Applying an Exhaustification Operator in Update Semantics

Henk Zeevat
University of Amsterdam

Contents

1 Introduction

This paper presents an exhaustification operator suitable for update semantics and discusses a series of applications of this operator. The exhaustification operator takes an open formula and assigns (if this is possible) values to the free variables such that the formula is true as a result and entails any sentence that can be obtained from the formula by assigning other values to the variables that make the sentence true.

This operator is employed to provide an (update) semantics for questions. The Wh-elements of the question correspond to discourse markers and the discourse markers are exhaustified with respect to the question. (Positive) answers to the question present extra constraints on the same discourse markers.

The theory of questions that results is then used to formalize the theory of topic and focus that equates the topic with a question and the focus with its answer. As we use a standard DRT-like representation of the complete sentence to represent the focus, the semantic effect of the topic-focus division is that certain discourse markers in the sentence receive an interpretation that is exhaustive with respect to the topic. The same assumption also makes it possible for the theory to allow multiple topics.

The topic-focus theory is applied to obtain certain scalar implicatures and to explain the Evans-effects. The indeterminacy of the topic focus division is exploited to explain the "cancellation" of the implicatures and the definiteness effects.

It is moreover used to salvage a discarded theory of plurals in DRT, which analyses those plurals that are not monotone decreasing as definites or indefinites (with an internal structure), deriving those properties which are in conflict with this assumption from the topic-focus distinction.

Dealing with the applications in the way sketched here involves me in a number of non-orthodox positions. First of all, I will have to take some distance from the generalised quantifier approach to plural determiners, or at least with its direct application to natural language semantics. The treatment I adopt is close to early Discourse Representation Theory[1]. The problems that face such an approach will be solved by assuming a topic-focus mechanism which exploits the exhaustification operator[2]. Second, the application to topic-focus places me in the camp of those who try to assimilate the meaning of topic and focus to a question (with exhaustive interpretation) representing the topic and an answer to that question representing the focus[3]. The treatment of scalar implicatures, or at least a part of them, no longer makes use of scales and involves some special assumptions about the semantic representation of various NL-expressions. In contrast, the theory of questions developed in section 4 tries to stay as close as possible to the classical view of Groenendijk & Stokhof (1984), although it completely changes the framework in which that theory is couched. Perhaps controversial is

[1] In contrast, Kamp and Reyle (1993) is orthodox in its treatment of quantifiers.

[2] The effects of this operator are standardly assumed in the characterisations of the quantifiers in the theory of generalised quantifiers. A disadvantage of that approach is a multiplication of readings, in the absence of a mechanism that chooses between them. The two level approach chosen here, a basic semantic representation strengthened by an essentially pragmatic mechanism keeps down the number of readings assumed by e.g. Scha and moreover provides a disambiguation mechanism.

[3] Van Kuppevelt (1991) is a systematic defence of this view.

the commitment to exhaustiveness that derives from Groenendijk & Stokhof.

Why use update semantics? It will be clear from my discussion that it is quite possible to define exhaustification operators outside the context of an update semantics. Elsewhere (Zeevat & Scha (1992)), we have defended the view that update semantics is particularly suited for developing pragmatics and semantics within a single theory. The successful treatment of presuppositions in update semantics goes back to Karttunen (1974) with important additions by Heim (1981). Certain pragmatic implicatures of assertions have been shown by Stalnaker (1978) to be directly expressible as conditions on updates. Here, we attempt to do the same for certain implicatures arising from quantity and relevance. There are also two advantages in the treatment of questions. In the first place, it is unnecessary to make a distinction between semantic and pragmatic answers to questions: the information state that is updated always forms a suitable context. (With the semantic notion recoverable as updates of the empty information state.) A second advantage is that updating a single information state provides us with a simple semantics for dialogue, if we take the information state to be the common ground between the conversation partners.

Information states are here conceived as in Stalnaker (1978) to be a representation of the apparent common ground between speaker and hearer(s): that body of information which partners have purported to believe in the conversation. Some have proposed to take the hearer's information or the hearer's picture of the common ground. It would seem to me that this makes it hard to deal with non-monologic phenomena such as e.g. conversations, as we would be obliged to jump from one information state to another one that cannot be derived from the first one, every time the speaker changes. As it is my purpose to deal with questions and answers, I see no reason to depart from Stalnaker. Note, however, that the development of partners' information states can in principle be easily described in terms of changes to the common ground, at least for honest partners. (Intersection should suffice in that case.)

2 Exhaustification

What is the exhaustive interpretation of a variable in a formula? Intuitively, it is a value for a variable such that taking it to be the value rather than something else makes the formula true and makes it entail all the true formulas that can be obtained by assigning another value to the variable in the formula. It is the strongest interpretation that the open formula allows. Of course, there need not exist an exhaustive interpretation for a formula. This is indeed a common situation. Suppose five boys are asleep. It is then impossible to have an exhaustive reading for sentences like (1a/b) .

(1) a. One boy sleeps.
$x \wedge boy(x) \wedge one(x) \wedge sleep(x)$
b. Less than three boys sleep.
$x \wedge boy(x) \wedge \#x < 3 \wedge sleep(x)$

None of the values we can find for these sentences is exhaustive: if x denotes one sleeping boy, x can also denote another sleeping boy without there being a logical connection between the statement about the one boy and the other. The same holds if x denotes sets with cardinality less than three: there are variants for the denotation of x that are logically unconnected.

Exhaustification is thereby a combination of the statement that exhaustive readings are possible together with the assignment of the exhaustive value to the free variable[4] When exhaustification is possible, it gives minimal or maximal elements with respect to some order, e.g. the inclusion order on sets or the natural order on natural numbers.

The concept of interpreted open formulas entailing one another is not standard and the following remarks are intended to make it precise. The problem is that we must define a notion of entailment over interpreted formulas. Interpretation normally involves a single model and does not make sense over arbitrary classes of models (e.g. the value may not be available in another model, or it may play a completely different role). Entailment however essentially involves a quantification over models.

To give content to interpretation, we will limit our models to a class K which contains the expansions to a language L of a given model M_0 for a language $L_0 \subseteq L$. The given model fixes the domain and some privileged relations. For the examples considered in this paper, it suffices to take the basic model M_0 to be the powerset of some given non-empty set (without $\emptyset$) together with the set of natural numbers (without 0), with the privileged relations *inclusion* between the sets, *smaller than* between numbers and *cardinality* relating sets and numbers. Object variables will range over sets of objects, number variables over numbers. A reasonable extension would be the inclusion of quantities of stuff and reals among the domain entities with the basic relations between the two. Part-whole relationships and measurement are other obvious candidates.

In addition, we assume a set MP of postulates about the non-privileged relations. With entailment, we will mean K-entailment from now on.

K can be written as $K = \{M : M \models MP \text{ and } M|L_0 = M_0\}$. $|$ is the restriction operator.

Let φ be a formula with some free variable x and K a class of models M as described above. The exhaustification of φ in K with respect to the variable x is that object u in the domain U_M of M such that (2).

(1) $M \models \varphi < u >$ and
(2) $\forall v \in U_M \forall M_1 \in K(M \models \varphi < v > \text{ and } M_1 \models \varphi < u >$
$\Rightarrow M_1 \models \varphi < v >)$.

Example 1.

Let K be as described above. Let MP be given by:

$Px \wedge y \subseteq x \rightarrow Py$

(gloss: If John has sheep x then John has sheep y for $y \subset x$.)

Let φ be Px

Then an exhaustive value for x in the model M is the set of all P in M. (John's sheep.)

[4] Classically, we would have to say that exhaustification binds the variable. That the variable is available as a name for the exhaustive value outside the scope of the operator is a non-classical dynamic effect. Unlike systems like DPL we do not assume that the variable only functions in this way to the right of the scope of the operator.

Example 2.

The postulates are given by:

$Pn \wedge m > n \to Pm$

(gloss: If John runs the mile in n minutes then John runs it in m minutes if $m > n$)

Let φ be Pn

The exhaustive value is the smallest number m such that Pm in M. (gloss: John's time for the mile.)

Example 3.

$Pn \wedge n > m > 0 \to Pm$

(gloss: If Bill has four chairs then Bill has three chairs.)

The exhaustive value is the largest number m such that Pm. (gloss: The number of Bill's chairs.)

2.1 Update Semantics

Update semantics is a general name for any theory of language that explains the semantic properties of its expressions in terms of the information change that they bring about on information states.

There is room for a general theory of update semantics: one that tries to abstract from any assumptions about the nature of the information states and the changes that they allow. (See e.g. Veltman(to appear)). Notions of logical consequence typically belong to this level. A natural notion is to define $\varphi_1, \ldots, \varphi_n \models \psi$ as $\forall\sigma\sigma[\varphi_1]\ldots[\varphi_n][\psi] = \sigma[\varphi_1]\ldots[\varphi_n]$ (for other notions, see Veltman).

