Pragmatic Meaning and Non-monotonic Reasoning: The Case of Exhaustive Interpretation

Katrin Schulz and Robert van Rooij

Abstract. In this paper an approach to the *exhaustive interpretation* of answers is developed. It builds on a proposal brought forward by Groenendijk & Stokhof (1984). We will use the close connection between their approach and McCarthy's (1980, 1986) *predicate circumscription* and describe exhaustive interpretation as an instance of *interpretation in minimal models*, well-known from work on counterfactuals (see for instance Lewis (1973)). It is shown that by combining this approach with independent developments in semantics/pragmatics one can overcome certain limitations of Groenenedijk & Stokhof's (1984) proposal. In the last part of the paper we will provide a Gricean motivation for exhaustive interpretation building on work of Schulz (to appear) and van Rooij & Schulz (2004).

Keywords: Conversational Implicatures, Exhaustive Interpretation, Predicate Circumscription, Relevance, Only knowing.

1. Introduction

The central aim of this paper is to find an adequate description of the particular way in which we often enrich the semantic meaning of answers. To illustrate the phenomenon, consider the following dialogue.

(1) Ann: Who passed the examination?
Bob: John and Mary.

In many contexts Bob's answer is interpreted as *exhausting* the predicate in question, hence, as stating not only that John and Mary passed the examination, but also that these are the only people that did. This reading is called the *exhaustive interpretation* of answers (see e.g.

Groenendijk & Stokhof (1984), von Stechow & Zimmermann (1985)) which we will study in this paper.[1]

The term 'exhaustive interpretation' has not only been used in connection with the interpretation of answers. Aspects of the meaning[2] of sentences containing 'only' (compare 'Only John and Mary passed the examination'), cleft constructions ('It was John and Mary who passed the examination') and free intonational focus ('[John and Mary]$_F$ passed the examination'), for instance, have been characterized in this way as well. In this paper, however, we will limit ourselves to a description of the exhaustive interpretation of answers. We will discuss semantic analyses of these other constructions only insofar as they have to do with problems that arise with the exhaustive interpretation of answers as well.

In their dissertation from 1984, Groenendijk & Stokhof proposed a very promising approach to the exhaustive interpretation of answers. We will introduce this approach in section 3 and discuss its merits. However, Groenendijk & Stokhof's (1984) description of exhaustive interpretation also faces certain shortcomings. The main goal of the remaining sections is to overcome these limitations.

In section 4 we will discuss the close relation between Groenendijk & Stokhof's (1984) approach, McCarthy's (1980, 1986) theory of predicate circumscription, and the latter's model-theoretic variant: interpretation in minimal models. We will then switch to a description of exhaustive interpretation as interpretation in minimal models and show that this already allows us to address some of the problems Groenendijk & Stokhof (1984) have to face.

---

[1] As will become clearer in section 2, we will treat the particular reading examplified in (1) as only a special case of exhaustive interpretation.

[2] In this paper 'meaning' will be used as referring to all the information conveyed by an utterance in a particular context.

In section 5 another modification is added: we will combine the new approach with dynamic semantics. Given the developments in semantics during the last 20 years, this is an alteration of the original static approach of Groenendijk & Stokhof (1984) that would have been necessary anyway. It will turn out that it solves some problems, already discussed by Groenendijk & Stokhof (1984) themselves, concerning, for instance, the interaction of exhaustive interpretation and the semantics of determiners.

In section 6 we will address the context-dependence of exhaustive interpretation. As will be illustrated in section 2, exhaustive interpretation can come in other forms than the reading we discussed for example (1). We will argue that this should be explained by taking a contextual parameter of relevance into account.

In the final section we will go beyond our primary aim to provide an adequate description of exhaustive interpretation. The need of such a description arises because standard semantics cannot handle the phenomenon. That means that if we want to maintain standard semantics, exhaustive interpretation cannot be explained as a semantic phenomenon. But where does it come from if not from semantics? One answer to this question that seems to be particularly attractive is to analyze it as a Gricean conversational implicatures. We will sketch a formalization of parts of Grice's theory brought forward by Schulz (to appear) and van Rooij & Schulz (2004). It can be shown that when combined with a principle of competence maximization, this formalization indeed accounts for exhaustive interpretation (as described in section 4).

2. The phenomenon

Before we can start thinking about how to formulate a general and precise description of the exhaustive interpretation of answers, we first need to get a clearer picture of what we actually have to describe. Therefore, this second section is devoted to a closer investigation of the properties of exhaustive interpretation.

2.1. Interaction with the semantic meaning of the answer

The first thing to notice is that an exhaustive interpretation does not always completely resolve the question the answer addresses. Consider, for instance, example (2) (all sentences discussed in this section should be understood as answers to the question 'Who passed the examination?').

(2) Some female students.

This answer can be interpreted exhaustively as stating that just a few students passed the examination and that they are all female. However, also on this reading the answer does not identify the students that passed the examination and therefore does not resolve the question.[3] Hence, even though the exhaustive interpretation strengthens the standard semantic meaning of the answer – and therefore makes them arguably better answers – it does not turn all answers into resolving ones. This point is also nicely illustrated with the following example.

[3] This is true, in particular, for the notion of resolving questions introduced by Groenendijk & Stokhof (1984). According to them, a question is resolved if the extension of the question-predicate is fully specified. One may argue that resolvedness should also depend on the information the questioner is interested in when asking her question (see Ginzburg (1995) and van Rooij (2003) for proposals along these lines). However, even if one modifies the notion of resolving questions accordingly, it will not be the case that the exhaustive interpretation of an answer like (2) will always resolve the question.

(3) John or Mary.

On its exhaustive interpretation this answer states that either only John or only Mary exhaust the set of people who passed the examination. Again, in most contexts this information will not fully resolve the question asked.[4] Another point that should be noticed in connection with this example is that its exhaustive interpretation is not completely described by taking the exclusive interpretation of 'or'.[5] One would miss the additional inference of the exhaustive reading that no-one else besides Mary and John passed the examination.

Careful attention should be paid also to the way exhaustive interpretation interacts with the semantics of determiners. Compare, for instance, (4) and (5).

(4) Three students.

(5) At least three students.

The exhaustive interpretation of (4) allows us to conclude that not more than three students passed the examination. (5), however, cannot be read in this way.[6] So there is a difference between (4) and (5) that exhaustive interpretation is sensitive to. However, it will not be adequate to propose that 'at least' simply cancels an exhaustive interpretation.

[4] Sometimes, however, a disjunction can be resolving. Consider, for instance, (i) reading Ann's utterance as a polar question.

(i) Ann: Did Mary or John pass the examination?
Bob: Yes, Mary or John passed the examination.

[5] Under the exclusive interpretation of 'or', 'A or B' is true iff one of the disjuncts is true but not both.

[6] We only discuss here 'at least' as a modifier of the numeral. The occurrence of 'at least' in (5) can also be read as particle, with a syntactical behavior similar to 'even'. This use is not discussed in the present paper. Readers who have problems getting the exhaustive interpretation for (5) should try 'John and at least three of his friends passed the examination'.

(5) can give rise to the inference that nobody besides students passed the examination, and, thus, can show effects of exhaustification.

For (6), just as for (5), we will not infer a limitation on the number of students that passed the examination, if it is interpreted exhaustively.

(6) Students.

Notice that nevertheless it *can* be concluded that, besides students, no one else passed the examination. Thus, also in this case certain effects of exhaustive interpretation are present. In contrast to (6), the exhaustive interpretation of (7) implies additionally that not all students passed the examination. So, again, something distinguishes (6) and (7) with respect to exhaustive interpretation.

(7) Most students.

How can these observations be explained? We will propose in section 5 that exhaustivity is sensitive to the different dynamic semantics of the answers and this leads to the different interpretations.[7]

2.2. The Context-dependence of Exhaustivity

The examples discussed above show how exhaustive inferences change depending on the answer given. Interestingly enough, even the same answer (following the same question form) can give rise to different exhaustive interpretations in different contexts. First of all, it seems that sometimes answers should not be interpreted exhaustively at all. A typical example is the dialogue given in (8).

[7] Some of the inferences attributed in this section to exhaustive interpretation are standardly analyzed as conversational implicatures. This is no accident, as we see it. In section 7 we will discuss the relation between exhaustive interpretation and conversational implicatures in some detail.

(8) Ann: Who has a light?
Bob: John.

Here, Bob's answer is normally not understood as 'John is the only one who has a light'. Instead, it seems that no information other than its semantic meaning is conveyed. We call this interpretation of answers the *mention-some* reading, while we will refer to the one discussed until now as exhaustive interpretation as the *mention-all* reading. It appears that mention-some readings occur precisely in those contexts where the questioner is intuitively not interested in the exact specification of the question-predicate and the semantic meaning of the answer already provides her with all the information she needs.[8]

Aside from the *mention-all* and *mention-some* readings, there also seem to be situations with *intermediate* exhaustive interpretations. In these cases some of the typical inferences of mention-all readings are allowed, but not all of them.

Perhaps the best known limitation is *domain restriction.* There are contexts in which an answer to a question with question-predicate P specifies those and only those individuals that have property P - but only for a *subset* of all objects to which P may apply. Imagine Mr. Smith asking one of his employees:

(9) Mr. Smith: Who was at the meeting yesterday?
Employee: John and Mary.

[8] It is important to distinguish mention-some readings from a different interpretation an answer can get, for instance, if the speaker adds, for instance, 'as far as I know'. In contrast to mention-some readings, in the latter cases the information one receives is not exhausted by the semantic meaning of the answer. Instead it is additionally inferred that the information given in the answer exhausts the *knowledge* of the speaker. Of course, this latter reading can also occur if the questioner is interested in a full specification of the question-predicate. See section 7 for more discussion.

There is a reading of this answer implying that John and Mary are the only *employees* of Mr. Smith who were at the meeting yesterday. There may have been others besides employees of Mr. Smith, but nothing is inferred about them. For the choice of interpretation it seems to be relevant, again, what is commonly known about the information Mr. Smith is interested in. Suppose, for instance, that it is mutually known that Mr. Smith would like to know whether one of his rivals from other companies was at this meeting. Then one would infer from (9) that John and Mary are the only *rivals* of Mr. Smith who were at the meeting yesterday.

Exhaustive interpretation is limited in other ways in so-called *scalar readings* of answers (cf. Hirschberg, 1985). As in the example above, also here exhaustivity seems to apply only to parts of the question-predicate. Imagine Ann and Bob playing poker.

(10) Ann: What cards did you have?
Bob: Two aces.

Here, Ann will interpret Bob's answer as saying that he did not have three aces or two additional kings (a double pair wins over a single one). Still, the answer intuitively leaves open the possibility that Bob additionally had, for example, a seven, a nine, and the king of spades. Just as in the previous case, Ann's interest in information here is different from the case in which an answer gets a mention-all reading. She is not interested in the exact cards that Bob had. She wants to know, however, how *good* (with respect to an ordering relation induced by the poker rules) Bob's cards were. And the scalar reading tells her that Bob did not have additional cards that would raise this value.

To give a final example of a context where the force of exhaustive interpretation seems to change depending on the context, consider (11).

(11) Ann: How far can you jump?
Bob: Five meters.

If it is commonly known that Ann wants to have precise information about Bob's jumping capacities, the exhaustive interpretation of his answer will imply that he cannot jump a centimeter further than 5 meters. If, however, a rough indication is sufficient, one may infer just that he cannot jump 6 meters. This illustrates how – depending on the needs of the questioner – exhaustive interpretation can select the domain of the question-predicate with different degrees of fine-grainedness.

In this subsection we have discussed some examples where an answer does not obtain the strong interpretation that is traditionally associated with the name 'exhaustive interpretation'. Sometimes only parts of the mention-all reading were observed, sometimes nothing was added to the semantic meaning of the answer at all. But in all cases the contextual parameter on which the strength of the exhaustive interpretation depends seems to be what is commonly known to be relevant for the questioner. If the questioner is known not to be interested in certain information, then it will not be provided by the exhaustive interpretation of the answer. For instance, in a typical context where (8) is used it is clear that for the questioner, it is sufficient to know of somebody who has a light that she has a light. This interest is fully satisfied with the semantic meaning of the answer given by Bob. We will take this observation seriously and describe in section 6 exhaustive interpretation as depending on the information the questioner is interested in. It will turn out that in this way we can account for domain restrictions, granularity effects, scalar readings, as well as the mention-some interpretation.

