SCALAR IMPLICATURES: EXHAUSTIVITY AND GRICEAN REASONING

BENJAMIN SPECTOR
Laboratoire de Linguistique Formelle (Paris VII)/Ecole Normale Supérieure/Institut Jean-Nicod
benjamin.spector@ens.fr

ABSTRACT: This paper shows that both scalar implicatures and exhaustification of answers can be understood as the outcome of a pragmatic reasoning based on Gricean maxims. I offer a formalization of the Gricean reasoning that solves the problems (cf. Chierchia 2001) faced by standard neo-Gricean accounts. I further show that positive and non-positive answers pattern very differently, in a way that can be predicted by stating carefully, for a given question-answer pair, what counts as an "alternative answer" - this notion plays the same role as that of “scalar alternative” in previous approaches.

1. Imperfections of Standard Neo-Gricean Accounts

According to neo-Gricean accounts, scalar implicatures are computed as follows: given a sentence S containing a scalar term t, S is to be compared to all sentences which can be obtained from S by replacing t with a term belonging to t’s *scale*. For any such *scalar alternative* S’ such that S’ asymmetrically entails S, the hearer infers that S’ is not part of the speaker’s beliefs. (hereafter, rule R1; this derives the so-called *clausal*[1] or *primary*[2] implicatures). The underlying principle motivating this inference is Grice’s first maxim of *Quantity*. Assuming further that the speaker is maximally informed, the hearer infers that S’ is in fact false according to the speaker (hereafter, rule R2).

(1) A or B (2) A and B

Suppose the speaker utters a sentence of the form of (1). Its unique scalar alternative is (2). Since (2) is logically stronger than (1), (2) is not part of the speaker’s belief. Moreover, if the speaker is maximally informed, (2) is false, so that *or* in (1) is interpreted as exclusive, even though its literal linguistic meaning is that of inclusive disjunction.

Whatever the merits of this approach (in particular, the fact that it predicts that the exclusive reading of *or* should disappear in monotone decreasing contexts, due to the reversal of entailment patterns), it has been shown to be inaccurate in many cases, especially when a scalar term is interpreted under the scope of some operators. For instance, Chierchia (2001) points out that the neo-Gricean procedure yields too weak results for a sentence like (3):

(3) Each of the students read *Othello* or *King Lear*

(3) (sometimes) implicates (4)[3]:

(4) Each of the students read *Othello* or *King Lear* and not both.

The neo-Gricean account predicts a much weaker implicature, namely (5):

(5) It is not the case that each of the students read *Othello* and *King Lear*.

Another problem is that the neo-Gricean account can also lead to too strong predictions. Take a sentence of the following form:

(6) (A or B) or C

Scalar alternatives of (6):
a. (A and B) or C b. (A or B) and C c. (A and B) and C

All these alternatives are stronger than (6), so that (6) should implicate that they are all false (by rule R2). In particular, a. should be false, in which case C is. But (6) certainly does not implicate that C is false[4]. Let me call this problem, which is actually very general, that of *unwanted negations*. If, on the other hand, we find a way of blocking this inference, we remain unable to predict that (6) normally implicates that only one of the three disjuncts is true.

1. 1 Chierchia's localist solution

Chierchia (2002) presents a solution based on a recursive interpretation function which computes "strengthened meanings" in tandem with the interpretation function that computes "literal meanings". For him, scalar implicatures are simply an additional dimension of meaning, and the link between scalar implicatures and general principles of conversational rationality becomes less clear, even though some basic aspects of the neo-Gricean approaches are retained.

Hereafter, I will defend a "globalist" approach to scalar implicatures, in the sense that it relies on the natural hypothesis that pragmatic processes operate at least at the sentential level.