Another distinction that can be made is the one between monotonic systems, allowing only updates, and non-monotonic systems that allow the information state to decrease. The latter kind are important for theories of belief revision and have also been used for giving an update semantics for DPL (Groenendijk (1993)).

The changes that can be considered depend on the information states that are allowed. The structure of the information states in turn limits the possible operations on these states.

Two main options are possible. We let the information states grow as they acquire new information. This is the constructive approach. A classical model would be to take complete theories in some logic. Information growth would be the addition of a new sentence to the theory and closing off. (Another model of this approach is the DRS construction algorithm: the natural language defines the updates, the information states are the DRSs.) The other road starts from taking a set of information carriers as given and proceeds by eliminating carriers. This is eliminative update logic. A third approach is a combination of elimination and construction. This has been considered by Dekker(1993), in the footsteps of Heim(1983).

Our approach here is purely eliminative. In an eliminative update semantics, we increase the information in an information state by eliminating information carriers: those in which the new information does not hold. Both the appearance of new discourse markers and the appearance of new facts will be modelled by elimination.

Information carriers for a language $L =< P, F, C >$ (with P a set of relations, F a set of function symbols, and C a set of constant symbols) will be models for languages $L' =< P, F, D >$ with $D \subseteq C$. C is made up of two sorts: sets of objects and natural numbers. We make no distinction between constants and variables. Among the ranges of the individual terms we do not include the empty set and the number zero. (This reflects natural language: there is no group of zero elephants.)

The language introduced is a version of the DRT-formalism and is close to Vermeulen (to appear).

(2) Terms:
a. basic terms for numbers and sets.
b. $ft_1, \ldots, t_n$ is a term iff $t_1, \ldots, t_n$ are terms, f is a function symbol and $t_1, \ldots, t_n$ match the signature of f.

Formulas are defined in (3).

(3) Formulas:
a. basic terms are formulas
b. $t_1 = t_2$ is a formula iff t_1 and t_2 are terms of the same sort.
c.$Pt_1, \ldots, t_n$ is a formula iff $t_1, \ldots, t_n$ are terms, P is a predicate symbol and $t_1, \ldots, t_n$ match the signature of P.
d. $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \rightarrow \psi$ are formulas iff φ and ψ are.

The function of the terms as formulas is similar to the discourse markers of Kamp (1981). Below, we define discourse markers by a recursion over the formulas, though we will not use this definition in the semantics.

1. $DM(x) = \{x\}$

2. $DM(\varphi) = \emptyset$ if φ is atomic or $\varphi = \neg\psi$ or $\varphi = \psi \rightarrow \chi$

3. $DM(\varphi \wedge \psi) = DM(\varphi) \cup DM(\psi)$

Information states are sets of information carriers. We can now define the update $\sigma[\varphi]$ of an information state σ by a formula φ in the following way.

1. $\sigma[x] = \{i \in \sigma : ix \text{ defined}\}$

2. $\sigma[Pt_1,\ldots,t_n] = \{i \in \sigma : \neg < it_1,\ldots it_n > \notin iP\}$

3. $\sigma[t_1 = t_2] = \{i \in \sigma : \neg\exists u \exists v(it_1 = u, it_2 = v \wedge u \neq v)\}$

4. $\sigma[\varphi \wedge \psi] = \sigma[\varphi][\psi]$

5. $\sigma[\neg\varphi] = \mathbf{neg}(\sigma[\varphi], \sigma)$

6. $\sigma[\varphi \rightarrow \psi] = \sigma[\neg(\varphi \wedge \neg\psi)]$

For the negation we need the definition (4),

(4) $\mathbf{neg}(\sigma, \tau) = \tau \setminus \sigma^{dm(\sigma,\tau)}$

which in turn requires (5) and (8) .

(5) $\sigma^X = \{i : \exists j \in \sigma \; i =_X j\}$

(5) makes use of (6).

(6) $i =_X j$ iff $\forall a$ ($a \notin X \Rightarrow$ ($ia = ja$ or ia and ja are both undefined.))

(7) $dm(\sigma, \tau) = \{c \in C : \sigma \models c \wedge \tau \not\models c\}$

The first three clauses of the definition of updates are set up in such a way that there is a distinction between an atomic formula (with *free* terms) eliminating information carriers and updating the *conjunction* of the free terms with the atomic formula: only in the latter case it is guaranteed that each of the variables will be defined throughout the information state. The atomic formulas only eliminate those carriers that overtly contradict them. This allows a notion of the discourse markers of an information state: the terms that are everywhere defined in that information state and, thereby, of the negation of an information state σ_1 with respect to another information state σ: the subtraction of the closure of the first information state σ_1 with respect to those of its discourse markers that are not markers of σ from σ. This semantic definition allows the development of the semantics as a proper algebra over information states.

Our treatment of discourse markers may cause some worries. An update with a term c makes the term into a complete object, but does not add interesting claims about it, other than that it is a possible object. On arbitrary σ, we can add $square(c)$, then $\neg square(c)$ without causing σ to become the inconsistent information state. Only when we add c as a final update, will inconsistency be reached. Natural language names are of course quite different, as their use presupposes their existence. Here, the update with c is the presupposed existence, the other occurrences are non-presupposing.

The fact that the update c is so uninteresting makes the update $\neg c$ necessarily inconsistent, even if another occurrence of c is accessible in the sense of Kamp. (In that case, the local state contains c, whereas $negc$ denies it.) This makes $\neg\neg c$ a tautology.

Information states can be in three minds about a discourse marker: it can contain it, i.e. $\sigma \models c$, it can reject it ($\sigma[c] = \emptyset$) and it can accept it as possible ($\emptyset \subset \sigma[c] \subset \sigma$).

Kamp's accessibility can be faithfully expressed as $\sigma \models c$. This should not be confused with the property of being an old discourse marker which is much weaker. This notion cannot be defined along these lines, since one can be old by being a non-accessible discourse marker on another path or by being constrained without being a discourse marker. The safest option for a natural language interpreter is therefore to stick to the rule of using new terms unless an old term is explicitly required.

2.2 Exhaustive Updates

Exhaustive updates are updates with a formula whose discourse markers in the update are exhaustified with respect to the formula. To achieve this we think of the marker as just another proper name. We eliminate the information carriers in which the formula does not hold and those in which the carrier does not give an exhaustive value to the variable. The first elimination is conventional, for the second we need to import exhaustification.

Information carriers are models. Quantification is dealt with by considering other information carriers in the information state which are almost exactly the same except for the value assigned to certain variables. This relation is standard. For exhaustiveness, we will introduce a relation which is similar to variation with respect to a set of variables, but which allows everything else to vary instead. The two information carriers must have the same domain, the same interpretation of all individual constants and have the same extension for a set of privileged relations. (the analogue of the L_0 relations considered above). They can vary with respect to the interpretation of non-privileged predicates and functions. We call this relation object-identity: the two information carriers agree with respect to the objects and their basic ontological relationships. They may disagree about everything else.

As before, we expect the information state (i.e. every carrier in it) to satisfy a set of meaning postulates MP. The set K we had before can now be equated with those parts of the information state that have the form $\{j \in \sigma : j \text{ is object-identical with } i\}$.

We obtain by this relation for each of the information-carriers i in the information state σ a set of models K_i that can support the entailment relation needed for judging whether i assigns an exhaustive value to the variable. Because σ contains the conceptual information in MP, so does each of the K_i.

The following diagram shows the demand of exhaustiveness on the variable x with respect to the information state $\sigma[\varphi]$.

$$
\begin{tikzcd}
i \arrow[r, "\texttt{object-identical}"] \arrow[d, "\texttt{x-variant}"'] & j \arrow[d, "\texttt{x-variant}"] \\
k \arrow[r, "\texttt{object-identical}"'] & l
\end{tikzcd}
$$

i : $sleepers = \{j, b, m\}$
$x = \{j, b\}$

j : $sleepers = \{j, b\}$
$x = \{j, b\}$ (obj.id. i)

k : $sleepers = \{j, b, m\}$ (x-var. i)
$x = \{j, b, m\}$

l : $sleepers = \{j, b\}$ (x-var. j)
$x = \{j, b, m\}$ (obj.id. k)

So, l does not satisfy $sleep(x)$

Exhaustivity Diagram

i is exhaustive for x iff whenever i has an x-variant i_x and an object-identical variant i_o in $\sigma[\varphi]$ then i_o has an x-variant $i_{ox} \in \sigma[\varphi]$ that is object-identical to i_x.

To see that this is correct, consider what we mean by exhaustive values. φ according to i should entail all of the φ-meanings in x-variants i_x of i. When would it not do so? Intuitively, if the carrier i has an x-variant i_x, but there is a world w, in which φ is the same as in i, but which lacks the corresponding x-variant.