2.3. Other types of questions

The examples discussed so far all were answers to some *wh*-question where the question-predicate is of type $\langle s, \langle e, t\rangle\rangle$. The exhaustive interpretation is, however, not restricted to this class of answers. There are questions of other types whose answers also seem to show exhaustiveness effects. For instance, there is a well-known tendency to interpret conditional answers to polar questions, exemplified in (12), as bi-conditional.

(12) Ann: Will Mary win?
Bob: Yes, if John doesn't realize that she is bluffing.

Thus, one infers that Mary will win just in case John does not realize that she is bluffing. Intuitively, in this reading the same thing is going on as in the cases of exhaustive interpretation discussed so far: the worlds where Mary will win are taken to be exhausted by those where the antecedent of Bob's answer is true. Therefore, it seems reasonable to expect that a convincing approach to exhaustive interpretation should be able to deal with this observation as well.

Various approaches to exhaustive interpretation already exist in the literature. However, the domain of application differs markedly from theory to theory. As far as we know, none of the existing approaches can account for all the observations discussed above. Moreover, none of these theories gives a satisfactory explanation for why the scope of exhaustive interpretation should be restricted to those cases that they can actually handle.

In this paper a unified approach to the exhaustive interpretation of answers is presented which is able to deal with the whole list of examples discussed so far. This account essentially builds on a description of exhaustive interpretation proposed by Groenendijk & Stokhof (1984) (abbreviated by G&S). We will therefore start by discussing their work.

There is one final point that should be made clear. The reader may have noticed that some of the inferences attributed here to exhaustive interpretation are standardly analyzed as conversational implicatures.

As we will argue in section 7, this is to be expected because exhaustive interpretation is by itself a conversational implicature. But then one may wonder what the relevance of this paper is, for we already have Grice' s theory to account for conversational implicatures. However, the well-known problem of this theory is that it does not make clear predictions, and although many people have tried, we are not aware of any fully satisfying formalization of Grice's proposal. Therefore, if we want to use it to account for exhaustive interpretation, we need to provide at least a partial formalization of Grice's theory. This will be the topic of section 7. But to evaluate whether this formalization indeed accounts for exhaustive interpretation, and also as a starting point for theories that do not agree with our opinion that exhaustive interpretation is a conversational implicature, one first needs an adequate description of this rule of interpretation. The sections 4 to 6 of this paper will deal with this issue.

But let us start with discussion the classical approach to exhaustive interpretation of Groenendijk & Stokhof (1984).

3. Groenendijk and Stokhof's proposal

Assume that W is a class of models (possible worlds) for our language and let $[\phi]$ denote the intensional semantic meaning of expression ϕ. Hence, $[\phi]$ is a function mapping elements w of W on the extension of ϕ in w (in case ϕ is a sentence we use $[\phi]^W$ to denote the set of models in W where ϕ is true). Groenendijk & Stokhof (1984) propose to describe the exhaustive interpretation of answers to 'Who'-questions by the operation exh_{GS} taking as arguments the generalized quantifier denoted by the term answer T and the property denoted by the question-predicate P.[9]

DEFINITION 1. *(The exhaustivity operator of Groenendijk & Stokhof)*

$$exh_{GS}([T],[P]) =_{def} \lambda w.[T]([P])(w) \wedge \neg\exists \mathcal{P}' : [T](\mathcal{P}')(w) \wedge \mathcal{P}'(w) \subset [P](w)$$

Set-theoretically, the above formula applied to a generalized quantifier and a property allows the property only to select the *minimal* elements of the generalized quantifier. To illustrate, assume that Bob's response to Ann's question 'Who passed the examination?' is 'John'. Analyzed as a general quantifier 'John' denotes $\lambda w \lambda \mathcal{P}.\mathcal{P}(w)(j)$, which is true of some property $\mathcal{P}$ if in every world $\mathcal{P}$ denotes a set containing j. Applying exh_{GS} to this function turns it into a generalized quantifier that is true of $\mathcal{P}$ if in every world it denotes the *minimal* set containing j, which is the set $\{j\}$. Thus, it is correctly predicted that by exhaustive interpretation we can conclude from the answer 'John' that $\{j\}$ is the set of individuals that passed the examination.

Reading term answers as generalized quantifiers in combination with the exhaustivity operation defined above allows us to account for the interpretation effects observed in examples as (1), (2), (3), (4) and (7) discussed in sections 1 and 2. Actually, G&S can do even more. They show that the above stated operator for terms can be generalized easily to n-ary question-predicates.[10]

Although these results are very appealing, G&S's exhaustivity operator has still been criticized. For instance, Bonomi & Casalegno (1993) have

[9] As mentioned in G&S (1984), this operator has much in common with Szabolcsi's (1981) interpretation rule for 'only'. Though similar in content, G&S's version provides the more transparent formulation.

[10] Their general exhaustivity operator for n-ary terms looks as follows:
$exh^n_{GS}([T_n],[P_n]) = \lambda w.[T_n]([P_n])(w) \wedge \neg\exists \mathcal{P}'_n : [T_n](\mathcal{P}'_n)(w) \wedge \mathcal{P}'_n(w) \subset [P]_n(w)$.

argued that G&S's analysis is rather limited because it can be applied only to noun phrases. To account for examples in which 'only'[11] associates with expressions of another category, they argue that we have to make use of events. We acknowledge that the use of (something like) events might, in the end, be forced upon us. But perhaps not exactly for the reason they suggest. Crucial for G&S's analysis is that (ignoring the intensional parameter) their exhaustivity operator is applied to objects of type $\langle\langle\phi, t\rangle, t\rangle$. It is normally assumed that noun phrases denote generalized quantifiers of type $\langle\langle e, t\rangle, t\rangle$, which means that denotations of noun phrases are in the range of the exhaustivity operator. However, it is also standardly assumed that an expression of *any* type ϕ can be *lifted* to an expression of type $\langle\langle\phi, t\rangle, t\rangle$ without a change of meaning. But this means that – after type-lifting – G&S's exhaustivity operator can be applied to the denotation of expressions of any type, and there is no special need for events.[12]

Although Bonomi & Casalegno's (1993) criticism does not seem to apply, G&S's analysis faces some other limitations. First, it is quite obvious (and has been noticed by themselves) that they cannot account for mention-some readings, domain restriction, granularity effects, and scalar readings (see section 2.2). This is inevitable given the functionality of exh_{GS}. By taking only the semantic meaning of the predicate of the question and the term in the answer as arguments,

[11] They discuss exh_{GS} as a description of the semantic meaning of 'only', but their criticism applies with the same force to exh_{GS} as a description of the exhaustive interpretation of answers.

[12] The reason that we still might need (something like) events is that for questions as 'What did you do last summer?' a possible-world approach may not provide enough fine-structure to properly describe the meaning of the question-predicate, thus, the set of 'things' one did last summer. Making use of events may be one way to achieve this required fine-grainedness. But this is not a problem of G&S's approach to exhaustive interpretation but rather for the general conception of meaning in which this proposal is situated.

exh_{GS} is too rigid to account for differences that can occur involving the same question-predicate and the same answer. The limited functionality of the operation exh_{GS} also seems to be responsible for other problems of the approach (called 'the functionality problem' by Bonomi & Casalegno (1993)). Because G&S assign to (4) 'Three students' and (5) 'At least three students' the same meaning, exh_{GS} predicts for these pairs of answers the same exhaustive interpretation. However, as discussed in section 2.1, intuitively the interpretations differ.[13] Something similar is the case for answers like (3) 'John or Mary' and (13).

(13) John or Mary or both.

Standard semantics takes both answers to be equivalent, but their exhaustive interpretation differs. While (3) implies that John and Mary did not pass the examination, this is not true for (13). Exh_{GS} predicts the exclusive reading in both cases.

The next problem (discussed in G&S, pp. 416-417) concerns the way exh_{GS} operates on its arguments. If we allow for group objects, interpret ['passed the examination'] as a distributive predicate[14] and ['John and Mary'] in (14) as the quantifier $\lambda w.\lambda \mathcal{P}.\mathcal{P}(w)(j \oplus m)$, then G&S predict that on the exhaustive interpretation of (14) ['passed the examination'] denotes the set $\{j \oplus m\}$.

(14) Ann: Who passed the examination?
Bob: John and Mary.

Because ['passed the examination'] is distributive, this cannot be fulfilled in any world: there can be no model w of the language where $j \oplus m \in$ ['passed the examination'](w) but $j \notin$ ['passed the examination'](w).

[13] In both cases the application of exh_{GS} implies that not more than three students passed the examination. This problem has also been noted by G&S themselves.

[14] A predicate P with domain D is distributive in a set of models W if for all $w \in W$, $(\forall x, y \in D)([P](w)(x) \wedge [P](w)(y) \leftrightarrow [P](w)(x \oplus y))$.

Hence, when applying exh_{GS} to (14) Bob's answer is interpreted as the absurd proposition. This is inadequate given that (14) can be interpreted straightforwardly in an exhaustive way.[15]

Finally, negation is a problem for exh_{GS}. Apply, for instance, this operation to Bob's answer in (15).

(15) Ann: Who passed the examination?
Bob: Not John.

Then Bob's response is interpreted as implying that *nobody* passed the examination: the smallest extension of predicate 'passed the examination' such that the answer is true is the empty set. This is clearly not a possible reading for this answer.

The aim of this paper is to overcome the problems discussed above. We claim that this can be done without radically changing the basic idea behind G&S's exhaustivity operator.

What do we understand this basic idea to be? According to G&S, to interpret an answer exhaustively means to minimize the question-predicate of the answer: from the fact that the answerer did not claim that a certain object has property $\mathcal{P}$ it is inferred that the object does not have property $\mathcal{P}$. Thus, the hearer makes the absence of information meaningful. She interprets it as negation. This we take to be an essentially correct perspective on what exhaustive interpretation is about.

[15] There is a solution to this problem, already sketched by G&S, ibid. For independent reasons one is driven to allow the interpreter to choose freely between a distributive and non-distributive reading for predication to plural objects. If one additionally assumes that distributive predicates allow only for the second reading, (14) is interpreted as $\forall x \leq j \oplus m : P(x)$. Minimization of P relative to this answer does not give rise to complications. Later on (section 4.2) we will propose another solution. It has the advantage to carry over to a different kind of problem that the proposal sketched here cannot capture.

However, G&S were not aware of the fact that this reasoning pattern – negation as failure – was starting to get a lot of attention in artificial intelligence as well. It lead (among other things) to the development of a whole new branch of logic: *non-monotonic* logic. When we now try to improve on the proposal of G&S we can build on the work done in this area.[16]

4. Exhaustivity as Predicate Circumscription

4.1. Predicate Circumscription

Only a few years before Groenendijk and Stokhof came up with their description of exhaustive interpretation, McCarthy impressed the artificial intelligence community by introducing *Predicate Circumscription*, one of the first formalisms of non-monotonic logic. McCarthy's goal was to formalize common sense reasoning. More specifically, Predicate Circumscription was intended to solve the *qualification problem*: if we would use classical logic to derive every-day conclusions, we would need an "impracticable and implausible" (McCarthy, 1980, p. 145) number of qualifications in the premises. For instance, if one wants to predict that if we would throw our computers out of our windows, they would smash on Nieuwe Doelenstraat, one would have to specify that gravitation will not stop working, the computers will not develop wings and fly

[16] A question one often hears in this context is 'Do we really need non-monotonic logic?'. Indeed, we do. Non-monotonicity is simply a property of exhaustive interpretation. Therefore, no matter how one describes exhaustive interpretation, it will also be a property of the description. Recall that reasoning is non-monotone if certain inferences might be given up under the presence of more information. It is easy to see that this holds for exhaustive interpretation. From the answer 'John' we can conclude that Mary did not pass the examination. This inference is lost when the speaker also tells us that Mary passed as in (14).

away etc. - in short: nothing extraordinary will happen. The solution McCarthy proposes is to strengthen the inferences one can draw from a theory by adding to the premises a *normality* assumption. It says that nothing abnormal is the case that is not explicitly mentioned in the theory. Or, to restate it somewhat more abstractly, the extension of certain predicates (the abnormality predicates) is restricted to those and only those objects that are explicitly stated by the premises to be in the extension. To come back to the example above, if there is no explicit information about abnormalities in the gravitation of the earth the normality assumption adds the premise that the gravitation is working as normal. Thus, abnormality is negated as failure.

McCarthy (1986) formalizes this idea[17] by defining a syntactic operation on a sentence (the premise) that maps it to a new second-order sentence (the premise plus the normality assumption) in the following way.