1. 2. Sauerland (2004)

Sauerland (2004)[5] proposes a globalist approach to the puzzle of multiple disjunctions which relies on two modifications of the standard neo-Gricean account. First, he expands the set of alternatives for a given sentence, and second, he motivates a modification of the inference rules. A sentence S of the form '(A or B) or C' will have the following alternatives:

{A, B, C, (A or B), (A or C), (B or C), (A and B), (A and C), (B and C), (A and B) or C, (A or B) and C, (A and B) and C, (A or B) or C}

The first inference rule is meant to capture what Sauerland calls *primary implicatures,* i.e. inferences of the form 'The speaker does not hold the belief that….'. For any alternative S' that asymmetrically entails S, it follows from the maxim of

quantity that the speaker does not believe that S' is true. In this particular case, all the alternatives (as defined above) entails S. So we derive, among other things, the fact that the speaker does not believe A to be true, nor B nor C, and also that he does not believe (A or B) to be true. Sauerland uses the following notation, borrowed from epistemic logic :

¬KA, ¬KB, ¬KC, ¬K (A or B), etc.

Next, all the logical consequences of these statements are computed. For instance, since the speaker does not know '(A or B)' to be true, and since, on the other hand, he believes '(A or B) or C' to be true (maxim of quality), it follows that he cannot know C to be false. Indeed, if he both believed C to be false and '(A or B) or C' to be true, then he would believe '(A or B)' to be true, which contradicts '¬K(A or B)'. Therefore ¬K¬C can be added to the set of primary implicatures. More generally, it follows that the speaker must be uncertain about the truth of each of the disjuncts. After all the logical consequences of primary implicatures are derived, and added to the set of primary implicatures, so-called *secondary implicatures* are computed as follows:

If ¬Kf is a primary implicature and ¬K¬f has not be derived as a primary implicature or a logical consequence of primary implicatures, then infer the following : K¬f.

In the case of multiple disjunctions, you get the intended reading by deriving first ¬K(A and B), ¬K(A and C), ¬K(B and C) and then, by using the second inference rule : K¬(A and B), K¬(A and C), K¬(B and C).

A few comments on Sauerland's procedure, which is I believe basically on the right track: One question is how exactly the alternatives are defined. Sauerland needs to say that for any sentence of the form (A or B), the set of its alternatives is {A, B, (A or B), (A and B)}. But because he wants the relation "being an alternative of" to be an equivalence relation (as in the standard view), he runs into the following problem: any two sentences X and Y are alternatives of each other; indeed, since (X or Y) is an alternative of X and of Y, X and Y are alternatives of each other (by symmetry and transitivity of the relation "being an alternative of"). But if this were true, then no scalar implicature would ever be derived.[6] Sauerland solves this problem by an entirely *ad hoc* move: he introduces two binary connectors c_L and c_R such that 'A c_L B' 'is equivalent to A, and 'A c_R B' is equivalent to B, and then stipulates the following *scale* : <or, and, c_L, c_R>. Then the alternatives of (A or B) are the following : {A c_L B, A c_R B, A or B, A and B}. This set is *semantically* equivalent to the previous one, but it is not the case anymore than any two sentences X and Y are alternatives of each other, even though 'X c_LY' and 'X c_R Y', which are equivalent to 'X' and 'Y', are, for any X and Y. As the author notes, these two connectors are actually never *used* –and this is accounted for by the maxim of *manner*, which says that one is supposed to be brief. Not only is this an *ad hoc* move; it is also not in accordance with the Gricean intuition that sentences are to be compared to other sentences that *could* have been uttered instead. Instead of resorting to this move, one could have simply given up the constraint that alternative sets be equivalence classes. After all, it would seem quite natural to say that 'A' is an alternative of 'A or B', but not the other way around, based on the plausible view that a given sentence should be compared only to sentences that are not more complex (one could interpret Grice's maxim of *manner* in this way). But this is not the solution I will advocate. Rather, I

will *nearly* claim that any two sentences X and Y are alternatives of each other. More precisely, what I will claim is that any two *elementary answers* to a given *question under discussion* are alternatives of each other. Consider the following dialogue:

(7) Who came ?
Peter came.