Object-identity here guarantees two things: it guarantees that the same value is given to x and that x is not just formally the same: it plays the same role in the ontology of the other world. So w must be object-identical to i. The corresponding x-variant must similarly be object-identical to i_x. A counterexample to i being exhaustive for x and φ with respect to some σ is therefore an x-variant i_x and an object-identical i_o, both in σ, which lack an element i_{ox} that is object-identical to i_x and x-variant to i_o.

To go back to our earlier example:

We need the meaning postulate (8):

(8) $$Pz \wedge y \subseteq z \rightarrow Py$$

i.e. we assume that (9)

(9) $$\sigma[Pz \wedge y \subseteq z \rightarrow Py] = \sigma$$

and $\varphi = Px$. x must be new to the information state, i.e. $\not\sigma \models x$.

Suppose i assigns $pow(A)$ to P, and $B \subset A$ to x. Take i_x such that i_x assigns A to x. $i_x \in \sigma[Px]$ since x is new and by the assumption.

Consider i_o such that i_o assigns $pow(B)$ to P. $i_x \in \sigma[Px]$ as x is new.

Then there is no i_{ox} such that $i_{ox} \in \sigma[Px]$, i_{ox} is an x-variant of i_o and i_{ox} is object-identical to i_x.

By object-identity: i_{ox} assigns A to x.

By x-variance: i_{ox} assigns $pow(B)$ to P.

But then $i_{ox} \notin \sigma[Px]$

So x must be the maximum if i is exhaustive.

The other examples follow by the same reasoning.

An exhaustification operator q can be defined with the above semantics. The operator will take the discourse referents of a formula and deliver an exhaustive interpretation for all of them if such a reading exists. By the semantic definition of discourse markers, the discourse markers of the argument of the operator are the same as those of the result.

(10) $\sigma[q(\varphi)] = \{i \in \sigma[\varphi] : \forall j, k \in \sigma[\varphi]\ (j =_{dm(\sigma[\varphi],\sigma)} i \wedge k$ is object-identical to $i\ \exists l \in \sigma[\varphi]\ (l =_{dm(\sigma[\varphi],\sigma)} k \wedge l$ is object-identical to $j))\}$

3 Questions

The aim of this section will be to consider the combination of exhaustivity and update semantics as a tool for reformulating the theory of Groenendijk en Stokhof on questions. Only direct questions will be be treated, indirect questions will be only speculatively considered.

In the theory of Groenendijk & Stokhof (1984) (GS), standard answers are characterised by three properties.

(11) They are true
They are rigid
They are exhaustive

Two semantic characterisations of questions follow: questions are concepts of their true answers (i.e. the functions that assign to a possible world the true propositional answer to the question) and partitions over possible worlds, induced by the relation $\{< i, j >$:the propositional answer to the question in index i contains $j\}$).

The informational perspective and the employment of update semantics precludes taking over the Montague grammar formulation of these concepts. In update semantics, we do not have the notion of truth (though it can be added), we have only expressions of type t and e, and it is only by information change that we can define meaning. Within our monotonic update semantics, it holds that if questions mean anything at all, we have to characterise this meaning in terms of the new information they bring to the information state.

The theory of questions I am proposing is simple: it applies the exhaustification operator to the formula representing the question that contains the question's Wh-elements as its discourse markers. A question update is an auxiliary update with the formula so obtained. The answer will determine how to proceed with the auxiliary information state.

An auxiliary update leaves the original information state intact and constructs a second information state. (We have seen an example in the treatment of negation, in which we update with the negated formula, and determine the update of the negated sentence in terms of the information state so obtained.)

There are three ways in which we can deal with the auxiliary state: we can negate it with respect to the original information state, in case the answer is negative (e.g *no one, no, no animals*), we can replace the original state by the auxiliary state updated by the answer if it is positive and finally we can forget it[5], in case the interlocutor declines to answer it (e.g. *I don't know.*). In the following two examples, we illustrate these three cases.

(12) Did John come to the party?
a. Yes.
b. No.
c. I do not know.

(13) Who came to the party?
a. John's friends.
b. Nobody.
c. I do not know.

A positive answer can be reconstructed as a sentence (by some mechanism for ellipsis resolution), or we can assume a mechanism for interpreting sentence fragments. In both cases, we need only one thing: that the variables for the referents of the expressions in the answer corresponding to the Wh-expressions in the sentence are the same (by unification) or are stated to be identical. (In the sketch of a grammatical treatment in the next section, I assume sentence fragments and unification: the answer *John* to the question *Who sleeps?* is presented as the unification of x and $John$, the answer *A boy* as the statement $boy(x)$. A positive answer adds its contents to the auxiliary information state, which replaces the original information state. In the following table (14), we give the sequence of events for a question that is asked and then positively answered, negatively answered or declined.

(14) **Positive answers** 1. σ
2. $\sigma.\sigma$
3. $\sigma[question].\sigma$
4. $\sigma[question][answer].\sigma$
5. $\sigma[question][answer]$

In step (1), the conversation partners have a common ground σ. The fact that a question is asked puts (2) a copy of the common ground to the foreground, keeping the original information state in the background (the dot is the stack forming operation). The foreground is now updated (3) with the question and with the positive answer (4). Acceptation of the positive answer makes the foreground into the new common ground (5).

(15) **Negative answers.** 1. σ
2. $\sigma.\sigma$
3. $\sigma[question].\sigma$
4. $neg\ \sigma[question].\sigma$

[5] The ignorance of the interlocutor will be part of the common ground, which makes it strictly speaking wrong to just obliterate the question update. A proper treatment of the common ground assumptions about the speaker and the hearer is however beyond the scope of this paper.

In (15) , steps (1) to (3) are the same. In (4), the new common ground becomes the negation of the foreground, with respect to the background[6].

(16) **Declining to answer.** 1. σ
2 $\sigma.\sigma$
3. $\sigma[question].\sigma$
4. σ

In (16) , finally step (4) reverts to the information state of (1).

3.1 Adapting Questions

The choice between a positive answer and declining to answer is not always a sharp one: we can know the answer partially. Though there are answering strategies that provide for proper answering (*John and others* or *John and maybe others* can be proper answers to a Wh-question, reflecting the speaker's ignorance), another strategy is to tacitly change the question. In case the question was *Who is asleep?* and we only know that John sleeps but fail to know anything about the others, we may answer the weaker question *Is John asleep?*. In this case, there are some means of expression that help to indicate that we are answering a different question. Twiddly intonation on *John* is one of these devices, but also more elaborate locutions may be chosen (e.g. *John is asleep, but I do not know about the others*).

Overanswering is the phenomenon that the answer gives more information than the question was -strictly speaking- asking for. This again is a question of tacitly changing the question, sometimes combined with an answer to the original question.

(17) Did any stock rise yesterday?
Yes, Alcatel and Telefonos Mexicanos.

In (17) the answer to the *yes-no*-question is followed by an answer to the Wh-question *Which stock rose yesterday?*, a question that was not explicitly asked, but one which the interpreter obviously thought would be the next one the speaker would ask. That this question must be reconstructed in a grammatical treatment follows within our treatment from the need to obtain the exhaustivity effects.

Questions come in an obvious order. The weakest ones are the yes-no-questions. Stronger questions can be obtained by replacing standard NPs by Wh-elements and by replacing more restricted Wh-phrases by less restricted ones. Underanswering can be seen as answering a question derived from the original one by filling in a more concrete Wh-element for one of the Wh-elements in the question or by replacing it by an non-Wh-element altogether. Overanswering can analogously be understood as adding Wh-elements to the question or as making the Wh-elements less specific. The ordering strongly resembles the unification

[6] The treatment must allow for the addition of a denied answer. The answer *not more than three books* to the question *What did John read?* can be modelled by updating with *more than three books*, before negating. An alternative treatment treats negative answers on a par with positive answers, by unifying the internal variable of the negative quantifier with the Wh-variable. Such a treatment however does not work for *no*, unless that is considered to be an anaphor that takes the question as its antecedent and denies it.

semilattice of the elements subsuming a given ground term. The semi-lattice can be grounded in semantics as well: knowing the answer to a stronger question always entails knowing the answer to the weaker question, under the assumption that the knowledge subject knows that the stronger question is stronger than the weaker one.

Of course, a speaker does not change the question without good cause. Going to a weaker question is allowed if the speaker cannot reply to the stronger question or if the speaker realises that her partner is really looking for an answer to weaker question. Answering a stronger question results from the realisation of the speaker that she can do so and that the stronger question is the one her conversation partner is really after. Recognising the speaker's intention is as important in understanding a question as it is in understanding an assertion.

An application of shifting questions are non-exhaustive answers: they can be understood as answers to the weaker question. In terms of our theory, the topic of a non-exhaustive answer is a weakening of the explicit question. The exchange (18):

(18) Where can I get some coffee?
One floor down, second door left.

does not entail that coffee cannot be had elsewhere (though sometimes it does). We can explain this by assuming an implicit condition *around here* inside the *where* or a more specific meaning of the word *where*: *which is the closest place where* to obtain a weaker Wh-question or a shift to the yes-no-question: Can I get some coffee one floor down,second door left?