DEFINITION 2. *(Predicate Circumscription)*
Let A be a second-order formula and P a predicate of some language $\mathcal{L}$ of predicate logic. Then the circumscription of P relative to A is the formula $\textsc{Circ}(A, P)$ *defined as:*

$$\textsc{Circ}(A, P) := A \wedge \neg\exists P' : A[P'/P] \wedge P' \subset P,$$

where $A[P'/P]$ describes the substitution of all free occurrences of P in A by P'.

Looking at this formalization of Predicate Circumscription, our reader will immediately recognize the following striking fact: G&S's exhaustivity operation is – roughly speaking – just an instantiation of McCarthy's predicate circumscription! The circumscribed predicate is now the question-predicate, and the circumscription is relative to the sentence one gets by combining term-answer and question-predicate - or

[17] This is a simplified version of his formalization.

simply the sentential answer. This important parallelism was first noticed, as far as we know, by Johan van Benthem (1989).[18]

Predicate circumscription has a model-theoretic pendant: interpretation in *minimal models*. First, the model-theory for classical logic is enriched by defining an order on the set of models W: a model v is said to be more minimal than a model w with respect to some predicate P, $v <_P w$, in case they agree on everything except the interpretation they assign to P and it holds that $[P](v) \subset [P](w)$. It can be shown that if W is the class of all models the P-minimal models of a theory A, hence the set $\{w \in [A]^W | \neg\exists v \in [A]^W : v <_P w\}$, are exactly the models where the circumscription formula $\textsc{Circ}(A, P)$ holds.[19]

[18] There are certain differences between exh_{GS} and $\textsc{Circ}$ that should be mentioned. First, G&S took exh_{GS} to be a description of an operation on semantic representations while $\textsc{Circ}(A, P)$ is an expression in the object language. Second, $\textsc{Circ}$ takes as arguments a predicate and a sentence, while exh_{GS} applies to the predicate and the sentence without the predicate. $\textsc{Circ}$, therefore, relies on less syntactic information. But, as Ede Zimmermann pointed out to us, it looks as if there are cases where exhaustive interpretation relies on this information. Consider, for instance, the answer 'Men that wear a hat' to a question 'Who wears a hat?', where the question-predicate P appears in the term answer part T. The circumscription of $A = T(P)$ w.r.t. P minimizes P in all occurrences of A and interprets the answer as implying that nobody wears a hat – which is obviously wrong. Exh_{GS} only minimizes occurrences of P outside T and correctly predicts that exactly those people wear a hat that are men that wear a hat. In this paper we will assume that the question-predicate does not occur in the term answer part.

[19] This set of minimal models can be described relative to a set of alternatives of A, $Alt(A)$, as well. If we say that $v <_{Alt(A)} w$ if and only if v is exactly like w except that $\{B \in Alt(A) | v \in [B]^W\} \subset \{B \in Alt(A) | w \in [B]^W\}$, we can define the following set of minimal models: $\{w \in [A]^W | \neg\exists v \in [A]^W : v <_{Alt(A)} w\}$. This set is the same as the one described in the main text if we define $Alt(A)$ as follows: $\{P(a) | d \in D \ \&\ a \text{ is the name of } d\}$, and assume that every individual has a unique name. A similar notion of alternatives is used in alternative-semantics approaches

This formulation of predicate circumscription – as interpretation in minimal models – is not a stranger to linguists. The Lewis/Stalnaker approach to counterfactuals (see, for instance, Lewis (1973)) also makes use of it. The application at hand differs mainly in the way the *order* is defined.

4.2. The basic setting

It is this later, model-theoretic formulation that we will use to describe exhaustive interpretation. Here comes our basic definition.

DEFINITION 3. *(Exhaustive interpretation - the basic case)*
Let A be an answer given to a question with question-predicate P in context W. We define the exhaustive interpretation $exh^W_{std}(A, P)$ of A with respect to P and W as follows:

$$exh^W_{std}(A, P) \equiv \{w \in [A]^W | \neg\exists v \in [A]^W : v <_P w\}$$

To illustrate the working of this interpretation function, let us go back to example (12) here repeated as (16).

(16) Ann: Will Mary win?
Bob: Yes, if John doesn't realize that she is bluffing.

In this case the question-predicate P = 'Mary will win' is of arity 0.[20] But this means that $v <_P w$ iff v is exactly like w, except that whereas w makes P true, v makes it false. Now it can be checked that $exh^W_{std}(A \rightarrow P, P)$ is true only in those worlds where either both A and P are true, or both A and P are false. Worlds where A is false and P

to the meaning of 'only' (e.g. Rooth (1996)). For more discussion see van Rooij & Schulz (to appear).

[20] We assume that the extension of an n-ary predicate P^n in world w is the set of n-ary tuples that verifies sentence $P^n(\vec{x})$ in w. If P^0 is true in w, it denotes $\{\langle\rangle\}$, otherwise $\emptyset$.

true are ruled out because there are other worlds that verify $A \rightarrow P$, but do not make P true (worlds where both A and P are false) and, hence, are more minimal. The possibility that A is true and P is false is excluded by the semantic meaning of the answer. Thus, by applying exh_{std} the conditional answer gets the desired bi-conditional reading.[21]

The change from G&S's approach to the one given in definition 3 is rather subtle – mainly one of perspective. But, as we will see in the rest of the paper, this model-theoretic description of exhaustive interpretation proves to easily admit the amendments we have to make to deal with the limitations of G&S's approach. It also allows us to improve on exh_{GS} directly. Remember our earlier discussion of applying the rule of exhaustive interpretation to distributive predicates (enriching the domain with group objects). We discussed it using our very first example, repeated here as (17).

(17) Ann: Who passed the examination?
Bob: John and Mary.

Let us calculate once more the exhaustive interpretation of Bob's answer, but now using exh_{std}. Again, Bob is taken to be talking about a plural object $j \oplus m$. To determine $exh^W_{std}(P(j \oplus m), P)$ we first eliminate all worlds where Bob's answer is false. Then, we select those worlds where the extension of the question-predicate P is minimal. At first one may think that these are the worlds where the extension of 'passed the examination' contains only $j \oplus m$. However, such worlds do not exist. The predicate is distributive and already G&S account for this by letting meaning postulates impose restrictions on the class of proper models. But then, the smallest extensions P can receive in worlds where Bob's answer is true are such that besides the plural object $j \oplus m$ also j and m are in the extension of question-predicate 'passed the examination'. Thus, we obtain the right result.

[21] This prediction is, of course, already made by G&S's operation exh_{GS}.

But why can we solve this problem just by taking exh_{std} instead of exh_{GS}? Did we not claim above that $exh_{GS}([T],[P])$ is roughly the same as $\textsc{Circ}(T(P),P)$ and the latter (more particularly $[\textsc{Circ}(T(P),P]^W)$ is equivalent to $exh^W_{std}(T(P),P)$? Well, one has to be careful. Remember that the latter equivalence only holds if W is the class of *all* models. Meaning postulates impose restrictions on W. Exh_{std} is sensitive to these restrictions because they influence the set of possible worlds it quantifies over. Exh_{GS}, however, quantifies locally over alternative extensions for the question-predicate. It does not check whether these alternatives are realized in some world. Only if the meaning postulates are taken to be part of the answer, exh_{GS} and $\textsc{Circ}$ are forced to respect them and predict correctly.

To sum up, distributive predicates show that circumscribing just the answer may not be enough. The exhaustive interpretation is sensitive to information available in the context set W, in particular to meaning postulates. Because exh_{std} quantifies over W it can immediately account for this dependence.[22]

Actually, there are even more striking examples in favor of a notion of exhaustivity which respects meaning postulates and they do not rely on particular premises such as the group analysis of Bob's answer in (17). For instance, the proposed formalization also allows us to account for some puzzles connected with the meaning of 'only'. For limitations of space, however, we cannot discuss this issue in detail here.[23]

[22] The variable W makes our interpretation function very context dependent. If W is understood as the respective common ground then all the information presented there will influence what counts as a minimal model in a particular context. It still has to be seen to what extent exhaustive interpretation is context sensitive in this sense. See also the discussion in section 7.

[23] One particularly famous example that this approach can account for is the following from Kratzer (1989).

(i) Bob: I only [painted a still-life]$_F$.

5. Exhaustivity and dynamic semantics

Another problem of G&S's approach that we discussed in section 2.1 is that it makes incorrect predictions for answers like (5) and (6), here repeated as (18b) and (18c).

(18) (a) Three Students.
(b) At least three students.
(c) Students.
(d) Most students.

As we pointed out earlier, it is standardly assumed in generalized quantifier theory (adopted by G&S) that 'three students' has the same semantic meaning as 'at least three students'. Because the operation exh_{GS} (the same is true for exh_{std}) takes only the semantic meaning of the answer and the question-predicate into account, it predicts for (18a) and (18b) the same readings. However, the exhaustive interpretation of the first answer gives rise, intuitively, to an *at most* inference, while the exhaustive interpretation of the latter does not. Something similar has been observed comparing (18c) and (18d). Thus, there is a difference between these answers exhaustive interpretation is sensitive to which exh_{GS} (and exh_{std} as well) fails to observe.

Different perspectives are possible on this dilemma. An interesting proposal is made by Zeevat (1994), who incorporates the *at most* inference (18d) comes with in the semantics of 'most'. In this paper, however, we stick to the traditional analysis of this determiner.[24] Others have proposed that expressions containing 'at least' or bare nominals should not be interpreted exhaustively. However, as observed in section 2.1, also for these expressions we observe *some* exhaustivity effect.

(ii) Lunatic: No. You also [painted apples]$_F$.

For a closer discussion see van Rooij & Schulz (to appear).

[24] Still, our final explanation will have some similarity with Zeevat's proposal.

Hence, total absence of exhaustification is no option. We will propose instead that exhaustive interpretation *does* take place but that it will not give rise to the *at most* inference.

There is a difference between, for instance, (18c) and (18d) that can be made responsible for their unequal exhaustive meanings. But – or so we propose – it is a difference in their semantic meaning.[25] The answers diverge in their dynamic discourse contribution. In consequence, to be able to make the correct predictions we have to adopt a dynamic perspective on semantics and describe exhaustive interpretation as an operation that is sensitive to dynamic information.

We will not introduce full-blooded dynamic semantics but restrict ourselves to some of its essential features, leaving the exact implementation to the reader's favorite dynamic theory. We assume a dynamic interpretation function that maps an information state σ and a sentence ϕ to the new information state $\sigma[\phi]$. An information state is a set of possibilities, i.e., a set of world-assigment pairs. Discourse referents are interpreted as fixed variables of the assignments. The definition of the order $<_P$ comparing the extensions of the question-predicate is extended to the case of possibilities by adding the condition that the assignments have to be identical to make possibilities comparable.[26] Dynamic exhaustive interpretation is then defined as a context change function that selects minimal possibilities instead of worlds.

DEFINITION 4. *(Dynamic Exhaustive Interpretation)*
Let A be an answer given to a question with question-predicate P in context σ. We define the exhaustive interpretation $exh^{\sigma}_{dyn}(A, P)$ of A with respect to P and σ as follows:

$$exh^{\sigma}_{dyn}(A, P) \equiv \{i \in \sigma[A] | \neg\exists i' \in \sigma[A] : i' <_P i\}.$$

[25] Thus, in this case it is not the functionality of $exh_{G\&S}$ (or exh_{std}) that causes the mispredictions, but the semantic analysis of the determiners G&S adopt.

[26] Thus, we redefine $\langle w, g\rangle <_P \langle v, h\rangle$ iff$_{\text{def}}$ $g = h$ and $w <_P v$.

How does this straightforward extension of exh_{std} to dynamic semantics solve the problems discussed above? The crucial point is that the interpretation of introduced discourse referents becomes a fixed property of our possibilities. These variables can no longer be varied freely when the extension of the question-predicate is minimized. That makes it more difficult for possibilities to be minimal.

This will become much clearer after discussing some examples. First, let us consider the answers (18c), 'Students', and (18d), 'Most students'. It has been argued that the determiners occurring in these answers belong to different classes. While the first (together with 'A man', 'Some$_1$ girls', 'Five girls' and 'At least five girls') contains a weak determiner, the determiner of the second (together with 'all ducks', 'most students', and 'some$_2$ girls') is strong. Adopting a standard assumption of dynamic semantics (e.g. Kamp & Reyle, 1993), we treat only the latter type of NPs as two-place generalized quantifiers. NPs with weak determiners, in contrast, do not denote generalized quantifiers and directly introduce discourse referents. For anaphoric reference to strong quantifiers, discourse referents have to be constructed afterwards from the intersection of nucleus and restrictor. It turns out that if we adopt this treatment of weak and strong quantifiers the new function exh_{dyn} can account for the differences in the exhaustive interpretations.