I define the set of *elementary answers* to the question in (7) as the set of all propositions of the form 'x came', where x ranges over the contextually given domain of quantification. The proposition that, say, 'John came', is therefore an alternative to 'Peter came'. Using Sauerland's scale, it is also the case that 'John came and Peter came' is among the alternatives. This alternative asymmetrically entails the proposition expressed by the answer actually uttered. From which the hearer derives (in two steps) that the speaker believes that 'John came and Peter came' is false, which, together which the fact that he believes 'Peter came' to be true, entails that he believes 'John came' to be false. More generally, for any individual C distinct from Peter in the quantificational domain, it follows that the speaker believes that C didn't come. What I have informally derived is the so-called *exhaustive* interpretation of answers. My main claim in this paper is that exhaustivity can be derived from Gricean assumptions.

1. 3. Is exhaustification the solution ?

Van Rooy (2002), on the other hand, uses exhausitivity as his starting point, and claims that scalar implicatures are just a sub-case of it. He proposes to derive scalar implicatures from the fact that, if a certain question Q is under discussion and a certain sentence S is given as an answer to Q, S is generally interpreted as "exhaustive".

The exhaustivity operator (Groenendijk & Stokhof 1984) operates on answers of the form 'GQ P', where GQ stands for a generalized quantifier and P for a predicate. The question under discussion is understood as "for which objects is P true of these objects ?".

The exhaustivity (*exh*) operator works as follows[7]:
$[\![\text{exh (GQ P)}]\!] = 1$ iff $[\![P]\!] \in (\text{Min } [\![GQ]\!])$, where $(\text{Min } [\![GQ]\!])$ is the set that includes only the minimal members of $[\![GQ]\!]$, i.e:
$\text{Min } [\![GQ]\!] = \{x \mid x \in [\![GQ]\!] \text{ and there is no } x' \text{ in } [\![GQ]\!] \text{ such that } x' \subset x\}$
($\subset$ = "is a proper subset of")

Example:
(7) a. Among John, Mary and Peter, who came?
b. John or Mary came
$[\![\text{John or Mary}]\!] = \{\{J, M, P\}, \{J, M\}, \{J, P\}, \{J\}, \{M, P\}, \{M\}\}$
$(\text{Min } [\![\text{John or Mary}]\!]) = \{\{J\}, \{M\}\}$
$[\![\text{exh (John or Mary came)}]\!] = 1$ iff $[\![\text{came}]\!] \in \{\{J\}, \{M\}\}$ i.e. iff only John came or only Mary came.

Van Rooy shows that when exhaustification is applied to monotone increasing contexts, it can solve some of Chierchia's puzzles.

1. 4. When exactly do we exhaustify answers ?

However, if exhaustification is applied to a sentence 'GQ P' where GQ is decreasing, exhaustification as defined above leads to unrealistic implicatures: "less than two chemists came" should implicate that nobody came! So Van Rooy uses a second exhaustivity operator (*exh'*) in these cases, following Stechow & Zimmermann (1984):

$[\![\text{exh' (GQ P)}]\!] = 1$ iff $[\![P]\!] \in (\text{Max } [\![GQ]\!])$, where $(\text{Max } [\![GQ]\!])$ is the set that includes only the maximal members of $[\![GQ]\!]$

There are several problems with this account. First, the second rule of exhaustification makes wrong predictions:

(8) a. Among the chemists and the philosophers, who came?
b. Less than two of the chemists

Exhaustification leads to b':
b' . Exactly one chemist and all the philosophers came.

But b. does not seem to implicate b'; b. actually *does* indeed suggest that some chemist came, but does not implicate anything regarding non-chemists. It rather suggests that the speaker does not know much about them.

Second, these two rules are unable to account for cases where the speaker combines increasing and decreasing quantifiers, thus creating a non-monotone GQ, as in (9)b:

(9) a. Among the chemists, the philosophers and the linguists, who came?
b. Less than two chemists but one philosopher came

If we apply the first exhaustivity operator, what we get is that b. implicates that no chemist and no linguist came, while exactly one philosopher came. If we apply the second exhaustivity operator, what we get is that exactly one chemist, all the philosophers and all the linguists came. None of these predictions is in fact borne out. Rather, it seems that (9) implicates that at least one chemist came, exactly one philosopher came, and that the speaker does not know much about linguists.