This is however not the only way in which we could deal with non-exhaustive questions. We could have a special speech act of non-exhaustive answering. Rather than an identity between discourse markers, we would have a strict inclusion between the markers of the answer and the marker of the Wh-element.

A non-exhaustive exchange would then be treated as (19):

(19) Who is at home?
John.
$q(x \wedge \; x \textit{ is at home}) \wedge j \subset x$

For the time being, I prefer the first approach, which is closer to the way we treat topic and focus in general. The first question would be forgotten and replaced by the weaker question *Is John at home?* which is then positively answered.

3.2 Wh-elements

A logical representation of questions needs to have a question operator and a way for marking Wh-elements.

My proposal would be to let Wh-phrases be represented as indefinites: a new discourse marker and possibly a new condition. That they are Wh-markers is then indicated by the fact that they are bound by the q-operator. The meaning of the q-operator is to give an exhaustive interpretation to the discourse markers that it binds. Within a DRT-context, the main syntactic problem is then to protect possible indefinites occurring in the syntactic scope of the Wh-phrase from being bound as well. A simple proposal is to add an operator, closing off the syntactic scope of the Wh-phrase, to the semantics of the Wh-phrase. Operators with this property are readily available: the double negation or $true \rightarrow \varphi$. The q-operator itself is unsuitable as it does not preserve the non-dynamic meaning.

A disadvantage of this procedure is that it makes the indefinites unavailable for future anaphora. This is incorrect as such anaphora does occur, when the question is answered in a positive way. This is a strong argument for following GS in assuming full propositional answers using ellipsis resolution for constituent answers and unifying the discourse markers deriving from the Wh-phrases the relevant discourse markers in the answer. The semantic representation of the answer could then be standard, i.e. omitting both the q-operators and the double negation(s).

This is not the place to go into full detail about the ellipsis mechanism but we can imagine a process in which the constituents in the answer replace the corresponding Wh-phrases and the result is interpreted in the normal way and which as a side-effect unifies the markers of the Wh-phrases with the markers of the constituents that reply to them. The matching process involved in the definition of correspondence is equally important in the interpretation of complete sentential answers: also in this case, the Wh-phrase markers and the corresponding markers must be unified[7].

The question can remain as proposed here, with the indefinites unavailable, but as they are repeated in the answer, they become available after a positive answer. The assumption of multiple topics in the next section also points in the direction of propositional answers: if we assume multiple topics for a single sentence, we cannot have constituent answers as the basic case.

This gives us the scheme (20) for the semantics of a Wh-phrase.

(20) $x \wedge A \wedge \neg\neg(B)$

Here A is a restriction possibly incorporated in the Wh-phrase and B is the scope of the Wh-phrase.

In case we are dealing with a scope that does not contain a Wh-phrase, this is fine. The combination should get a plus-value for the feature *wh*. For type s, there is then a syntactically empty operation that adds a q-operator to the semantics of a wh+-sentence. The results of this operation can form the scope of both Wh-phrases and certain other NPs.

If a Wh-phrase applies to wh+-expression to which a q has been added, it must not apply a double negation. In order to make the internal markers of the wh+-phrase available for

[7]The operation needed here is the unification of two partial sentential structures: one which derives from the question by abstracting away all the material deriving from the Wh-phrases with the exception of sortal restrictions on their marker, the other complete when a full answer is given, but a partially determined sentence in which the answer fragments occur (in the order in which they occur). This is different from the process described by Prüst et al.(1994) as the traces of the question semantics would interfere (an elliptic answer would become a question) and also from the approach in Gardent(1991) as the process is symmetric: material must flow from the answer to the (abstracted) question and may flow from the question to the answer. Unlike VP-ellipsis, elliptical answers never override values in the antecedent.

unification with markers in the answer, it is necessary that these markers are available and not made unaccessible by a double negation. Next to this process, Wh-phrases can also direct apply to expressions containing another Wh-phrase. In this case, the Wh-phrases start behaving as if together they were forming a single large Wh-phrase. This leads to a second scheme, which applies to [wh+]-arguments, with or without a q.

(21) $x \wedge A \wedge B[wh+]$

Yes-no-questions can be incorporated into this scheme, but need not, although we want to avoid that exhaustivity applies directly to their discourse referents. We can let quantifiers[8] have wider scopes than Wh-phrases. In this way we can obtain the two readings of (22). Some examples:

(22) Which woman does every man like most?
$x \wedge x = MANdist(x, q(y \wedge woman(y) \wedge like\_most(x, y)))$
$q(y \wedge woman(y) \wedge \neg\neg(x \wedge x = MAN \wedge dist(x, like\_most(x, y))))$

(23) Who sleeps?
$q(x \wedge \neg\neg sleep(x))$

(24) Who meets a professor?
$q(x \wedge \neg\neg(y \wedge professor(y) \wedge meet(x, y)))$
Who meets which professor? (embedding)
$q(x \wedge q(y \wedge professor(y) \wedge \neg\neg meet(x, y)))$

(25) Who meets which professor?(lumping)
$q(x \wedge y \wedge professor(y) \wedge \neg\neg meet(x, y))$

3.3 Multiple Exhaustification

We predict that (26) has two readings.

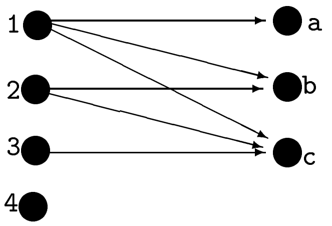
(26) Who loves who?

In the first case, we obtain a representation (27).

(27) $q(x \wedge q(y \wedge love(x, y))$

Let us consider the double operation on a small domain. The internal q-operator gives us the interpretion under step 1. Possible interpretations like $\{1\} : \{a, b\}$ or $\{2\} : \{b\}$ are eliminated.

[8]That (non-plural) indefinites always have a narrower scope than Wh-phrases needs an explanation. Perhaps this must be found in the nature of such indefinites (indefinites like to be bound) or in the unsuitability of asking about things the speaker knows but the hearer does not know yet.



step 1	step 2
$\{1\} : \{a, b, c\}$	$\{1\} : \{a, b, c\}$
$\{2\} : \{b, c\}$	
$\{3\} : \{c\}$	
$\{1, 2\} : \{b, c\}$	$\{1, 2\} : \{b, c\}$
$\{1, 3\} : \{c\}$	
$\{2, 3\} : \{c\}$	
$\{1, 2, 3\} : \{c\}$	$\{1, 2, 3\} : \{c\}$

Fig. 2 Love in carrier i

The doubly exhaustive reading makes for a compact representation of the positive part of the relation as a relation between sets.

The other reading of the question is (28):

(28) $q(x \wedge y \wedge love(x, y))$

The relational MP gives us the assignment : $x = \{1, 2, 3\}$ and $y = \{a, b, c\}$. If we interpret *love* under this MP, it will be the only exhaustive assignment. This reading requires an atomic basic representation: if some of the basic relata are sets, we are in trouble with the members of these sets: they are not related. What fails is the implication (29):

(29) $< x, y > \in R \wedge z \subseteq x \wedge v \subseteq y \Rightarrow < z, v > \in R$

There are two ways out of this problem. The first would be to follow several recent proposals and allow groups of groups, i.e. let basic sets be elements of the values of the exhaustive reading. The other would be to let exhaustivisation apply to a flattened predicate *love** which transforms the relation in a relation among singletons. The first option is preferable and can be argued for independently (see, e.g. Scha & Stallard(1985)).

3.4 Comparison

When does an information state answer a question? (In terms of such a definition we can define proper answers: $\sigma[Q][A]$ must answer Q.) This is easy in Groenendijk & Stokhof (1984): an information state σ answers a question if the question partitions it into the partition $\{\sigma\}$. Here we use a similar idea, but things are one degree more complex.

When is a question answered on an information state σ? The question update eliminates the carriers that lack an appropriate exhaustive value for the Wh-elements. What they then do for the other worlds is select the appropriate *dm*-variant: this is the answer according to a carrier i for which an answer exists. So the question update divides the carriers into two classes: one in which there is no answer, another in which the carrier gets mapped to its exhaustive *dm*-variant. That an information state answers a question means first of all that the first class is empty: every carrier should determine an answer. Second, they should all determine the same answer.

This can be expressed as demands on $\sigma[q(\varphi)]$. If σ answers $q(\varphi)$, the only effect of the update is that the allowed values for the discourse markers of $q(\varphi)$ are restricted. This implies that every $i \in \sigma$ should have a *dm*-variant in $\sigma[q(\varphi)]$. It follows from the definition of the question-update that only one *dm*-variant survives the update. So, sets of carriers that are *dm*-variants of each other collapse to a single element (this follows from the nature of the question update). In this case, we have a function f_q mapping σ to $\sigma[q(\varphi)]$. This does not yet guarantee that the same answer is given. But we can express this adequately by the demand that f_q preserves object-identity: f_q maps object-identical carriers to object-identical carriers.