Assume an information state: $\sigma = \{i_1, i_2, ..., i_8\}$, where $i_k = \langle w_k, g_k\rangle$. In the worlds of all possibilities we have the same interpretation for 'students', the set $\{a \oplus b, a \oplus c, b \oplus c, a \oplus b \oplus c\}$. For the interpretation of 'passed the examination' we assume: $[P](w_1) = \{a, b, c, a \oplus b, ...\}, [P](w_2) = \{a, b, a \oplus b\}, ..., [P](w_4) = \{b, c, b \oplus c\}, ..., [P](w_8) = \emptyset$ - we simply take every possible distributive set given the three atoms $\{a, b, c\}$. Hence, predicate P is assumed to be distributive. Furthermore, notice that only in w_1 but not in w_2, w_3, and w_4 it is true that all students passed the examination. First, we calculate the exhaustive interpretation of answer (18c). After updating with the semantic meaning of the sentence ‘Students passed the examination’, $\exists X : S(X) \wedge P(X)$,[27] we end up with an information state σ' containing successors of the possibilities i_1, i_2, i_3, and i_4 whose variable assignments now are defined for X[28] , the newly introduced discourse referent.[29] In σ' there will be a possibility for every possible mapping of X to a group of students that passed the examination in one of the worlds w_1, w_2, w_3, and w_4. So σ' contains, for instance, the possibility $\langle w_1, X : a \oplus b\rangle$, because the object $a \oplus b$ is in the extension of P in w_1. However, $\langle w_2, X : a \oplus b \oplus c\rangle$ will not be an element of σ'. Given the assignment $X : a \oplus b \oplus c$, answer (18c) would not be true in w_2. The following tableau lists all possibilities in σ' plus the way they are ordered by $<_P$, where $\langle ...\rangle_1 \rightarrow \langle ...\rangle_2$ means that the second possibility is P-smaller than the first.

$$\begin{tikzcd}
\boxed{\langle w_1, X : a \oplus b \oplus c\rangle} & \langle w_1, X : a \oplus b\rangle \arrow[d] & \langle w_1, X : a \oplus c\rangle \arrow[d] & \langle w_1, X : b \oplus c\rangle \arrow[d] \\
 & \boxed{\langle w_2, X : a \oplus b\rangle} & \boxed{\langle w_3, X : a \oplus c\rangle} & \boxed{\langle w_4, X : b \oplus c\rangle}
\end{tikzcd}$$

To determine the exhaustive interpretation of answer (18c) we collect the minimal elements of this ordering (marked by a box in the picture) and obtain the set: $\{\langle w_1, X : a \oplus b \oplus c\rangle, \langle w_2, X : a \oplus b\rangle, \langle w_3, X : a \oplus c\rangle, \langle w_4, X : b \oplus c\rangle\}$. This interpretation still allows for the possibility that *all* students passed the examination - even though it would be excluded that anybody other than students passed the examination. The reason is that after updating σ with $exh_{dyn}(\exists X : S(X) \wedge P(X), P)$ there is still a possibility that takes the world to be w_1: the possibility

[27] We take a standard approach to the dynamic meaning of ‘$\exists$’ and interpret it as introducing a new discourse referent for the variable it binds.

[28] This suggests that we adopt a particular perspective on dynamic semantics where new discourse referents extend assignment functions. However, eliminative approaches to dynamic semantics work as well.

[29] The others are excluded by the truth conditions of the answer.

$\langle w_1, X \rightarrow a \oplus b \oplus c\rangle$. And this is so because there will be no possibility in $\sigma[\exists X : S(X) \wedge P(X)]$ where the extension of P is smaller than in w_1 *and* which still maps X to $a \oplus b \oplus c$. Such a possibility would not make the answer true. Hence, we correctly predict no *at most* inference for the answer (18c).[30]

However, doing the same calculation with 'Most students', the example (18d), will lead to a different result. Because strong determiners do not immediately introduce discourse referents, we obtain as the semantic meaning of the answer in context σ the set $\{i_1, i_2, i_3, i_4\}$ (in all possibilities of σ it is true that most students passed). But i_2, i_3, i_4 are all $<_P$-smaller than i_1. Thus, after exhaustive interpretation we end up with a new information state containing only i_2, i_3 and i_4. The possibility that all students passed the examination is excluded.

Dynamic semantics also helps to account for the difference in exhaustive interpretation of (18a) 'Three students passed' and (18b) 'At least three students passed' (or answers like 'Three or more students passed'). Within dynamic frameworks (e.g. Kamp & Reyle, 1993) it is standard to represent (18a) as $\exists X : S(X) \wedge card(X) = 3 \wedge P(X)$. This formula has the same 'at least three' truth conditions that we obtain with the classical generalized quantifier interpretation of numerals. In particular, this sentence is true if four students passed, because then there is still a set of three students that passed. Thus, from a truth-conditional perspective we could have represented the semantic meaning of (18a) as well by $\exists X : S(X) \wedge card(X) \geq 3 \wedge P(X)$. Dynamically, however, the two formulas are not equivalent: the former introduces discourse referents that denote groups of exactly three individuals, while the groups introduced by the latter formula might be larger. As a consequence, if

[30] Notice, by the way, that after exhaustifive interpretation, if we refer back to the newly introduced discourse referent, we are talking about all students that passed the examination. This is on a par with intuition.

we apply exh_{dyn}, the former formula gets the 'exactly three' reading, while the latter does not. This suggests that the former one correctly represents (18a), while the latter formula is the natural representation of answer (18b). And indeed, that was proposed by Kadmon (1985) (for related, but still somewhat different reasons). Hence, adopting Kadmon's analysis of the two determiners 'three' and 'at least three' allows us to account for their different behavior under exhaustive interpretation.[31]

To sum up the discussion in this section so far: the behavior of determiners is not a problem that forces us to give up the circumscription account for exhaustive interpretation or to propose that certain determiners have to come with special cancellation properties with respect to this mode of interpretation. It suffices to make the description sensitive to dynamic information.[32]

[31] Kamp & Reyle (1993), in fact, would not represent a sentence like (18b) by $\exists X : card(X) \geq 3 \wedge S(X) \wedge P(X)$, but rather by $\exists X : card(X) \geq 3 \wedge X = \lambda y[S(y) \wedge P(y)]$. For our purposes, however, this does not matter. They still predict that (18b) directly introduces a discourse referent and that is all we need for our analysis to go through.

[32] One may argue that free focus is generally interpreted exhaustively. However, certain examples suggest that in so-called topic-focus constructions, or sentences with a hat-contour, the focal-part should not be read exhaustively, even if it is used as an immediate response to a question. Consider (ii) and (iii) as answers to question (i).

(i) What did the boys eat?

(ii) [Some boys]$_T$ ate [broccoli]$_F$.

(iii) [One boy]$_T$ ate [broccoli]$_F$.

If we would interpret 'broccoli' exhaustively, and 'some boys' or 'one boy' as the generalized quantifier 'at least some/one boy(s)', it would mean for (iii) that for all alternatives x distinct to broccoli, the sentence '(At least) one boy ate x' has to be false. But this gives the wrong result that from (iii) we can conclude that *none* of the boys ate anything other than broccoli. As it turns out, also this problem disappears once we adopt a dynamic perspective. We interpret (iii), for instance, as $exh^{\sigma}_{dyn}(\exists X[Boy(X) \wedge card(X) = 1 \wedge Ate(X, Broccoli)], \lambda y.Ate(X, y))$. Sentence (iii)

In the last part of this section we will discuss how dynamic information may also help to solve another part of the functionality problem of exh_{GS}. Remember example (13) 'John or Mary or both'. The application of exh_{GS} to this answer (the same hold for exh_{std}) excludes the last disjunct, hence, predicts that *either John or Mary is the only one who passed the examination.* But even though this answer can be interpreted exhaustively (implying that nobody besides John or Mary passed the examination) the possibility that both of them passed should not be excluded. Similarly, the sentence 'John owns 3 or 5 cars' is on G&S's analysis (and by exh_{std} as well) falsely predicted to mean that John owns exactly 3 cars (the question is 'How many cars does John own?' and we assume an *at least* interpretation of numerals). What we would like to end up with, however, is the prediction that John owns either exactly 3 cars, or exactly 5.

Intuitively, what both operations $exh_{G\&S}$ and exh_{std} miss seems to be that in exhaustively interpreting an answer we are not allowed to exclude any possibility explicitly mentioned in the answer.[33] We can account for this using exactly the same strategy as for the closely related problem concerning determiners. One can simply propose that

is now interpreted as stating that one boy ate broccoli, and that this one boy has eaten nothing else. We correctly predict that it is still possible that non-members of the denotation of the discourse referent X ate something other than broccoli, e.g. beans. Thus, we predict that examples (ii) and (iii) do not provide good arguments against an exhaustive interpretation of free focus.

[33] This was also the basic idea behind Gazdar's (1979) solution for this problem. He was not addressing the exhaustive interpretation of answers but analyzed the exclusive interpretation of 'or' as scalar implicature. To account for the cancellation of this implicature in a sentence like 'John or Mary or both passed the examination' Gazdar proposes that a disjunctive sentence additionally triggers the clausal implicatures that each of its disjuncts is considered possible. If the clausal and scalar implicatures of a sentence contradict each other – as is the case in the example at hand – clausal implicatures overrule scalar ones.

while 'John or Mary passed the examination' and 'John or Mary or both passed the examination' have the same truth conditions, their dynamic semantic meanings are, again, different. Maria Aloni (2003), for instance, argues for independent reasons that the first sentence should be represent by something like $\exists q: {}^{\vee}q \wedge (q = {}^{\wedge}P(j) \vee q = {}^{\wedge}P(m))$, where '$q$' is a propositional variable and '${}^{\vee}$' and '${}^{\wedge}$' have their usual Montagovian meanings. Notice that this formula has the same truth conditions as the standard representation of the sentence: $P(j) \vee P(m)$. Following Aloni's lead, we should then, of course, represent 'John or Mary or both passed the examination' by $\exists q: {}^{\vee}q \wedge (q = {}^{\wedge}P(j) \vee q = {}^{\wedge}P(m) \vee q = {}^{\wedge}(P(j) \wedge P(m)))$, which also gives rise to the same truth conditions. Still, with a dynamic interpretation of the existential quantifier the dynamic semantic meanings of the two formulas differ, because the latter allows for a verifying world-assignment pair where the assignment maps q to the proposition that both John and Mary passed the examination, while the former formula does not.[34] In almost exactly the same way as for the examples (18a) and (18b), this difference in dynamic semantic meaning has the effect that the two formulas give rise to different exhaustive interpretations: the former, $exh^{\sigma}_{dyn}(\exists q: {}^{\vee}q \wedge (q = {}^{\wedge}P(j) \vee q = {}^{\wedge}P(m)), P)$, allows only for possibilities (and thus worlds) in which either only John or only Mary passed the examination; the latter, $exh^{\sigma}_{dyn}(\exists q: {}^{\vee}q \wedge (q = {}^{\wedge}P(j) \vee q = {}^{\wedge}P(m) \vee q = {}^{\wedge}(P(j) \wedge P(m))), P)$, allows for possibilities

[34] Philippe Schlenker (p.c.) came up with a direct 'anaphoric' argument for why sentences of the form 'A or B' and 'A or B or both' indeed should have different dynamic semantic meanings:

(i) We'll invite John or Bill, and *he*'ll have a good time.

(ii) *We'll invite John or Bill or both, and *he*'ll have a good time.

(iii) We'll invite John or Bill or both, and *they*'ll have a good time.

These sentences suggest that the first conjunct of (i), for instance, should be represented by $\exists x: Inv(x) \wedge (x = j \vee x = b)$ rather than by $\exists q: {}^{\vee}q \wedge (q = {}^{\wedge}Inv(j) \vee q = {}^{\wedge}Inv(b))$. But this does not make any difference for our explanation.

where both passed the examination. In a similar manner we can account for the exactly-reading of 'John owns 3 or 5 cars', if we represent it by $\exists q : {}^{\vee}q \wedge (q = {}^{\wedge}[\text{'John owns 3 cars'}] \vee q = {}^{\wedge}[\text{'John owns 5 cars'}])$.