1. 5. Goal of this paper: deriving exhaustivity

In the next sections, I show that both scalar implicatures and exhaustification of answers can be understood as the outcome of a pragmatic reasoning that is based on the Gricean maxims. I will first offer a precise formalization of the Gricean reasoning, meant to replace the two rules R1 and R2. I will then show that it is possible to predict the facts reviewed above by defining carefully what counts as an "alternative answer" for a given answer to a certain question under discussion[8].

2. Formalizing the Gricean reasoning

I now assume that a certain sentence A is uttered as an answer to a (maybe implicit) question Q, and I adopt a partition semantics for questions

(Groenendijk & Stockhof 1984): Q induces an equivalence relation R_Q, over the set of worlds.

Notation:

- $w\ R_Q\ w'$ — w and w' belongs to the same « cell»
- $R_Q(v) = \{w \mid w\ R_Q\ v\}$ — (= the set of worlds equivalent to v, or v's cell)
- $\alpha(w)$ — _ is true in w (alternatively: $w \in$ _)
- $\alpha \subseteq \beta$ — _ is a subset of _; _ entails _
- $\alpha \subset \beta$ — _ is a proper subset of _; _ asymmetrically entails _

The proposition α expressed by A is supposed to meet the condition of *strong relevance*:

Def 1 (strong relevance): A proposition α (= set of worlds) is strongly relevant with respect to a question Q if

a) $\exists w, (R_Q(w) \cap \alpha) = \emptyset$ (i.e.: α excludes at least one cell)

and

b) $\forall w, (\alpha(w) \leftrightarrow R_Q(w) \subseteq \alpha))$ (α does not distinguish between two worlds that belong to the same cell, i.e. provides no irrelevant information)

The speaker's information state is modeled as a set of worlds, i.e. a proposition. As an agent believes a lot of things that are irrelevant in the context of a given question, it is useful to define what counts as the *relevant information* contained in a certain information state:

Def 2 (relevant information): Let i be an information state and Q a question. Then we define i relativized to Q, written as i/Q, as follows:

$i/Q = \{w \mid \exists w', (w'\ R_Q\ w \text{ and } w' \in i)\}$ $(= \cup_{w \in i} R_Q(w))$.

The Gricean reasoning is based on the idea that α (the proposition given as an answer) must be compared to a certain set of *alternative propositions*[9] which the speaker could have chosen instead of α. This **alternative set**, call it S, must contain α itself, and be such that all its members are relevant[10]. The hearer's task when interpreting the speaker's utterance is to address the following question: *given that the speaker has preferred α to all the other members of S, what does this entail regarding his information state i_0?* First, the speaker must believe α to be true (Grice's maxim of quality), i.e. i_0 must entail α. Second, α must be optimal in the sense that there must be no more informative proposition in S entailed by the speaker's beliefs (Grice's maxim of quantity), i.e. there must be no proposition α' such that i_0 entails α' and α' asymmetrically entails α. Put differently, i_0 must belong to the following set I(S, α, Q):

Def 3: $I(S,\alpha, Q) = \{i \mid i/Q \subseteq \alpha \text{ and } \forall \alpha' (\alpha' \in S \text{ and } i/Q \subseteq \alpha') \rightarrow \neg (\alpha' \subset \alpha)\}$

So if a certain proposition β is entailed by no member of I(S,α,Q), the hearer can conclude that β is not part of the speaker's belief. This reasoning plays the role of rule R1. It is immediately predicted that if the speaker utters a sentence P of the form "A or B" and if the propositions expressed by A and by B belong to the alternative set S, as I will assume (so does (2002)), then the speaker cannot know A to be either true or false: if A were true, then A would have been a better answer than P; if A were false,

B would be true (since P is), and B would have been a better answer than P[11]. Now, let the hearer assume that the speaker is *as informed as possible given the answer he made.* This means that his information state i_0 is maximal in I(S,α,Q) in the following sense: there is no i' in I(S,α,Q) such that i' (relativized to Q) asymmetrically entails i_0 (relativized to Q). In other words, i_0 belongs to Max(S,α,Q), defined as follows:

Def 4: $Max(S,\alpha,Q)=\{i \mid i \in I(S,\alpha,Q) \text{ and } \forall i' \, (i' \in I(S,\alpha,Q)) \rightarrow \neg \, (i'/Q \subset i/Q)\}$

From this the hearer can conclude that if a proposition β is entailed by all the members of Max(S,α,Q), then β is believed by the speaker. This reasoning plays the role of R2, but is not equivalent to it: there is no way of deriving an "unwanted negation". In the case of a disjunctive statement in which the disjuncts are logically independent, the disjuncts and their negations are entailed by no member of I(S,α,Q), as shown above, so that they cannot be entailed by any member of Max(S,α,Q) either, since Max(S,α,Q) is included in I(S,α,Q).

From now on, whenever it clear what the question under discussion is, and considering that the content of an alternative set only depends on the question under discussion and the sentence uttered, I will simply write I(α) and Max(α) instead of I(S,α,Q) and Max(S,α,Q). S(α) will denote the alternative set of α.

3. Alternative sets and Exhaustification

3. 1. An example

Let P be of the form '(A or B) or C', where A, B and C are logically independent. Assume that (1) is uttered in a context in which A, B and C's truth-values are what is relevant i.e. as an answer to a question Q amounting to "Which sentence(s) are true among A, B, and C?"

For any information state i, the relevant part of i in this context (i.e i/Q) belongs to the boolean closure of {A,B,C}. So we will loose nothing if we view information states as *sets of valuations of* {A, B, C}, i.e. as propositions of the propositional language based on {A,B,C}, where any such proposition actually stands for a class of propositions that are all equivalent when relativized to Q. Let S(P) (the alternative set of P) be *the closure under union and intersection of* {A,B,C}[12]. Intuitively, S(P) is the *set of positive answers to Q*: S(P)={A,B,C,A∨B,A∧B,A∨C,A∧C,B∨C,B∧C,(A∨B)∨C,(A∧B)∨C,A∨(B∧C)...}Assume i_0 = ((A∨B)∨C))∧(¬(A∧B) ∧ (¬(A∧C)∧¬(B∧C))). Then $i_0 \in$ I(P), since P is the only – and therefore best - proposition in S(P) entailed by i_0[13]; i_0 can also be described as the set of the three following valuations:

	A	B	C
W1	T	F	F
W2	F	T	F
W3	F	F	T

I now show that Max (P) = $\{i_0\}$, e.g. that i_0 entails all the members of I(P). Suppose i_1 is an information state that is not entailed by i_0 and that belongs to I(P). There is then an element of i_0 that does not belong to i_1. Suppose W_1 does not belong to i_1. Then i_1 entails P': $P' = \neg(A \wedge (\neg B \wedge \neg C)) = \neg A \vee (B \vee C)$

But i_1 belongs to I(P), and therefore entails P. Hence i_1 also entails P'':

$P'' = ((A \vee B) \vee C)) \wedge (\neg A \vee (B \vee C)) = (B \vee C)$

But P'', which belongs to S(P), would have been a better answer than P in information state i_1, so that i_1 does not belong to I(P), contrary to the hypothesis. Things are similar if W2 or W3 does not belong to i_1 (by symmetry). Hence Max(P) = $\{i_0\}$, and P implicates i_0.

This proof can be generalized to all formulas whose only logical operators are disjunctions.

3. 2. Background concepts

As shown in section 1.4., answers lead to different kinds of implicatures, especially regarding exhausitivity, depending on whether they are, intuitively speaking, positive or negative. But this cannot make sense so far, as I have not said precisely what it is for an answer to be "positive". This is the goal of the present section.

I now assume that questions are all equivalent to something like:

Q: "For which x is P(x) true?", where x is of any semantic type, and P is a certain predicate (simple or complex) that can be built in a natural language.