In (30), a definition is given that expresses these two demands directly in terms of object-identity and *dm*-variance. If σ does not answer the question, there are carriers without an answer or there are two carriers that have the same ideas about the structure of the world, but different ideas about the facts which determine different answers.

(30) σ answers $q(\varphi)$ iff
1. $\forall i \in \sigma \; \exists! j \in \sigma[q(\varphi)] \; i =_{dm} j$
2. $\forall i, j \in \sigma \; (oid(i,j) \rightarrow \forall k, l \in \sigma[q(\varphi)] \; (i =_{dm} k \wedge j =_{dm} l \rightarrow oid(k,l)))$

In essence, our definition is the same as that in GS. Object-identity composed with *dm*-variance for a fixed set of discourse markers is an equivalence relation.

(31) $iRj \Leftrightarrow_{def} \exists k \; (oid(i,k) \wedge k =_X j)$

Within an equivalence class of that relation we can form a partition by looking at the actual values for the discourse markers.

(32) $iQj \Leftrightarrow_{def} \forall x \in X \; ix = jx$

Consider $\tau = \sigma[q(\varphi)]$ and let $X = dm(\varphi)$. Consider $i/R/Q$ with respect to τ. Now, $i/R = i/R/Q$, i.e. the partition determined by Q over i/R is $\{i/R\}$, if σ answers $q(\varphi)$.

What about truth, rigidity and exhaustivity?

The point of update semantics is that we do not normally know what the truth is. There are no privileged carriers[9]. Experience also teaches us that our information is fallible. So there seems to be no content here to the idea that an answer must be true, apart from the fact that if we have loaded the question, i is in the set of worlds that give the same answer and mean the same thing with that answer, as in GS. But here this is only a triviality.

If we demand that all carriers are related by our relation R, answers are rigid: if σ answers $q(\varphi)$ the question update guarantees that the discourse markers of φ are defined and rigid over σ. What we have in fact for answering a question is rigidity modulo R.

[9]Following Heim(1983), we could add a priviliged information carrier in standard information states. This would not help us with the problem at hand however as we have no notion of the denotation of a question (its answer) in a carrier. The function f_q can be considered as an approximation of this notion, but it gives carriers, not propositions as its result. There is no guarantee that any carrier will be mapped to itself by this function.

Exhaustivity is clearly present in the current approach.

I have little to say about embedded questions, which formed the discovery ground of the GS theory. The problem is primarily that no fully satisfactory semantics for the attitudes has so far been given in update semantics.

Let me sketch a possible approach to knowledge. We assume that for $i \in \sigma$ we have a set of carriers K_i representing John's knowledge. We can then define knowledge updates in the following way[10]:

(33) $\sigma[K\varphi] = \{i \in \sigma[\varphi] : K_i[\varphi] = K_i\}$

This definition guarantees by means of an extra update that φ also holds in σ after the update. It would be more proper to arrange this by presupposition, as the projection behaviour is not correctly characterised by putting an extra update inside the larger update, but we will not bother about this here. We demand[11] that $\sigma \subseteq \bigcap_{i \in \sigma} K_i$. We must also guarantee that K_i and i have the same ideas about discourse markers. A tentative demand lets i and $j \in K_i$ be related by the composition of object-identity and variance for variables outside the discourse markers of σ that are also markers of K_i.

How do things work out with an example like (34)?

(34) John knows who sleeps.

We must form $\sigma[\textit{who sleeps}]$ and reduce that to those i in which K_i is a fix-point for the update with $q(x \wedge sleep(x))$. Since i is in $\sigma[q(x \wedge sleep(x))]$, i agrees with K_i on the value of x.

In σ, however, it need not be clear who sleeps, so there may be object-identicals j of i that have an x-variant in $\sigma[q(x \wedge sleep(x))]$ which is also a fix-point for the update with $Kq(x \wedge sleep(x))$.

If the answer is known this cannot happen. So like in GS, we have (35):

(35) Bill sleeps.
John knows who sleeps.
ergo
John knows that Bill sleeps.

and its negative variant.

How about other environments of indirect questions? I only have a version of *wonder whether* here, that is definable as the negation of our knowledge operator (in fact, precisely because we did not treat knowledge as presuppositional).

[10]This is the same treatment as the one suggested by Kamp for belief according to Heim (1992). There are alternatives for this definition.

[11]We know that what somebody knows is true, but we normally do not know everything that this person knows: of all the person knows, we know that it must be consistent with our information. If we know the total of someone's knowledge, the information state determined by it must be a superset of our information state.

(36) John wonders who sleeps

(36) comes out as an update that eliminates the carriers in which John knows who sleeps and their *dm*-variants. It is clear that this time the corresponding inferences are blocked.

So, I conclude that nothing indicates that it is impossible to develop indirect questions on the lines we have indicated so far for direct questions.

4 Topic and Focus

The idea that topic and focus are related to exhaustivity goes back to Szabolcsi (1981). In her theory (but in my terms), a focussed constituent is interpreted as supplying the answer to the question resulting from omitting it in the sentence and replacing it by a suitable Wh-element. This same theory is also defended -but without the exhaustification- in the work of Van Kuppevelt (1991), who extends the theory with a connection to the theory of discourse: any sentence should be viewed as an answer to an explicit or implicit question.

How do we find out about the question? A popular view, suggested by work on operators like *only* and *even* and on subjunctives (Kasper (1992)), is that it derives from a binary division coded into the form of the utterance by a variety of devices in different languages: syntactic position, case-marking, intonation etc. Others (e.g. Vallduvi (1991)) assume a tripartite structure.

I want to suggest that it is not necessary to assume a formal division and that indeed this is a view that is hard to maintain. I will also argue against the view that we are dealing with a binary or ternary division. Kasper (1992) convincingly shows that in many cases, we must divide the semantical content of a word into a presupposition and an asserted part in order to obtain a sensible construction of the meaning of the subjunctive sentences. He equally convincingly argues that this division cannot be made once and for all in the lexicon: different contexts lead to different divisions. It follows that these divisions cannot be formally marked by any device unless we assume syntax beneath the word level, intonational patterns that select part of a word meaning, lexical marking of focus, or similar unconvincing stratagems. What we are left with for the interpretation of the formal devices are just constraints: in particular constraints that tell us what cannot be topic, e.g. the NP marked by *wa* in Japanese cannot be in the topic of the sentence in its entirety, post-Wackernagel material cannot appear (with the exception of verbal material) entirely in a topic, focus-intonation similarly indicates that some of its material must stay out of the topic. Binary (or ternary) divisions are easy to mark in natural languages (compare quantification, subordination). So the variety of means of expression indicates that we are not dealing with a binary division. The fact that ternary divisions have been proposed also points in this direction.

Our current context suggests a simple solution. We adopt a semantics for the sentence as a whole. This semantics allows a set of abstractions, the questions that the sentence could possibly answer. Certain of these questions are ruled out by focus marking. Other questions are already answered by the information state. The remaining questions together form the topic or topics of the sentence. We obtain the informational contribution of the sentence by asserting the sentence (its representation with slots unified with Wh-elements in the topics) in an information state to which we have added each of the topics as a question.

This predicts a series of exhaustification effects, which indeed we find in some cases. Sometimes however it appears as if there is a unique question. These are the cases like (37).

(37) John likes MARY

The reconstruction into a question and an answer to the question can be performed in a number of ways, depending on the Wh-element chosen. (38) lists some possibilities.

(38) a. Who does John like? Mary.
b. Which girl does John like? Mary.
c. Which of Jane and Mary does John like? Mary.

It is the information state that determines which one is chosen. If we know John likes a girl, or that he likes one of Jane and Mary, the last two questions are the topics that apply. If we know nothing about the answer to (38a) , we must choose that as the topic.

The variation in possible topics increases if we consider larger foci, as in (39):

(39) John [LIKES MARY]

which may question John's emotional attitudes towards girls, John's liking of people in general, etc.

The assignment of a focus to a sentence is not unique, and even when it is unique, it does not give rise to a unique question.

Suppose we know has John has a farm and we are wondering about his life-stock. The assertion (40)

(40) John has 5 sheep.

is then naturally interpreted as answering a series of questions

(41) a. Does John have life-stock?
b. What life-stock does he have?
c. How many X does he have?

where the (b) and (c) answers are responsible for the implicatures that John has no goats or cows and that he does not have 6 or 12 sheep.

The update semantical framework provides some support in identifying the topic. First of all, topics can be regarded as accommodatable presuppositions. If they are found in the context, either directly or as a subquestion of an earlier question, they are preferred over topics that must be accommodated. Second, topics must be proper questions with respect to the information state. It does not do to select a topic that is already answered in the information state. This can filter out topics that are improper questions.

