentence is quite heavy even in Norwegian. It may therefore be useful to give a context for it:
"Mari vet lite om hester, men tror at enhver hest kaster sin eier av, iallfall hvis eieren er en mann. Hittil har hun bare fått bekreftet sin tro, idet enhver hest Mari kjenner en mann som eier, kaster ham av."
The first som may or may not be omitted.

9) Names of the syntax categories will in the following be as in Kamp (1981) and not as in Cooper (1983).

10) Only one restrictive relative is permitted for each CN in L_1. The use of non-atomic qualifiers is therefore not necessary. We could have followed the Kamp-technique of splitting the basic common noun from the relative clause when reducing the expression in which the common noun with the relative clause is a constituent of the main term. However, the use of non-atomic qualifiers also serves other purposes:

i) The rules for construction of DRSs can be made simpler and more systematic, with smaller and more general reduction steps.

ii) Every DR will have a main occurrence which is the ancestor to all the other occurrences in the DR.

iii) It opens for a uniform treatment of more complicated CN phrases in a larger fragment.

iv) Scope manipulations will be easier with one occurrence representing the entire CN-part of terms on the form $\left\{\begin{matrix}\text{a(n)}\\ \text{every}\end{matrix}\right\}\beta$, where $\beta \in CN$.

v) The DRSs are intended also to represent meaning as what the listener "grasps" at hearing the text. The entire CN plays a role in forming the concept in our minds. This is better represented by using an occurrence of the entire CN in the DRSs than by splitting up the CN at once.

(See definitions 22 and 25 for exact definition of DR and DRS.)

11) At deixis, v must be chosen from some other given set.

12) The function c indicates one way of implementing deictic use of pronomina; as connections (set by the language user by means of the context or by non-linguistic means) between some particular discourse referents and individuals in U_M. A verifying function for a DRS for a discourse, relative to some preset connections c, must then be compatible with c.

13) An extra index will be necessary to separate different occurrences of the same expression. We omit it here for readability reasons, as each expression will only occur once in one examples.

14) See under example 2 of section 3.3 for a description of additions.

References

Barwise, Jon and John Perry: 1980 "The Situation Underground" in Barwise, Jon and Ivan A. Sag (eds): "Stanford Working Papers in Semantics", Volume 1, Stanford.

Barwise, Jon and John Perry: 1983 "Situations and Attitudes". Bradford Books, M.I.T press.

Cooper, Robin: 1979 "The Interpretation of Pronouns" in Heny, Frank and Helmut S. Schnelle (eds): "Syntax and Semantics. Selections from the Third Groningen Round Table", Volume 10, Academic Press, New York.

Cooper, Robin: 1983 "Quantification and Syntactic Theory". D. Reidel Publishing Company, Doordrecht.

Engdahl, Elisabet: 1980 "The Syntax and Semantics of Questions in Swedish" unpublished Ph.D dissertation, University of Massachussets, Amherst.

Geach, Peter T.: 1962 "Reference and Generality", Itacha.

Hellan, Lars: 1980 "On Semantic Scope" in Heny, Frank "Ambiguities in Intensional Contexts". D. Reidel Publishing Company, Doordrecht.

Kamp, Hans: 1981 "A Theory of Truth and Semantic Representation" in Groenendijk, J.A.G., T.M.V. Janssen and M.B.J. Stokhof (eds): "Formal Methods in the Study of Language", Matematisch Centrum, Amsterdam.

Rodman, R.: 1976 "Scope Phenomena, "Movement Transformations" and Relative clauses" in B. Partee (ed), "Montague Grammar", Academic Press, New York.

Sem, Helle Frisak: 1983 "Kvantifikasjon, Syntaks og Semantikk", unpublished cand.scient-thesis, Dept. of Mathematics, University of Oslo.