I further assume that the domain of quantification is fixed and finite, and known to all participants. Thus any relevant answer to Q can be translated into the following propositional language L_Q: let $(c_i)_{0< i < n+1}$ be an enumeration of names for each of the individuals of the domain. Then L_Q is the propositional language with disjunction and conjunction as its only binary connectors and based on the atomic sentences $(P_i)_{0< i < n+1}$, where P_i translates $P(c_i)$.

Now, relevant answers to Q can be seen as sets of valuations of $(P_i)_{0< i < n+1}$. And the relevant part of any information state can also be seen as a set of valuations. So we can assimilate information states to sets of valuations, without loosing anything.

Definitions (see appendix):

1. **Literal**: a **literal** is an atomic sentence or the negation of an atomic sentence. A literal is **positive** if it is an atomic sentence, **negative** otherwise
2. **Sentence P *favors* literal L**: a sentence or a proposition P favors a literal L iff there is a valuation V such that V(P) = V(L) = 1 and $V_{-L}(P) = 0$, where V_{-L} is defined as the valuation which is identical to V except for the value it assigns to L.
3. **Sentence P *essentially mentions* literal L**: A sentence P essentially mentions a literal L iff L occurs without a negation preceding it in every P' equivalent to P and such that the scope of all negations occurring in P' is an atomic sentence.
4. **positive sentence/positive proposition**: a sentence or a proposition is positive (resp. negative) iff it favors at least one positive (resp. negative) literal and no negative (resp. positive) literal.

We can then prove the following theorems (see Appendix):

Theorem 1: For any sentence P and any literal L, P favors L iff P essentially mentions L

Theorem 2: A sentence P is positive (resp. negative) iff P is equivalent to a sentence which belongs to the closure of positive (resp. negative) literals under conjunction and disjunction

Corollary: A sentence P is positive iff it is equivalent to a sentence P' which contains no negation.

We therefore have two characterizations of positive answers: an answer is positive if it is equivalent to a sentence which contains no negation, or, equivalently, if it favors at least one positive literal and no negative literal. This equivalence will prove helpful.

3. 3. The case of positive propositions: predicting exhaustification

The alternative set of any positive proposition is defined as **the set of all positive propositions**[14].

3. 3.1. An Example

Consider the following dialogue:

(10) Among John, Peter, Mary and Sue, who will come?
- Well, John will come, or Peter and Mary will come

I translate the answer into a propositional language containing four atomic sentences A, B, C and D:

$P = A \vee (B \wedge C)$

P quite clearly implicates Q : $Q = (A \wedge \neg B \wedge \neg C \wedge \neg D) \vee (B \wedge C \wedge \neg A \wedge \neg D)$ i.e. "either only John will come, or only Peter and Mary will", which is exactly what exhaustification in Groenendijk & Stokhof's sense would yield.

What I will now prove is that Max(P) = {Q}, from which it indeed follows that P implicates Q. First, I show that $Q \in I(P)$, i.e. P is an optimal answer in S(P) in information state Q. Suppose the speaker's information state is Q. Q can be represented as the following set of valuations, where a valuation is itself represented as the set of atomic sentences that this valuation makes true: Q = {{A},{B, C}}.

By hypothesis, the speaker has to choose a proposition that belongs to the alternative set. This proposition must be entailed by Q and be such that there is no better proposition in the alternative set. Let Q' be a positive sentence entailed by Q. Necessarily the valuation represented by {A} is in Q'. But then, the valuation {A,B} must be in Q' too: if {A,B} were not in Q', indeed, ¬B would be **favored** by Q', since there would be a valuation v making ¬B true in Q' (namely v = {A}) and such that the valuation v' identical to v except over B (v'={A, B}) would not be in Q'; so Q' would favor a negative literal and not be positive, contrary to the hypothesis. By the same reasoning, {A,C}, {A,D} {A,B,C}, {A,B,D}, {A,C,D} and {A,B,C,D} must belong to Q', and so does {B,C,D} (since {B,C} is in Q and therefore in Q'). So any positive proposition entailed by Q must include the following proposition, i.e. be entailed by it:

{{A}, {A,B}, {A,C},{A,D}, {A,B,C},{A,B,D}, {A,C,D},{A,B,C,D}, {B,C}, {B,C,D}} (= P)

But this set, **which turns out to represent P**, is a positive proposition which is entailed by Q and which entails all other positive propositions that are entailed by Q (as I have just shown). So P is the strongest positive proposition entailed by Q, i.e. $\underline{Q \in I(P)}$ (recall that I(P) is the set of all information states which make P an optimal answer among positive answers).