6. Exhaustivity and Relevance

One problem of G&S's approach that our operation exh_{dyn} still inherits is that it cannot account for the contextual dependence of exhaustive interpretation we have observed in domain restricted exhaustive interpretations, the scalar readings, the mention-some readings, and differences in the fine-grainedness of the interpretation (see section 2.2). The crucial observations made when discussing these readings were that (i) in all these cases exhaustive interpretation was not substituted by some other interpretation but simply weakened[35], and (ii) this weakening can be characterized as follows: inferences of the strong fine-grained mention-all reading of exhaustive interpretation disappear if they are commonly known in the context of utterance to be *irrelevant* for the questioner. This leads us to adopt the following strategy towards these readings: we extend our definition of exhaustive interpretation by making it dependent on what counts as relevant for the questioner. As it turns out, we can then correctly describe the intended variation in the strength of exhaustive interpretation.

Let us start with trying to understand what it means to be relevant information for the questioner and how it may play a role for the exhaustive interpretation of answers. If somebody poses a question, she is (normally) in need of certain information. A simple standard way

[35] Thus, in contrast to other analyses we propose for mention-some readings that exhaustivity is not absent in these cases, but that it does not do anything.

to describe this information is by a set DP of propositions.[36] For the questioner it is relevant to know which of these propositions actually hold. The semantic meaning Q of a question is also standardly described as a set of propositions, the appropriate, complete, or resolving answers to the question (see Hamblin (1973), Karttunen (1977), G&S (1984)). It seems rational to assume that for reasons of efficiency there might be a difference between the information asked for explicitly by the questioner and the information needed, described by DP. For instance, assume that Ann is interested in who of John, Mary, and Peter passed the examination.[37] The question directly corresponding to this DP is 'Who of John, Mary and Peter passed the examination?'. But it is arguably better for Ann to ask 'Who passed the examination?' A complete answer to this question would provide her with more information than she needs, but that does not bother her. However, the second question is shorter and thus spares her effort.

If it is commonly known what counts as relevant for the questioner, it would be reasonable for the answerer Bob to take this information into account as well and exhaustively specify only this part of the syntactic question-predicate that is relevant. Instead of listing all individuals that passed the examination, he only mentions whom of John, Mary, and Peter did. This spares *him* effort. Then, of course, a rational hearer will respect this factor as well when interpreting Bob's utterance and will not conclude from the answer 'John' that John was the only individual that passed the examination, but rather that he was the only one of John, Mary, and Peter who did so. And this is exactly what seems to be going on in the case of domain restricted exhaustive interpretation.

Before we can come to a general formalization of this relevance-dependence

[36] This is a simplification of the notion of a *decision problem*.

[37] Thus, DP in this case is the set $\{\{v \in W | [P](w) \cap \{j,m,p\} = [P](v) \cap \{j,m,p\}\} | w \in W\}$.

of exhaustive interpretation, one further question has to be addressed: Does relevance already affect the interpretation of the question (thus, does Ann's question 'Who passed the examination?' semantically mean 'Who of John, Mary and Peter passed the examination?') or is relevance independent contextual information? In the first case the description of exhaustive interpretation we have given can easily be made sensitive to relevance by proposing that the operation does not work on the syntactic question-predicate but rather on the predicate that the question is really about. The only thing that we have to do is to clarify how this predicate can be calculated given the semantic meaning of the question. If, however, the semantic meaning of the question is not affected by what is known about DP, then we cannot use this shortcut and have to incorporate relevance as a fourth argument into our definition of exhaustive interpretation.

This paper is not about the semantics of questions. Therefore, we will not make a decision on this subject and shortly sketch how one can proceed in both cases distinguished above.

6.1. The indirect approach

There are certain arguments that speak in favor of taking the semantic meaning of questions to depend on relevance. For instance, it seems that this factor also influences the interpretation of embedded questions. An extensive discussion of the pros and cons on this issue can be found in van Rooij (2003), followed by an approach to the meaning of questions that takes relevance into account. Also according to this approach the meaning of a question is a set of propositions – but now their semantics is underspecified with respect to contextual relevance. We will adopt this proposal here.

Given this position towards the semantics of questions we can make exh_{std}[38] dependent on relevance simply by manipulating its arguments. We propose that it does not apply to the syntactic predicate of the question asked but to that predicate whose extension the questioner is really interested in. We define it to be a minimal property $\mathcal{X}$ such that knowing the extension of $\mathcal{X}$ would resolve Q, whereby Q is the semantic meaning of the question asked. We say that $\mathcal{X}$ is at least as minimal as $\mathcal{Y}$, $\mathcal{X} \subseteq \mathcal{Y}$, iff $\forall v : \mathcal{X}(v) \subseteq \mathcal{Y}(v)$. Following G&S, we take 'knowing $\mathcal{X}$' to mean knowing which of the following propositions is true: $Q_{\mathcal{X}} = \{\{v \in W | \mathcal{X}(v) = \mathcal{X}(w)\} | w \in W\}$. Thus, someone knows $\mathcal{X}$ if she can specify for every object whether it has property $\mathcal{X}$ or not. 'Knowing the extension of $\mathcal{X}$ resolves Q' is now understood as the following relation between $Q_{\mathcal{X}}$ and Q: $\forall q \in Q \exists q' \in Q_{\mathcal{X}} : q' \subseteq q$, or informally, knowing the extension of $\mathcal{X}$ has to imply knowing which elements of Q are true.

With this definition of the property the question is about we can correctly account for the different readings of exhaustive interpretation distinguished in section 2.2. For instance, if the whole extension of the syntactic predicate P of the question is relevant, van Rooij (2003) predicts that the meaning of the question is exactly such a partition $Q_{[P]}$ as defined above. In this case, no smaller property $\mathcal{X}$ than $[P]$ itself will exist such that $Q_{\mathcal{X}}$ solves $Q_{[P]}$. Hence $\mathcal{X} = [P]$ and the speaker will by exhaustively specifying $\mathcal{X}$ specify $[P]$. Thus, we predict a mention-all reading.

How to account for exhaustive interpretation when domain restriction or the level of required granularity is at issue is straightforward. For a degree-question like (11) 'How far can you jump?', for instance, the question-predicate can in some contexts range over meters, rather

[38] To avoid unnecessary complications, we continue with exh_{std} in this section. However, the changes that will be proposed for this operation can be easily applied to exh_{dyn} as well.

than centimeters. We therefore address scalar readings next. Remember the poker game example, (10) ‘What cards did you have?’. Again there exists a uniquely determined minimal $\mathcal{X}$, i.e., the $\mathcal{X}$ that contains in every world those and only those cards that contribute to the value of the syntactic question-predicate P in terms of the card game. Exhaustifying this set tells us that the speaker had no additional cards that would increase the value of the cards she mentioned explicitly. Because in every world w $\mathcal{X}(w)$ is a subset of $[P](w)$, she may have had other, irrelevant, cards. About them, nothing can be inferred.

Finally, the mention-some case. Here, intuitively, the questioner does not care about what she learns about the extension of the syntactic predicate P, as long as she learns for one thing in its extension that it has property $[P]$. Q is predicted by van Rooij (2003) to be the set $\{\{w \in W | d \in [P](w)\} | d \in D\}$. Any predicate that in each world w applies to exactly one object in $[P](w)$ and nothing else will qualify as a property $\mathcal{X}$ that Q is about. In this case, $\mathcal{X}$ is not uniquely defined. But for the interpreter the choice does not matter. In any case one learns from the answer that some subset of the extension of P consists exactly of the things mentioned in the answer - nothing more and nothing less. But that’s exactly what one wants for an answer as in (8).

6.2. The direct approach

Assume that the semantic meaning of the question does not depend on relevance. How, then, can relevance influence the exhaustive interpretation of answers? A solution to this problem that still keeps the principal setting of our approach the same is to propose that the order $\leq_P$, on which the selection of minimal models is based, depends on relevance. In some sense this is what we have done in the indirect approach as well. We proposed that the predicate, or property, used in exh_{std} should be one that is partly defined in terms of relevance.

Because the order $<_P$ compares worlds with respect to the extension they assign to this predicate, the order becomes sensitive to relevance as well. Now, we have to change the definition of the order to make it directly dependent on some measure of relevance, not just on P.

We say that a world w_1 is more minimal than a world w_2, $w_1 <_P^{rel} w_2$, if (everything else being equal) the proposition claiming that all the objects in $[P](w_1)$ indeed have property $[P]$ is less relevant than the proposition saying the same for world w_2.

DEFINITION 5. *(The relevance order)*

$$w_1 <_P^{rel} w_2 \ \ \textit{iff}_{def} \ \lambda w.[P](w_1) \subseteq [P](w) <_R \lambda w.[P](w_2) \subseteq [P](w)$$

By substituting this new order in the definition of exh_{std} we obtain a relevance dependent description of exhaustive interpretation.

DEFINITION 6. *(Relevance Exhaustive Interpretation)*
Let A be an answer given to a question with question-predicate P in context W. We define the exhaustive interpretation $exh_{rel}^W(A,P)$ of A with respect to P and W as follows:

$$exh_{rel}^W(A,P) \equiv \{w \in [A]^W | \neg\exists v \in [A]^W : v <_P^{rel} w\}$$

This leaves us with the problem to define the order $\leq_R$ comparing the relevance of propositions. Fortunately, a lot of work has been done on this topic in decision theory that we can make use of. The account we propose here simplifies this work quite a bit, for we do not need full-blooded decision theory for our concerns. Remember that we described the information a questioner is interested in by a set of propositions DP. The questioner wants to know which one of these propositions is true. We define the *utility value* of a proposition p by how much it helps to select one of these propositions in DP as true. Let $\mathcal{P}(q|p)$ be the quotient $\frac{card([q]^W)}{card([p]^W)}$. Hence, $\mathcal{P}$ is a simplified measure of the probability of q given that p is true.[39]

DEFINITION 7. *(The utility value)*

$$UV(p) = max_{q \in DP}\mathcal{P}(q|p) - max_{q \in DP}\mathcal{P}(q|W)$$

The order $<_R$ on propositions can then be defined by comparing this utility value.[40]

Now everything is in place and we can start to test the proposal. Assume that the speaker wants to know what exactly the extension of the syntactic question-predicate P is. DP is then the set of propositions that exactly specifies the extension of P. But this means that $v <_P^{rel} w$ iff, everything else being equal, $[P](v) \subset [P](w)$ and, thus, iff $v <_P w$. Hence, exh^W_{rel} reduces to exh^W_{std} and we predict a mention-all reading.

Mention-some answers can be treated as well. As already discussed in section 2.2, answers get mention-some or non-exhaustive interpretations in cases where it is clear that the addressee only has to know for someone in the extension of P that she has property P. For (8) 'Who has a light?', for instance, it is normally enough for Ann to know someone who has a light. She just wants to know who to ask for lightning her cigarette. Let us discuss a concrete example. Assume that $D = \{j, m\}$, $W = \{w_1, w_2, w_3\}$, and $[P](w_1) = \{j\}, [P](w_2) = \{m\}, [P](w_3) =$

[39] It is perhaps useful to point out that if DP is a singleton set consisting only of a 'goal' proposition h, the utility value we assign to a proposition comes down – according to definition 7 – to the standard statistical notion of relevance and is also very similar to what Merin (1999) defines as the relevance of this proposition. The notion given in definition 7 has also been used by Parikh (1992) and van Rooij (2003) for linguistic purposes.

[40] Notice that in case the reverse of $<_R$ is one-sided entailment (see van Rooij (2004) for a discussion under which circumstances this will be so), it will be the case that $w_1 <_P^{rel} w_2$ exactly if $[P](w_1) \subset [P](w_2)$. Thus, in these circumstances $w_1 <_P^{rel} w_2$ if and only if $w_1 <_P w_2$: the old ordering between worlds is a natural special case of our new ordering. It follows that in these cases $exh^W_{rel}(A, P) = exh^W_{std}(A, P)$.

$\{j, m\}$. Given what we have said about the questioner, DP would in this situation be the set: $\{\{w_1, w_3\}, \{w_2, w_3\}\}$. To determine the order $<_P^{rel}$ we first have to calculate the utility values of the propositions $\lambda w.[P](w_i) \subseteq [P](w)$ for $i = 1, 2, 3$. It turns out that $UV(\lambda w.[P](w_1) \subseteq [P](w)) = UV(\lambda w.[P](w_2) \subseteq [P](w)) = UV(\lambda w.[P](w_3) \subseteq [P](w))$. The order collapses totally, i.e., it does not make any difference which proposition is given as answer. Hence, $exh_{rel}^W(P(j), P) = [P(j)]^W$: our exhaustification operator adds nothing to the semantic meaning of the answer.