A third filter, studied by Van Kuppevelt in recent work is the way in which one sentence can limit the range of topics allowed for the next sentence.

The main problem for this approach is that we need to explain he occurrences of *only*. Given our analysis, *only* applies to a focus. Adding *only* to a sentence with a given focus would be a semantically empty operation. This is illustrated in (42).

(42) Who does Mary love? She loves only John.
Mary likes ONLY BEANS.

In both examples, the *only* appears to be superfluous. In (42a) because of the exhaustivity of answers, in (42b) because of the exhaustivity of foci. If we do not assume that we are completely on the wrong track, an explanation must be available for these occurrences of *only*. There are two possibilities. One is that *only* here functions as a pragmatic marker indicating that the answer goes against the expectations of the interlocutor: he or she would expect that Mary loves more people or likes more vegetables. This would place *only* on a par with *even* which reverses such expectations. Another explanation could come from the underspecification involved in determining the precise topic: *only* could enlarge the extension of the restriction on the hidden Wh-phrase in the topic (maybe from the contextually given set of alternatives that would otherwise be picked up to the full range of the possibilities) and thereby strengthen the exhaustivity. In both cases, *only* would have a role that is much less semantic than has been generally assumed.

A last remark is in order. The actual syntactic possibilities for Wh-phrases are not a good guide in determining the range of topics. The formalism we developed for representing questions is richer and can provide better guidance in identifying topics.

5 Plurals

That the framework of generalised quantifiers offers an interesting framework for analysing plural determiners has been proven by a constant stream of publications. My aim in this subsection is to provide an alternative not for those insights but for the direct application of generalised quantifiers in natural language semantics. The theory I want to propose is close to early DRT. It has the advantage that it is simple-minded, the disadvantage that, without further additions, it is not correct.

In the naive theory, plural NPs are a special kind of definites and indefinites, introducing discourse markers for sets, generally of cardinality ≥ 2. The interpretation of these sets is constrained in a number of ways by the NP.

The most important constraint is that the set always belongs to the extension of the noun. The noun sometimes has an anaphoric role referring back to an earlier plural referent included in the noun extension. In these cases, the denotation of the NP-referent must belong to this subset of the extension.

The denotation also meets some conditions deriving from the determiner. These conditions constrain the set or the relation between the set and the noun denotation. There may also be contextual parameters, especially in the case of vague determiners (many, few, etc.).

A last type of constraint that the determiner can impose concerns the binding of the argument place occupied by the NP. Certain determiners prefer a distributive interpretation (*many, every*). This is however not the only source of distributivity.

Another claim about determiners are that certain of them are negations of definite and indefinite determiners. Negation makes them not definite or indefinite. Examples are *no* and *few*.

The last claim is not about NPs but about the predicates and relations in which they occur. Certain of these obey special constraints. In particular, I will assume that one place predicates come in three varieties. Here and below I will use $x \in y$ as an abbreviation for $x \subseteq y \wedge \#x = 1$, a harmless notation as set membership will not be used as such.

All one-place predicates will obey the postulate (43)(closure under union).

$$(43) \qquad \forall x \in y\ \exists z \subset y\ x \in z \wedge Pz \rightarrow Py$$

Some will allow distributive readings, but also have collective ones. This gives us (44)(a special case of closure under union).

$$(44) \qquad \forall x \in y\ Px \rightarrow Py$$

And finally there are ones that are only distributive, giving the MP (45), typical for nouns.

$$(45) \qquad \forall x \in y\ Px \leftrightarrow Py$$

The situation with many-place predicates is even more complicated. In this paper we will only consider the postulate (46), here formulated for a 2-place predicate.

$$(46) \qquad \forall v \in x\ \exists w \in y\ Rvw \wedge \forall v \in y\ \forall w \in x\ Rwv \rightarrow Rxy$$

Other postulates are obtained by applying the meaning postulates for one-place predicates in turn to each of the argument places.

An example of such a derived postulate for 2-place relations is (47). This relation allows distributivity over both coordinates.

$$(47) \qquad \forall v \in x\ \forall w \in y\ Rvw \rightarrow Rxy$$

5.1 Some constraints

Let x be the discourse referent of the NP, d be the extension of the noun (or the contextually determined restriction of that extension). The determiners *all, every* and *the* provide the constraint (48)

(48) $x = d$

all the others that are not negative, the constraint (49).

(49) $x \subset d$

Using strict subset here is essential as will become clear later.

A number of determiners provide cardinality constraints. Some of these are stated in the following table. The variable n is a number value provided by the context.

$\#x = 3$	three
$\#x > \#(d - x)$	most
$\#x \geq 2$	some
$\#x \geq 2$	a few
$\#x < 3$	less than three
$\#x > 5$	more than 5
$\#x \geq n$	many

This determines most of the determiner meanings. The negative ones can now just be defined in terms of the positive ones:

(50) no:= not a
no:= not some
few:= not many

Every, many, most and *each* bind the predicate in a special way: they demand that each of the members of the discourse referent meets the condition of the predicate. In (51) I provide an update definition for distributivity. There is also a definition for fullness (intended for the semantics of *all.* These definitions are the same but for the fact that fullness continues to work for collective readings and mass interpretations of the variables.

(51) distributivity: $\sigma[dist(x, \varphi)] = \{i \in \sigma : \forall j\ j =_x i \wedge jx \in ix \rightarrow j \in \sigma[\varphi]\}$
fullness: $\sigma[full(x, \varphi)] = \{i \in \sigma : \forall j\ j =_x i \wedge jx \subseteq ix \rightarrow j \in \sigma[\varphi]\}$

In (51) distributivity is defined by quantifying over x-variants j of i that assign members of ix to x, for fullness, we quantify over all parts of ix.

I do not consider bare plurals and *pars pro toto* readings of definite plurals.

Below a combination of the constraints is provided, combining with the verb *to run*.

A boy runs.	$x \wedge x \subset BOY \wedge \#x = 1 \wedge run(x)$
Some boys run.	$x \wedge x \subset BOY \wedge \#x \geq 2 \wedge run(x)$
The boy runs.	$x \wedge BOY = x \wedge \#x = 1 \wedge run(x)$
The boys run.	$x \wedge BOY = x \wedge \#x \geq 2 \wedge run(x)$
All boys run.	$x \wedge BOY = x \wedge \#x \geq 2 \wedge full(x, run(x))$
Every boy runs.	$x \wedge BOY = x \wedge dist(x, run(x))$
Three boys run.	$x \wedge BOY = x \wedge \#x = 3 \wedge run(x)$
Few boys run.	$\neg(x \wedge x \subset BOY \wedge \#x > n \wedge dist(x, run(x)))$
Many boys run.	$x \wedge x \subset BOY \wedge \#x > n \wedge dist(x, run(x))$
Most boys run.	$x \wedge x \subset BOY \wedge \#x > \#(BOY - x) \wedge dist(x, run(x))$

Why is the treatment inadequate as it stands? The reason is simple: it is incapable of dealing with exhaustive quantifiers like precisely 2 or readings of quantifiers like 2 in which they carry an exhaustive interpretation.

Suppose there are five sleeping boys. Then both (52)

(52) Less then four boys sleep
$x \wedge boy(x) \wedge \#x < 4 \wedge sleep(x)$

and (53) are true under a standard DRT-interpretation: just take a smaller subset of the sleeping boys.

(53) Precisely 2 boys sleep
$x \wedge boy(x) \wedge \#x = 2 \wedge sleep(x)$

Another example that comes out wrong is the infamous cumulative reading of e.g. (54).

(54) 4 boys danced with 5 girls

(54) is true (due to our relational meaning postulate) when the cumulative interpretation is true (the total number of boys who danced with girls is 4 and the total number of girls they danced with is 5). Unfortunately it also is true if five boys danced with six girls (in the cumulative reading), i.e. when it is intuitively false.

5.2 Repairs

This paper started out about 10 years ago with an attempt to develop a question theory in DRT and received a strong impulse from a clever solution of Hans Kamp (p.c.) of the problem of cumulative quantification. My final solution is still very close to his idea, which can be recapitulated in the following three steps, applying to the example (55).

(55) 200 Dutch firms own 600 American computers.

(56) 1. reinterpret the relation in a cumulative way (the $\forall\exists\forall\exists$ meaning postulate)
2. apply the "naive" approach (section 3) to obtain a DRS
3. exhaustify the resulting DRS

My one change is to do the exhaustification beforehand by updating with (57)which is an instance of topicalisation. (The question can be glossed as: How many Dutch firms own how many American computers.)

(57) $q(n \wedge m \wedge x \wedge y \wedge \#x = n \wedge \#y = m \wedge \neg\neg(dutch\_firm(x) \wedge american\_computer(y) \wedge own(x, y)))$

After this update we then add the naive (58) with next to the variable sharing the invisible unifications $n = 60$ and $m = 300$.