The other effects of exhaustive interpretation observed in section 2.2 can be treated in terms of $exh_{rel}^W(A, P)$ successfully as well. Consider the granularity effect, for instance. If Ann is known to be only interested in the amount of (full) *meters* that Bob can jump, a world u where he can jump 5.00 meters, a world v where he can jump 5.50 meters, and a world w where he can jump 5.80 meters are all equally relevant, $u \approx_P^{rel} v \approx_P^{rel} w$. It follows that by taking Bob's assertion 'I can jump five meters' exhaustively, we predict that in this case Ann can conclude only that he cannot jump six meters, but not that he cannot jump five meters and 10 centimeters.

This ends our excursion to a relevance-dependent notion of exhaustive interpretation. In the rest of the paper we will ignore this possible extension and continue with the basic version of our proposal.

7. Exhaustive interpretation as conversational implicature

As mentioned in section 2, some inferences we have analyzed under the heading of exhaustive interpretation have also often been explained as *conversational implicatures*, in particular as *scalar implicatures*. To give two examples, the exclusive interpretation of 'or' in (3): 'John or Mary', and the inference that not all students passed the examination from the exhaustive interpretation of (18d) 'Most students' (the question which (3) and (18d) address is again 'Who passed the examination?') are standard scalar implicatures.

It turns out that the description of exhaustive interpretation proposed in the previous sections also correctly predicts many other scalar implicatures. For instance, it also generates for (19b) the 'scale' reversal inference (19c).

(19) (a) Ann: In how many seconds can you run the 100 meters?
(b) Bob: I can run the 100 meters in 12 seconds.
(c) Bob cannot run the 100 meters in 11 seconds.

The reason for this 'scale' reversal is that in contrast to predicates like 'Bob owns n children' and 'Bob can jump n meters far', the question-predicate of (19a) behaves monotone increasingly in numbers:[41] if n is in the extension of the predicate and $m > n$, then m is in its extension as well.

Particularly pleasing is the observation that the approach to exhaustive interpretation defended here can also account for the well-known problematic cases of implicatures of complex sentences. For instance, using exh_{GS} or exh_{std} one can derive for multiple disjunctions as in answer (20) the inference that only one of the disjuncts is true (hence, only one of John, Mary, and Peter passed the examination).

(20) John, Mary, or Peter.

For example (21) we correctly predict that the interpreter can infer that John ate either only the apples or only some but not all of the pears.

(21) Ann: What did John eat?
Bob: John ate the apples or some of the pears.

[41] This has to be guaranteed by meaning postulates.

Given these observations it is not very surprising that at different places in the literature it has been suggested that exhaustive interpretation can be explained as a pragmatic phenomenon using Grice's theory of conversational implicatures (see, for instance, Harnish (1976), G&S (1984)). The central problem of such an approach is that there is no thoroughly satisfying formalization of Grice's theory and, hence, no precise description of the conversational implicatures an utterance comes with. But without such a rigorous description we cannot say whether Grice's theory indeed does account for certain (non-semantic) inferences an utterance comes with. In particular, we cannot make such a claim for the exhaustive interpretation of answers. Thus, before we can see whether Grice's theory can be used to explain exhaustive interpretation, we first need to formally describe at least parts of the conversational implicatures an utterance comes with.

In Schulz (to appear)[42] a new formalization of the Gricean reasoning leading to scalar implicatures is proposed. We will follow van Rooij & Schulz (2004) in adapting this approach to the formal situation at hand but also add some small improvements.

The following Gricean principle has – in different forms – often been taken to be responsible for scalar implicatures. It combines Grice's first subclause of the maxim of quantity with the maxims of quality and relevance.

The Gricean Principle
In uttering A a rational and cooperative speaker makes a maximally relevant claim given her knowledge.

In the special case we are interested in here, where the utterance given is an answer to some previously asked question, the principle comes

[42] This paper summarizes the findings of the first author's masterthesis published in 2004.

down to saying that the speaker will not withhold information from the audience that would help to resolve the question she is answering – she provides a[43] best (i.e. most relevant) answer she can, given her knowledge.

Our goal is to formalize the inferences an interpreter can derive if she takes the speaker of some sentence A to obey this principle. The solution proposed in Schulz (to appear) and van Rooij & Schulz (2004) is closely related to McCarthy's predicate circumscription and makes essential use of ideas developed by Halpern & Moses (1984) on the concept of *only knowing*, generalized by van der Hoek et al. (1999, 2000). We describe the possibilities where the speaker obeys the principle as those where she knows the sentence A she uttered to be true but knows as little as possible about the predicate in question besides what is semantically conveyed by her answer. Hence, as in the case of predicate circumscription, the enriched interpretation of answer A is described by selecting minimal models. Now, however, the selection takes place among those possibilities where the speaker *knows* A, and the order that determines minimality does not compare the extension of the question-predicate, but rather how much the speaker knows about this extension.

To formalize such an interpretation function, we have to refer to the knowledge state of the speaker. We will adopt a standard modal logical way of modeling knowledge. Let W be a set of models/possible worlds of our language.[44] We add to W an accessibility relation R that connects every element w of W with a subset $R(w)$ of W. This subset contains all worlds that are consistent with what the speaker knows in w. Then we can say that sentence $\mathbf{K}A$, 'the speaker knows A', is true in w (with respect to W and R) if A is true in every world in

[43] There may be more than one optimum.

[44] We allow multiple occurrences of the same interpretation function of the language in W.

$R(w)$. Because we want to model knowledge we demand that w is an element of $R(w)$. In this way we warrant that if the speaker knows A, the sentence is true in w.

Now, we can define an interpretation function that gives us besides the semantic meaning also the conversational implicatures due to the Gricean Principle. Assume that $\preceq_{P,A}$ is the order that compares how much the speaker who uttered A knows about the question-predicate P.

DEFINITION 8. *(Interpreting according to the Gricean Principle) Let A be an answer given to a question with question-predicate P in context $C = \langle W, R \rangle$. We define the pragmatic interpretation $grice^C(A, P)$ of A with respect to P and C as follows:*

$$grice^C(A, P) =_{def} \{w \in [\mathbf{K}A]^C | \forall w' \in [\mathbf{K}A]^C : w \preceq_{P,A} w'\}$$

Of course, this definition will only be of use if we can also give an explicit definition of the order $\preceq_{P,A}$, and hence, describe what it means that in one possibility the speaker knows more about the extension of the question-predicate than in another. But when is this the case? Informally, what we want to express is that a speaker has more knowledge about P if she knows of more individuals that they have property P. Thus, we say that $w_1 \preceq_{P,A} w_2$ if for every world v_2 considered possible by the speaker in w_2 (i.e. $v_2 \in R(w_2)$), she distinguishes some possibility v_1 in $R(w_1)$ where the extension of P is smaller than or equal to the extension of P in v_2.[45] But wait! It may be the case that the speaker makes in her utterance a claim about the extension of P which depends on some other facts. For instance, she may answer 'If they asked the same questions as last year then Peter passed the

[45] Some readers may notice that in this way we do not respect knowledge the speaker might have about some individuals not having property $[P]$. We would like to have some kind of motivation for why this information should not be taken into account, but until now we do not have a convincing explanation.

examination' to the question 'Who passed the examination?'. Of course, in this case we expect a speaker that obeys the Gricean Principle also to tell us whether – as far as she knows – they asked the same questions as last year. Therefore, we define the order as follows:[46]

DEFINITION 9.
Given a context $C = \langle W, R\rangle$ we define for $v_1, v_2 \in W$

$$\begin{array}{lll} v_1 \leq^*_{P,A} v_2 \ \textit{iff}_{def} & 1. & [P](v_1) \subseteq [P](v_2) \ \textit{and} \\ & 2. & \textit{for all non-logical vocabulary } \theta \textit{ occurring in } A \\ & & \textit{besides } P\textit{: } [\theta](v_1) = [\theta](v_2); \\ v_1 \equiv^*_{P,A} v_2 \ \textit{iff}_{def} & & v_1 \leq^*_{P,A} v_2 \ \textit{and} \ v_2 \leq^*_{P,A} v_1. \end{array}$$

DEFINITION 10. *(Comparing relevant knowledge)*
Given a context $C = \langle W, R\rangle$ we define for $w_1, w_2 \in W$

$$w_1 \preceq_{P,A} w_2 \ \textit{iff}_{def} \ \forall v_2 \in R(w_2)\ \exists v_1 \in R(w_1) : v_1 \leq^*_{P,A} v_2,$$
$$w_1 \cong_{P,A} w_2 \ \textit{iff}_{def} \ w_1 \preceq_{P,A} w_2 \ \&\ w_2 \preceq_{P,A} w_1.$$

Now that we have with *grice* at least a partial description of the conversational implicatures an utterance comes with, we can see whether the part of Grice's theory we have formalized can explain the exhaustive interpretation of answers. Unfortunately, it turns out that this is not the case. To illustrate the problem, let us calculate what *grice* predicts for example (22).

[46] $\leq^*_{P,A}$ is stronger than the order $\leq_P$ that we have used so far. If one would substitute the latter in the definition of $\preceq_{P,A}$, then *grice* would minimize the knowledge of the speaker about the extension of *all* non-logical vocabulary, which is inadequate for our purposes. In van Rooij & Schulz (2004) an even stronger order was used. There, condition 2 of definition 9 was dropped and non-logical vocabulary besides P did not play any role for the order. Then, however, one misses for the answer 'If they asked the same questions as last year then Peter passed the examination' the intuitive inference that the speaker does not know whether they asked the same questions as last year.

(22) Ann: Who passed the examination?
Bob: Mary.

Hence, let us determine $grice^C(P(m), P)$. We choose a model where for every individual there exists a unique name. To make things even simpler, we assume that in context C there are only four different worlds: w_1, w_2, w_3, and w_4 with R and P defined as given in figure 1.[47]

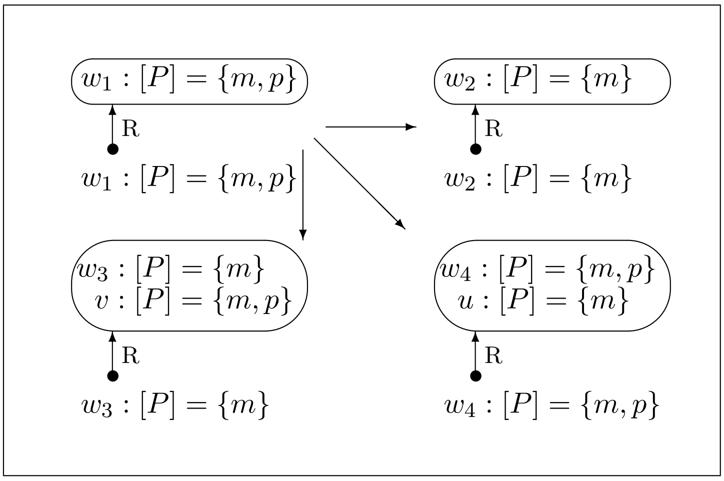

Figure 1.

What we would like to predict in such a situation is that all other individuals (in our example there is only one other individual: Peter (p)) did not pass the examination. To calculate $grice^C(P(m), P)$ according to definition 8, the first thing we have to do is to select those worlds w in W where the speaker knows that $P(m)$ is true. In turns out that this is the case for all elements of W. In a second step we select among those the possibilities where Bob knows least about the question-predicate P.

[47] Possible worlds are represented by points. Arrows annotated with R lead from a world w to the knowledge state $R(w)$ of the speaker in w. The arrows in the middle of the figure symbolize the ordering relation $\preceq_{P,A}$. Notice that in this example the worlds w_2 and w_3, for instance, differ, because in w_2 the speaker has a more definite opinion about the extension of P than in w_3.

The order tells us that the speaker knows more in w_1 than in w_2, w_3, and w_4, and that in the latter three worlds he knows equally much. Hence: $grice^C(P(m), P) = \{w_2, w_3, w_4\}$. A closer look reveals that this interpretation allows the interpreter to derive from Bob's answer the conversational implicature that he does not know that Peter passed the examination (i.e. $grice^C(P(m), P) \models \neg\mathbf{K}P(p)$). But we are not able to derive the desired inference that Peter, in fact, did not pass the examination. Hence, we have to conclude that the Gricean Principle, at least in the formalization given above, cannot explain exhaustive interpretation.