(58) $x \wedge y \wedge \#x = 60 \wedge \#y = 300 \wedge dutch\_firm(x) \wedge american\_computer(y) \wedge own(x, y)$

This is an ordinary instance of assigning a topic to a sentence.

The other quantifiers operate in much the same way. We will treat the exhaustivity effects under the heading of scalar implicatures.

6 Scalar Implicatures

Scalar implicatures is another area in which exhaustification does provide a direct explanation, independently of pragmatic maxims. If we analyse (59)

(59) John has four sheep.

as (60)

(60) $q(x \wedge have(j, x)) \wedge \#x = 4 \wedge sheep(x)$

or as (61)

(61) $q(n \wedge \neg\neg(\#x = n \wedge x \wedge have(j, x) \wedge sheep(x))) \wedge n = 4$

it cannot be that there are more than 4 sheep that John owns. If there are, we can form another set of 4 sheep owned by John who are not contained in the set chosen as value for x.

In this way, exhaustification explains all of the (pragmatic) implicatures of the form (62) for n a number greater than 4.

(62) John does not have n sheep

If we want to apply exhaustification to other cases of scalar implicature, things turn out to be more complicated than in this numerical case. Indeed, some of the cases discussed below can be regarded as arguments against the reduction of scalar implicatures to exhaustification. In general we have to make quite a number of special assumptions.

The first group of examples are formed by monotone increasing determiners like *some, most, at least three etc.* that seem compatible with the application of *all.* It seems to hold that if $A(det\ N)$ holds with *det* one of the mentioned determiners and A a simple context that does not bring the *det N* configuration into the scope of a quantifier or a negation, there is a scalar implicature that *not A(all N).* An example is (63).

(63) statement: Most sheep died.
implicature: Not all sheep died.

Now given an analysis of the determiners in question, exemplified in (64),

(64) $x \wedge sheep(x) \wedge die(x) \wedge \#(SHEEP - x) < \#x$

it is possible to assign the set of all sheep to the discourse marker introduced by the NP. It seems, however, that it is natural enough to introduce a condition to the semantics of these determiners that forces their referent to be subclass of the class indicated by the noun in the context[12]. This is a natural amendment as moving from a given class to a subclass is the discourse function of these determiners (a function that they share with the cardinals). In this way, we would have a semantic analysis of *some* that would work out on (65) as follows:

(65) Some sheep died.
$x \wedge x \subset SHEEP \wedge \#x \geq 2 \wedge die(x)$

Here *SHEEP* is whatever set is the denotation of the noun *sheep* in the context (a set of sheep introduced earlier on the discourse, the set of sheep within some context restriction or the set of sheep at the reference time and place).

Exhaustification does not work if all sheep died: unless SHEEP happens to have cardinality 2, none of the sets meeting the two conditions would be maximal. If it is 2, x is not a proper subset.

A similar analysis applies to the other determiners.

[12] We already introduced this constraint in an earlier section.

(66) Most sheep died.
$x \wedge x \subset SHEEP \wedge \#(SHEEP - x) < \#x \wedge die(x)$
At least three sheep died.
$x \wedge x \subset SHEEP \wedge \#x \geq 3 \wedge die(x)$

Notice that under the conditions that $SHEEP$ is plural, and has more than 3 members, if it holds (perforce exhaustively) that all sheep died, it still follows (non-exhaustively) that *some*, *most* and *at least three* sheep died. In our theory, it follows that the questions in (67)

(67) Did some sheep die?
Did most sheep die?
Did at least three sheep die?

must be answered in the positive, though the affirmative sentences can not be used with the NP or determiner in focus.

So the normal entailments[13] come out correctly and it even holds that in the restricted sense that answers to the corresponding questions must be answered positively, these quantifiers remain monotone increasing.

Sometimes, it is better to interpret scalar implicatures as part of another phenomenon. A case in point is the scalar implicature around *or* in (68).

(68) John has sheep or John has goats.

(68) seems to exclude that John has animals of both kinds. We could introduce a variable that can be classified by *or* and *and* (in a mutually exclusive way) and introduce a suitable partial ordering on the values of that variable, e.g. as in (69),

(69) $z \wedge connection(z, sheep, goats) \wedge or(z)$

The values of z could be taken from the domain of connectives ordered by implication.

To be precise:

f is a connective if $f \in 2^{2 \times 2}$,
$f \leq g$ iff $\forall x g(x) \leq f(x)$,
$or(z)$ iff $z = \{<< 0, 0 >, 0 >, << 1, 0 >, 1 >, << 0, 1 >, 1 >, << 1, 1 >, 1 >\}$ and
$connection(z, p, q)$ iff $z(< p, q >) = 1$.

If John has both sheep and goats, disjunction is a proper value for z, but not an exhaustive one. Conjunction is however exhaustive. *Or* itself can never be exhaustive, since whenever it

[13]Entailment intuitions can be reconstructed in two ways: (a) given that we know the premises can we answer *Yes* to the yes-no-question formed from the conclusion or (b) given that we know the premise can we sincerely and correctly assert the conclusion. For many examples in standard logic only the first interpretation can be maintained.

holds, there is a stronger connective (*and*, *left and not right* or *right and not left*) that wins out over *or*. So the explanation fails.

The approach is however not very natural to begin with. First, the addition of an extra variable is not warranted by anaphoric phenomena (the connection cannot be picked up by pronouns). Second, the scalar implicature can equally well be derived from the clausal implicatures associated with disjunction, though in a slightly weaker form (if the speaker does not know of either disjunct that it holds, it follows that she does not know their conjunction). This would be preferable.

Anaphora occurs with scalar implicatures like the ones in (70).

(70) John's sheep is rather heavy.
implicature: John's sheep is not extremely heavy.

After the first example we can continue with (71)

(71) Bill's sheep is just as heavy.

which seems to pick up the heaviness of John's sheep to apply it to Bill's sheep. In this way we can analyse (71) along the lines of (72) and obtain the implicature in the usual way.

(72) $q(w \wedge weight(w, s)) \wedge rather\_heavy(w)$

Here w can be thought of as a positive real, $weight(w, x)$ applies whenever weighing x gives a greater value than w[14], and *rather heavy* applies to an interval of weights distinct from that to which *extremely heavy* applies. Thus we maintain the entailment from *extremely heavy* to *rather heavy*, while obtaining a maximal value for the weight of John's sheep if exhaustification applies.

A final case is provided by the scale *know, belief.* Here it not evident that we can pick up the attitude anaphorically, though we come close. It is fine to have sequences like (73).

(73) John strongly resents that Bill's sheep have eaten his flowers.
Bill feels the same way about John's sheep eating his cabbage.

The corresponding *Have the same attitude towards* (the analogue of *Feel the same way* for *know* and *believe*) is however rather contrived. We can use the same scheme as we employed earlier on.

(74) $q(a \wedge attitude(a, j, p)) \wedge belief(a)$

Here a could be taken to range over mental states of subjects and we could take them to be ordered by content. Thus a regret of x towards p always contains a knowledge of x that p

[14] I weigh one kilo, non-exhaustively. This appears to be true. Similarly, I am one foot tall, but not seven.

which contains a belief of x that p. The regret, knowledge and belief are taken to be different states (no knowledge is a belief or inversely) that may partially constitute each other. This is the partial ordering we require for applying exhaustification.

As it is now fairly common to have a state parameter in the analysis of these sentences for the purpose of temporal processing, and others have argued for a view in which one state can be partial constituent of another, this analysis seems comparatively unproblematic.

As usual, the inference from a knowledge to a corresponding belief is maintained for the non-exhaustive reading of the belief.

6.1 Cancellation

There is a problem with what we have seen so far: the phenomenon of cancellation for scalar implicatures. So far we have pretended that exhaustification applies all the time. Cancellation tells us that the application of exhaustification must be limited to certain cases.

The point in treating scalar implicatures as pragmatic implicatures rather than entailments is precisely that there are exceptions to their application. Sometimes they apply, sometimes they do not.

(75) Does Leif have three chairs?
Yes, Leif has three chairs.

Following Kadmon (1986), the answer does not implicate that Leif has precisely three chairs. It may be that 3 chairs are needed for seating some extra guests, but that Leif owns 6 chairs in total.

Other means of cancelling the implicatures are connected with explicit cancellation and so-called twiddly intonation, which seems to be used to indicate that other things could be filled in as well.

(76) a. Leif has three chairs, allright, but he may have more.
b. Leif has three -even six- chairs.
c. Leif has thReE chairs.

As exhaustification is connected with the topic-focus division in the sentence, it follows that all kinds of cancellation must be related to means of influencing this division. An explicit question changes the division: if possible, the topic will coincide with the question. The explicit question thereby cancels the exhaustivity of the answer. Provisions also form a restriction on the topic-focus division. Constructing the topic as: *How many chairs does Leif own?*, i.e. making *three* the focus, for (76b) is contradicted by the interjection. Thereby, only the weaker question *Does Leif have three chairs?* can be the topic, with a treatment of the rejected topic included in the interjection. In (76a) , the proviso similarly forces a weaker topic. Finally in (76c) , the phenomenon of twiddly intonation is characteristic of a topic resetting and should here make it impossible to make *three* focus.