Actually, many students of conversational implicatures will find this a rather pleasing result. It has often been argued in the literature that the conversational implicatures due to the Gricean Principle[48] should be generated primarily with the weak epistemic force we predict (see, among others, Soames (1982), Leech (1983), Horn (1989), Matsumoto (1995), and Green (1995)). Hence, the conversational implicature of Bob's answer is indeed claimed to be that he does not know for people other than Mary that they passed the examination. Only in contexts where the speaker is assumed/believed to be competent/an authority on the subject matter under discussion, these authors propose, one can derive the stronger inference that what the speaker does not know to hold indeed does not hold (hence, in the example the desired inference that Peter did not pass the examination).

For our approach this would mean that we should be able to obtain the exhaustive interpretation by calculating *grice* with respect to the set C of contexts where the speaker Bob is competent/an authority on the question she is answering. However, in van Rooij & Schulz (2004) it is shown that this will not lead to an adequate description of the exhaustive interpretation of answers (or their scalar implicatures). The

[48] In particular, conversational implicatures due to the first subclause of the maxim of quantity.

problem is that some sentences that can be interpreted exhaustively (or give rise to scalar implicatures) cannot stem from a speaker that is at the same time competent/an authority and obeys the Gricean Principle. An example is the answer Bob gives in (3), here repeated as (23).

(23) Ann: Who passed the examination?
Bob: John or Mary.

In the present paper we have proposed to analyze the often observed exclusive interpretation of 'or' as due to exhaustive interpretation, and at many places in the literature the inference that not both disjuncts are true at the same time has been claimed to be a scalar implicature. However, the approach sketched above will not predict any conversational implicatures for such disjunctive answers. The reason is that a competent speaker should know which of the disjuncts is true and, if obeying the Gricean Principle, should have given this information. The fact that Bob nevertheless did not do so shows that he either is not competent or has disobeyed the principle. In neither case is the exclusive interpretation of 'or' predicted.

To overcome this problem but nevertheless stay faithful to the intuition that competence/authority plays a decisive role for the derivation of exhaustivity effects/scalar implicatures, van Rooij & Schulz (2004) propose to *maximize* the competence of the speaker when interpreting answers. However, it is only maximized in so far as this is consistent with taking the speaker to obey the Gricean Principle.[49]

[49] In this respect, our analysis bears resemblance to Gazdar's (1979) proposal that clausal implicatures with weak epistemic force can cancel scalar ones that have strong epistemic force, and with Sauerland's (2004) method of strengthening implicatures with weak epistemic force. Our approach is based on essentially the same ideas as Spector's (2003) Gricean justification of exhaustive interpretation. In contrast to all

There is an obvious way to extend the function *grice* such that it follows this idea. We introduce a second order on the set of possibilities that compares the competence of the speaker in different possibilities (a possibility is higher in the order if the speaker is more competent). Then we select maximal elements with respect to this order – but now only among those possibilities where the speaker obeys the Gricean Principle, i.e., among the elements in $grice^C(A, P)$.

Let $\sqsubseteq_{P,A}$ be the order that compares competence. The interpretation function defined below tells an interpreter what she can infer if she takes the speaker, first, to obey the Gricean Principle, and, second, to be as competent with respect to the question she answers as is consistent with the first assumption.

DEFINITION 11. *(Adding Competence to the Gricean Principle) Let A be an answer given to a question with question-predicate P in context $C = \langle W, R\rangle$. We define the pragmatic interpretation $eps^C(A, P)$ of A with respect to P and C as follows:*

$$eps^C(A,P) =_{def} \{w \in grice^C(A,P) | \forall w' \in grice^C(A,P) : w \not\sqsubset_{P,A} w'\}$$

$$= \{w \in [\mathbf{K}A]^C | \forall w' \in [\mathbf{K}A]^C : w \preceq_{P,A} w' \wedge (w \cong_{P,A} w' \rightarrow w \not\sqsubset_{P,A} w')]]\}.$$

Again, to make this definition useful we have to define the order $\sqsubset_{P,A}$ properly. We propose that in a world w_2 the speaker is as least as competent as in world w_1 if in w_1 the speaker considers as least as many extensions possible for question-predicate P as in w_2.[50]

these analyses, however, ours is more general, fully model-theoretic, and based on standard methods of non-monotonic reasoning.

[50] Here, again, we slightly deviate from the approach in van Rooij & Schulz (2004). There it is proposed to compare only what the speaker knows about objects that do not have property $[P]$. Because of the way that *grice* is defined, it does not make any difference for *eps* which of the two definitions is chosen.

DEFINITION 12. *(Comparing competence)*
Given a context $C = \langle W, R\rangle$ we define for $w_1, w_2 \in W$

$$w_1 \sqsubseteq_{P,A} w_2 \text{ iff}_{def} \forall v_2 \in R(w_2) : \exists v_1 \in R(w_1) : v_2 \equiv^*_{P,A} v_1.$$

To illustrate the working of the new strengthened interpretation function *eps*, let us reconsider our example (22) Ann: Who passed the examination? Bob: Mary. We calculate $eps^C(P(m), P)$, where W is defined as in figure 1. Remember that we want to obtain that Bob's answer implies that Peter did not pass the examination. Because we already know that $grice^C(A, P) = \{w_2, w_3, w_4\}$, the only thing that still has to be done is to select among the possibilities in this set those where according to $\sqsubseteq_{P,A}$ the speaker is maximally competent. Unsurprisingly, w_2 is the unique $\sqsubseteq_{P,A}$-maximum in $\{w_2, w_3, w_4\}$. Hence, $eps^W(P(m), P) = \{w_2\}$. But, as figure 1 shows, in w_2 the desired conclusion $\neg P(p)$ holds! So, for this example the new interpretation function combining Gricean reasoning with a principle of maximizing competence predicts correctly.

Of course, we would like to establish that *eps* can account for the exhaustive interpretation of answers in some generality. At least it would be pleasing if *eps* does not perform worse in describing exhaustive interpretation than does exh_{std}. It turns out that both notions are indeed closely related.[51]

FACT 1. *If A and ϕ do not contain modal operators and $C = \langle W, R\rangle$ is chosen such that there is no previous information in the context then*

$$eps^C(A, P) \models \phi \text{ iff } exh^C_{std}(A, P) \models \phi.$$

[51] The extension of the definition of exh_{std} to context $C = \langle W, R\rangle$ is straightforward. The proof of fact 1 goes very much along the same lines as the proof of a similar claim given in van Rooij & Schulz (2004). In the way the orders are defined here one needs the additional assumption that for all $w \in [A]^C$ there is some $w' \in exh^C_{std}(A, P)$ such that $w' \leq_{P,A} w$. Spector (2003) proves a closely related fact.

Hence, if the answer given does not contain modal operators and there is no information in the context, then both interpretation functions predict the same modal-free inferences. Until now, all examples discussed in this paper were of this kind. Thus, all the pleasing predictions made by exh_{std} are inherited by eps.

However, our richer modal analysis allows us, additionally, to describe exhaustivity effects that could not be accounted for in terms of exh_{std}. One kind of example are sentences that contain modal expressions, like belief attributions or possibility statements as in the answer 'John and perhaps also Mary'. This advantage of the Gricean derivation of exhaustive interpretation given above is explicitly discussed in van Rooij & Schulz (2004) and we will not repeat that discussion here. In this section we only want to indicate (as also discussed in the paper mentioned) how the modal approach can help us to solve a last problem of G&S's (1984) approach that has not been addressed so far: the exhaustive interpretation of negative answers.

Remember our discussion at the end of section 3: exh_{GS} predicts wrongly that the answer 'Not John' to question 'Who passed the examination?' means that *nobody* passed the examination. The same prediction is made for all (other) cases in which the question-predicate occurs under negation in the answer. To solve this problem, von Stechow & Zimmermann (1984) propose to modify G&S's exhaustivity operator by selecting in these cases not the *minimal* extensions of the predicate, but rather the *maximal* ones. In terms of our framework this means that now the speaker gives not the exhaustive extension of question-predicate P, but of its *complement*, $\bar{P}$, instead. Thus, in case P occurs negatively in A, we should not look for $exh^W_{std}(A, P)$ but rather for $exh^W_{std}(A, \bar{P})$. Then, the answer 'Not John', for instance, would be interpreted as implying that, except for John, everybody passed the examination. According to most of our informants, however, this kind of exhaustive interpretation is the exception, rather than the rule. They report instead that negative answers give rise to the conclusion that the semantic meaning of the answer is the only information the speaker has about the question-predicate. Interestingly enough, the same intuition is reported also for other answers, for instance, if the speaker uses special intonation or responds 'Well/As far as I know, Peter'.[52]

A very welcome side-effect of the Gricean explanation given to exhaustive interpretation in this section is that we can correctly describe this interpretation when we only apply the function *grice*, hence, take the speaker to obey the Gricean Principle, but not maximize her competence. This suggests the following explanation for the non-exhaustive interpretation of the answers discussed above. The speaker is always taken to fulfill the Gricean principle and, hence, *grice* is applied to the answer. We normally also take the answerer to be competent[53] and, hence, apply *eps*. However, the answerer can cancel this additional assumption by either mentioning that she is not competent or simply deviating from the standard form of answering a question (by using negation, special intonation, etc.). In this way we can correctly predict the weakening of exhaustive interpretation to 'limited-competence' inferences for such answers.

8. Conclusion and outlook

In this paper we did two things. First – and this was the central goal of the work presented – we propose a description of the exhaustive interpretation of answers. The main concept this description builds on is that of interpretation in minimal models, which we took from

[52] See also footnote 7.

[53] This seems to be a natural default assumption, given that only in such situations it makes perfect sense to ask a question to a certain addressee.

AI-research.[54] It constitutes the fundament of our formalization of exhaustive interpretation and holds the whole paper together. The second backbone of our description is dynamic semantics. It provides us with the semantic framework in which we embedded minimal interpretation. And finally, we use standard conceptions of relevance to bring communicational interests of the agents into play. Brought together, these three independent lines of research allow us to account for many observations on the phenomenon of exhaustive interpretation.

In the last section of the paper we have gone beyond the primary goal to provide an adequate description of exhaustive interpretation. We used a proposal made in Schulz (to appear) and van Rooij & Schulz (2004) to provide a pragmatic explanation for this rule of interpretation. Exhaustive interpretation is explained as based on the assumptions that first, the speaker obeys the Gricean Principle and, second, that she is competent on the question she answers (as far as this is consistent with the first assumption). We propose a formalization of these assumptions and the reasoning based on them that can be shown to perform as least as well in describing exhaustive interpretation as does exh_{std}. In fact, it turns out that this pragmatic explanation can account for a certain contextual weakening of exhaustive interpretation that none of our operations exh_{std}, exh_{dyn} or exh_{rel} could deal with.

This part of our work allows us to answer a question that accompanied us the whole paper: what is the relation between exhaustive interpretation and conversational implicatures? According to us, ex-

[54] We only know of one (other) attempt to use circumscription for (some of) the data we discuss in this paper: by Wainer in his dissertation (1991). When applying circumscription to utterances directly, he came across some of the same problems that we discussed for G&S's proposal. For this reason he opts, in the end, for a second description in which stipulated abnormality predicates are circumscribed. One of the main goals of this paper was to show that the direct approach without additional abnormality predicates can be pushed much further than Wainer assumed.

haustive interpretation refers to a class of conversational implicatures, and many scalar implicatures are among them. This interpretation is the result of a Gricean-like reasoning about rational behavior of cooperative speakers. This explains why so often in the literature the notions 'exhaustive interpretation' and 'scalar implicature' are used to describe the same observation.

In sections 5, 6, and 7 we have presented three different extensions of our basic description exh_{std} of exhaustive interpretation: exh_{dyn} was based on a dynamic approach to semantics, exh_{rel} took a contextual parameter of relevance into account, and *eps*, finally, modeled exhaustive interpretation as a consequence of a Gricean-like reasoning pattern. All of them addressed certain shortcomings of our initial account in terms of exh_{std}, but none of them overcomes all of them. The ultimate goal should be to combine all these extensions into one uniform description. We did not present the account in this way, because it would have made the paper much less readable. The interaction between the different extensions raises many additional questions that have to be addressed carefully. For instance, one can easily define *eps* based on a dynamic semantics with the aim of giving a Gricean motivation for exh_{dyn} as an extension of our justification for exh_{std}. But, then, new questions come up that one has to deal with. For example, in how far should information the speaker has about discourse referents involved in the answer be taken into account when comparing her knowledge? For some answers to these questions we would be able to present a dynamic version of fact 1, for others not. To keep these complications out of the already quite demanding discussion of the paper, we decided to split up our approach in different units and present them separately. However, this should not make the reader loose sight of the composite form of our proposal.