It is not the sentence as such that forms an exception to exhaustification. Cancellation can be limited to part of the sentence, while other quantifiers remain exhaustive. Compare (77).

(77) 3 boys kissed most—maybe all— girls.

One phenomenon that may be reduced to scalar implicatures, in our reconstruction, are the Evans-effects. Evans' observation is that there is a crucial difference between saying (78):

(78) John has sheep. Bill shaves them.

versus (79).

(79) John has sheep, that Bill shaves.

In the first, but not in the second case, Bill shaves all of John's sheep.

A treatment can be based on topic and focus. In the case of the single sentence, the focus can only include the whole NP, not the NP without the relative clause (this would only be allowed if the relative clause were non-restrictive). In the other case, we assume that the discourse referent of the NP *sheep* receives an exhaustive interpretation by being in focus.

7 Conclusions and Further Work

The picture that emerges is that every assertion either answers an existing question or existing questions or constructs the question(s) that it answers on the fly. The information that we gather from sentences is partially determined by the question we assume it answers. I have tried to present the case that various "implicatures" can be explained by referring to this question. Of course, they are no longer implicatures as they are not fallible. The apparent possibility of cancellation must be understood as different possibilities for finding the answered question.

The current approach to scalars together with update semantics approaches to clausal implicatures, makes it necessary to reinterpret the phenomenon of conversational implicatures. There is a class that is directly connected with basic interpretation (clausals and scalars) and that can be captured by a discourse grammar. Typical of this class is that does not require sophisticated reasoning. On the other hand, there can be no grammatical alternative for the implicatures generated by flouting maxims. Here we typically require reasoning about goals of the speaker and alternatives for reaching the communicative goals.

Notice that we manage to conform to Grice's original aim: maintain a simple logic and explain special effects by an additional mechanism. The mechanisms involved in clausal and quantity implicatures is simpler than the reasoning about communicative behaviour proposed by Grice. Such reasoning however remains indispensable for explaining the full range of perlocutionary effects.

Further work will be necessary to integrate the present results adequately in a grammatical framework.

Some of the themes discussed in this paper have the flavour of going back to old battlefields around DRT. The use of descriptions as an alternative for the DRT-analysis finds important arguments in the Evans-phenomena. With a mechanism like the one proposed here, the differences largely disappear. The fact that, for adequate analyses of the plural, it seems imperative to use the generalised quantifier structures proposed by Montague casts doubt on the general spirit of the analyses proposed in early DRT for the singular NPs. I believe these arguments no longer hold in a setting like the current one.

The theory of questions of section 4 is also simpler (on the formal level, not in the semantics as such) than the theory on which it is based (the Montague grammar approach of Groenendijk & Stokhof). It is possible to see the questions here as concepts of their true, rigid and exhaustive answers on the metalevel, by using functions f_q standing in for the denotation relation. Equally well, we can see the Karttunen theory embodied in this procedure. Equating a possible answer with an element of $\sigma[q(\varphi)]$ together with $q(\varphi)$ associates a set of possible answers with the question. (In case we perform the update on the empty information state, the set of all possible answers). Factual answers select possibilities from this set. On this level there is little to motivate questions as denoting their answer.

The real motivation for a denotational view in GS are formed by the indirect questions. The distinction between the predicate *know* as an extensional predicate of questions and *wonder* as an intensional one is elegant and hard to resist within the framework of Montague grammar and indeed within any framework based on the Fregean distinction. The alternative of making a distinction between predicates that (lexically) presuppose the existence (truth) of an argument and ones which do not is however equally general and correct, both for *know* and *wonder* and for *eat* and *want*.

Many questions have received only a tentative answer in this paper and require further work. I have tried to show how a simple mechanism can be applied in a number of important areas and avoid some of the usual complexities there.

Further research is necessary to come with an overall treatment of those Gricean implicatures that allow a grammatical treatment. More work is necessary to give an explicit treatment of clausal implicatures (following the lead of Stalnaker's theory of assertion) This will involve a treatment of speaker and hearer beliefs in the common ground, a treatment that is currently not available.

Within the question theory outlined here the two most important research questions seem to be the treatment of indirect questions within a more complete and adequate treatment of the attitudes and the treatment of complex answers. Answers typically can have the form of a list:

(80) Who likes which animals?
Mary likes poodles, Harry likes ants and Jane likes donkeys.

The problem is that we need to express that the answers together are exhaustive, while avoiding that any of the three is. The curious lumping of the elements in a list, discussed briefly in Prüst & al. seems to provide an answer, but this aspect should be studied in more depth. What we can do in our context is to update with the disjunction of the three answers. A precise treatment has to wait for a future occasion.

The sketchy topic-focus treatment needs a serious confrontation with the empirical facts before we can come to accurate predictions concerning the precise topics that operate in a particular utterance.

References

Dekker, P. *Transsentential Meditations. Upds and Downs in Dynamic Semantics.* PhD Dissertation, University of Amsterdam, 1993.

Evans, G.: 1977, 'Pronouns, Quantifiers, and Relative Clauses (I & II),' *The Canadian Journal of Philosophy* **7**, 467–536 and 777–797.

Gardent, C. (1991) Gapping and VP-ellipsis in a unification-based grammar. PhD Thesis, Cognitive Science, University of Edinburgh.

Groenendijk, J. & Stokhof, M. *Studies on the Semantics of Questions and the Pragmatics of Answers.* Academisch Proefschrift, Universiteit van Amsterdam 1984.

Groenendijk, J. & Stokhof, M. Two Theories of Dynamic Semantics, in J. van Eijck (ed.) *Logics in AI*, Springer, Berlin 1990.p. 55-64.

Heim, I. *On the projection problem for presuppositions*, *WCCFL* **2** (1983). p. 114-26.

Heim, I. File Change Semantics and the Familiarity Theory of Reference. In: Bäuerle, Schwarze and Von Stechow (eds.) *Meaning, Use and Interpretation.* De Gruyter, Berlin 1983.

Heim, I. Presupposition Projection and the Semantics of Attitude Verbs. *Journal of Semantics* 9, 3 (1992), p. 183-222.

Kamp, H. A Theory of Truth and Semantic Representation. In: J.A.G. Groenendijk, T.M.V. Janssen & M.B.J. Stokhof (eds.). *Formal Methods in the Study of Language.* Mathematical Centre, Amsterdam 1981. p. 277-322.

Kamp, H. & U. Reyle. *From Discourse to Logic.* Kluwer, Dordrecht 1990.

Karttunen, L. Presupposition and linguistic context.*Theoretical Linguistics* **1**,1974, p.181-94

Kasper, W. Presuppositions, Composition and Simple Subjunctives. *Journal of Semantics* **9**(1992), p. 307-331.

Kuppevelt, J. van. *Topic en Comment. Expliciete en Impliciete Vraagstelling in Discourse.* PhD Thesis, University of Nijmegen, 1991.

Prüst, H., R. Scha & M. van den Berg. Discourse Grammar and Verb Phrase Anaphora. In: *Linguistics and Philosophy* **17**:261-327 (1994).

Scha, R & D. Stallard. Multi-Level Plurals and Distributivity. In: Proceedings of the 26th Annual Meeting of the ACL, Buffalo, NY, June 1988.

Scha, R. Distributive, collective and cumulative quantification. In: J.A.G. Groenendijk, T.M.V. Janssen en M.B.J. Stokhof (eds.) Formal Methods in the Study of Language, part 2. 1981. p. 483-513.

Stalnaker, Robert. Assertion. In: Cole, Peter (ed.): *Syntax and Semantics 9: Pragmatics.*

New York 1978, p. 315 - 332.

A. Szabolcsi. The Semantics of the topic-focus articulation. In: J.A.G. Groenendijk, T.M.V. Janssen en M.B.J. Stokhof (eds.) Formal Methods in the Study of Language, part 2. 1981. p. 513-541.

Veltman, F. Defaults in Update Semantics. Accepted by *Journal of Philosophical Logic.* (ms. 1994, University of Amsterdam)

Vermeulen, C. F. M. *Merging without Mystery: Variables in Dynamic Semantics.* to appear in *Journal of Philosophical Logic.*

Zeevat, H. Presupposition and Accommodation in Update Semantics. In: *Journal of Semantics* **9** 1992, p. 379-412.

Zeevat, H. & Scha, R. Integrating Pragmatics in Update Semantics. In: A. Ortony, J. Slack & O. Stock (eds.) *Communication from an Artificial Intelligence Perspective: Theoretical and Applied Issues.* Springer Verlag Heidelberg, 1992.