Above we said that we take scalar implicatures to be a subclass of the inferences of exhaustive interpretation. At the beginning of the paper, however, we introduced exhaustive interpretation as the (normal) interpretation of answers. On the face of it this would mean that we predict scalar implicatures to be restricted to answers – what some of our readers may think a rather dangerous claim. But, first, the fact that exhaustive interpretation as discussed here was restricted to answers to overt questions does not necessarily mean that it occurs only in these contexts. This restriction was forced upon us mainly because we did not have sufficient empirical data to support a general statement about the contexts in which exhaustive interpretation occurs. Furthermore, one of the central issues in the recent literature on scalar implicatures is the context-dependence of these inferences. In particular, it has been claimed that questions can play an important role for the presence of scalar implicatures (see, for instance, Hirschberg (1985) and van Kuppevelt (1996)). Further research on this subject has to clarify in which contexts we do observe scalar implicatures, and whether they coincide with the contexts of exhaustive interpretation.

Another interesting question for further research is whether the given formalization of the Gricean Principle can be extended to a general implementation of Grice's maxims of conversation. Consider, for instance, the second subclause of the maxim of quantity. This subclause is taken to be the driving force behind another class of pragmatic inferences: those to the most stereotypical interpretation. For instance, that we normally interpret 'John killed the sheriff' as meaning that John murdered the sheriff in a stereotypical way, i.e. by knife or pistol, is often explained with reference to this maxim. Inferences to the stereotype/normal case (called I-implicatures by Atlas & Levinson (1981), and R-implicatures by Horn (1984)) are often analyzed as being in some sense opposite to scalar implicatures.[55] Against this background it is interesting to observe that the minimal model approach can be used naturally to account for the latter inferences as well.[56] The only thing that we have to change is how we instantiate the ordering. In this case it is not predicate minimization that counts – of which relevance minimization is a natural extension – but rather minimization of *normality* (or maximization of *plausibility* or *expectedness*). Thus, now we have to assume that $v \prec_A w$ iff v is a less surprising A-world than w is, and the interpretation of A w.r.t. $\prec_A$, $\{w \in [A]^W | \neg\exists v \in [A]^W : v \prec_A w\}$, results then just in the set of most plausible worlds that verify A. In the future we would like to see to what extent this formalization can account for the wide range of I or R implicatures described by Atlas & Levinson and Horn as due to this maxim,[57] and how they interact with scalar implicatures.

Of course, the observation that both types of implicatures may be captured by very similar interpretation rules does not make them necessarily the same phenomenon. In AI there has been an intense debate on the interpretations that non-monotonic reasoning formalisms can receive. One of the distinctions made there seems to show up here again. We have described exhaustive interpretation as a rule of negation as failure in the message: from the fact that the speaker did not say p for a certain class of propositions p, the interpreter infers that $\neg p$. Already McCarthy (1986) mentioned such rules of language use as examples of circumscription in action. A similar rule may also govern the I or

[55] The intuition being that while in case of scalar implicatures some stronger claim is excluded, in case of inference to the stereotype some stronger claim is assumed to hold.

[56] In fact, these are the inferences non-monotonic reasoning was originally made for. See, for instance, McCarthy's motivation for introducing Predicate Circumscription as briefly discussed in section 4.1.

[57] In several papers, e.g. Asher & Lascarides (1998), a sophisticated method of non-monotonic reasoning is used to account for some of these inferences.

R implicatures: if the speaker did not mention that the situation is in a certain way abnormal, then the interpreter can conclude that it is normal. But here we do not have to take the detour via language use. It may also simply be the case that the interpreter concludes to the stereotypical interpretation because it is for her the normal state of affairs given the information she has (including the message of the speaker). Note that this is not an admissible interpretation of exhaustive readings: if we learn that Mary has property P, only in very exceptional cases will general knowledge about how the world normally is allow us to infer that John does not have property P.

Hence, in summary, while the inference of negation as failure inherent in exhaustive interpretation is most plausibly due to rule-governed conversational behavior (which may be conventional or not), the inference to the stereotypical interpretation does not need to be anchored in language use.

Acknowledgements

The work on exhaustive interpretation goes back several years for both authors, but got a new boost when we realized the similarity between G&S's rule of exhaustive interpretation and McCarthy's predicate circumscription. Early versions of this paper were first presented in Amsterdam and later at the Fifth International Workshop on Computational Semantics in Tilburg (January 2003) and at several other workshops, summerschools (ESSLLI, NASSLLI), and lectures. We would like to thank the audiences at those meetings for their questions and comments. We also would like to thank Richmond Thomason for pointing us to Jacques Wainer's dissertation. Most of all, however, we would like to thank Johan van Benthem, Jeroen Groenendijk and Martin Stokhof for critical readings of earlier versions of this paper and Jason

Mattausch for correcting the English. Finally, we would like to thank two anonymous referees for their useful suggestions.

References

Aloni, M. (2003), 'On choice-offering imperatives', In: P. Dekker & R. van Rooy (eds.), *Proceedings of the Fourteenth Amsterdam Colloquium*, Amsterdam, pp. 57-62.

Asher, N. and A. Lascarides (1998), 'Bridging', *Journal of Semantics*, **15**: 83-113.

Atlas, J. and S. Levinson (1981), 'It-Clefts, Informativeness and Logical Form', In: P. Cole (ed.), *Radical Pragmatics*, Academic Press, New York, pp. 1-61.

Benthem, J. van (1989), 'Semantic parallels in natural language and computation', In: H. D. Ebbinghaus et al. (eds.), *Logic Colloquium '87*, Elsevier Science Publishers, Amsterdam, pp. 331-375.

Bonomi, A. and P. Casalegno (1993), 'Only: association with focus in event semantics', *Natural Language Semantics*, **2**: 1-45.

Gazdar, G. (1979), *Pragmatics*, Academic Press, London.

Ginzburg, J. (1995), 'Resolving questions, I', *Linguistics and Philosophy*, **18**: 459-527.

Green, M. (1995), 'Quantity, volubility, and some varieties of discourse', *Linguistics and Philosophy*, **18**: 83-112.

Grice, H. P. (1967), 'Logic and Conversation', typescript from the William James Lectures, Harvard University. Published in P. Grice (1989), *Studies in the Way of Words*, Harvard University Press, Cambridge Massachusetts, pp. 22-40.

Groenendijk, J. and M. Stokhof (1984), *Studies in the Semantics of Questions and the Pragmatics of Answers*, Ph.D. thesis, University of Amsterdam.

Halpern, J.Y. and Y. Moses (1984), 'Towards a theory of knowledge and ignorance', *Proceedings 1984 Non-monotonic reasoning workshop*, American Association for Artificial Intelligence, New Paltz, NY, pp. 165-193.

Hamblin, C.L. (1973), 'Questions in Montague English', *Foundations of Language*, **10**: 41-53.

Harnish, R. M. (1976), 'Logical form and implicature', In: T. G. Bever et al. (eds.), *An integrated theory of linguistic ability*, Crowell, New York, pp. 313-391.

Hirschberg, J. (1985), *A theory of scalar implicature*, Ph.D. thesis, Pennsylvania State University.

Hoek, W. van der, J. Jaspers and E. Thijsse (1999), 'Persistence and minimality in epistemic logic', *Annals of Mathematics and Artificial Intelligence*, **27**: 25-47.

Hoek, W. van der, J. Jaspers and E. Thijsse (2000), 'A general approach to multi-agent minimal knowledge', In: M. Ojeda-Aciego et. al. (eds.), *Proceedings JELIA 2000*, LNAI 1919, Springer-Verlag, Heidelberg, pp. 254-268.

Horn, L. (1984), 'Towards a new taxonomy of pragmatic inference: Q-based and R-based implicature', In: D. Schiffrin (ed.), *Meaning, Form, and Use in Context: Linguistic Applications*, GURT84, Georgetown University Press, Washington, pp. 11-42.

Horn, L. (1989), *A Natural History of Negation*, University of Chicago Press, Chicago.

Kadmon, N. (1985), 'The discourse representation of NPs with numeral determiners', In: S. Berman et al. (eds.), *Proceedings of NELS 15*, GLSA, Linguistics Dept. University of Massachusetts, Amherst, pp. 207-19.

Kamp, H. and U. Reyle (1993), *From Discourse to Logic*, Kluwer, Dordrecht.

Kratzer, A. (1989), 'An investigation of the lumps of thought', *Linguistics and Philosophy*, **12**: 607-653.

Karttunen, L. (1977), 'Syntax and semantics of questions', *Linguistics and Philosophy*, **1**: 3-44.

Kuppevelt, J. van (1996), 'Inferring from topics. Scalar implicatures as topic-dependent inferences', *Linguistics and Philosophy*, **19**: 393-443.

Leech, G. (1983), *Principles of Pragmatics*, Longman, London.

Lewis, D. (1973), *Counterfactuals*, Blackwell, Oxford.

Matsumoto, Y. (1995), 'The conversational condition on Horn scales', *Linguistics and Philosophy*, **18**: 21-60.

McCarthy, J. (1980), 'Circumscription - a form of non-monotonic reasoning', *Artificial Intelligence*, **13**: 27-39.

McCarthy, J. (1986), 'Applications of circumscription to formalizing common sense knowledge', *Artificial Intelligence*, **28**: 89-116.

McCawley, J. (1993), *Everything that linguists always wanted to know about Logic**, The University of Chicago Press, Chicago.

Merin, A. (1999), 'Information, relevance, and social decisionmaking', In: L. Moss et al. (eds.), *Logic, Language, and Computation, Vol. 2*, CSLI publications, Stanford, pp. 179-221.

Parikh, P. (1992), 'A game-theoretical account of implicature', In: Y. Vardi (ed.), *Theoretical aspects of Rationality and Knowledge: TARK IV*, Montery, California.

Rooth, M. (1996), 'Focus', In: S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, Blackwell Publishers, Malden Massachusetts, pp. 271-297.

Rooij, R. van (2003), 'Questioning to resolve decision problems', *Linguistics and Philosophy*, **26**: 727-763.

Rooij, R. van (2004), 'Utility, informativity and protocols', *Journal of Philosophical Logic*, **33**: 389-419.

Rooij, R. van and K. Schulz (2004), 'Exhaustive interpretation of complex sentences', *Journal of Logic, Language, and Information*, **13**: 491-519.

Rooij, R. van and K. Schulz (to appear), '*Only*: meaning and implicatures', In: M. Aloni, A. Butler & P. Dekker & (eds.), *The Semantics and Pragmatics of Questions and Answers*, Elsevier, Amsterdam.

Sauerland, U. (2004), 'Scalar implicatures of complex sentences', *Linguistics and Philosophy*, **27**, 367-391.

Schulz, K. (to appear), 'A pragmatic solution for the paradox of free choice permission', *Synthese*.

Szabolcsi, A. (1981), 'The semantics of topic-focus articulation', In: J. Groenendijk et al. (eds.), *Formal methods in the study of language*, Mathematical Centre, Amsterdam, pp. 513-540.

Soames, S. (1982), 'How presuppositions are inherited: A solution to the projection problem', *Linguistic Inquiry*, **13**: 483-545.

Spector, B. (2003), 'Scalar implicatures: Exhaustivity and Gricean reasoning?', In: B. ten Cate (ed.), *Proceedings of the ESSLLI 2003 Student session*, Vienna.

Stechow, A. von and T.E. Zimmermann (1984), 'Term answers and contextual change', *Linguistics*, **22**: 3-40.

Wainer, J. (1991), *Uses of nonmonotonic logic in natural language understanding: generalized implicatures*, Ph.D. thesis, Pennsylvania State University.

Zeevat, H. (1994), 'Questions and Exhaustivity in Update Semantics', In: H. Bunt et al. (eds.), *Proceedings of the International Workshop on Computational Semantics*, Institute for Language Technology and Artificial Intelligence, Tilburg.

Institute for Language, Logic and Computation (ILLC)
University of Amsterdam
Nieuwe Doelenstraat 15

1012 CP Amsterdam
The Netherlands
{K.Schulz,R.A.M.vanRooij}@uva.nl