<u>Second, I show that Max(P) = {Q}</u>. This amounts to proving that Q entails all the members of I(P). Assume there is an information state i which belongs to I(P) and is not entailed by Q. Since i is not entailed by Q, then either {A} or {B, C} does not belong to i. Suppose {A} does not belong to i. On the other hand, i belongs to I(P) and therefore entails P. From which it follows that i entails P-{A}, i.e. i is included in the following set of valuations:

P-{A}=
{{A,B},{A,C},{A,D},{A,B,C},{A,B,D},{A,C,D},{A,B,C,D},{B,C},{B,C,D}}

But this set is itself a positive proposition, since it can be checked that P-{A} favors no negative literal. In fact, P-{A} can be written as: $(A \wedge (B \vee C \vee D)) \vee (B \wedge C)$. So i entails a positive proposition that is stronger than P, namely P-{A}, which contradicts the hypothesis that i belongs to I(P). Things work similarly if {B,C} does not belong to i. Therefore there is no such i. From which it follows that Q entails all the members of I(P). Q.E.D

<u>3. 3.</u> <u>2. Predicting exhaustification</u>

In the general case, positive answers are predicted to be interpreted as **exhaustive**.

Definitions (see also the appendix):

1. Exhaustification:

Let P be any non-negative proposition, then the function Exhaust is defined as follows:

Exhaust(P) = {V | V∈P and there is no valuation V' in P such that V' ⊂ V}

This operator is the propositional counterpart of Groenendijk & Stockhof's exhaustivity operator.

2. Positive extension of a proposition P: for any non negative proposition P, there is a unique positive proposition Q such that P entails Q and Q entails all the other positive propositions that P entails (i.e. Q is **the strongest positive proposition that P entails**). This can be shown by using the same reasoning as in the previous section: namely, you get Q by adding to P all the valuations that are needed in order not to favor any negative literal (see the appendix). The result of this operation I call the *Positive Extension of P*, or **Pos (P)**.

For any P, Pos(P) = {V | there is a valuation V' in P such that V' ⊆ V} (recall that a valuation is seen as a set of atomic sentences)

Facts (proved in the appendix): for any non negative proposition P,

1. If P is positive, Pos(P) = P
2. Pos (Exhaust(P)) = P
3. If P is a positive proposition and V a minimal member of P, then P – {V} is a positive proposition too.

Theorem: if P is a positive proposition, then Max(P) = {Exhaust(P)}, and therefore P implicates Exhaust(P).

Proof: We prove a) that Exhaust(P) ∈ Max(P), and b) that Exhaust P is the only member of Max(P)

a) Let P be a positive proposition. I(P) is the set of states i making P an optimal answer, i.e. such that P is the strongest positive proposition entailed by i, i.e. such that P = Pos(i). Hence I(P) = {i|Pos(i)=P}.

Since Pos(Exhaust(P))=Pos(P)=P (by facts 1 and 2), Exhaust(P) ∈ I(P).

b) *Ad absurdum*: we want to show that Exhaust(P) entails all the other members of I(P). Let's assume, to the contrary, that there is a member i_1 of I(P) such that Exhaust(P) does not entail i_1. Then there is a valuation in Exhaust(P) which does not belong to i_1, call it V. Given that i_1 entails P, i_1 also entails P-{V}. Since V, belonging to Exhaust(P), is a minimal member of P, P-{V} is positive (by fact 3), and P-{V} is therefore a positive proposition entailed by i_1, from which it follows that P cannot be the strongest positive proposition entailed by i_1, i.e. P≠Pos(i_1). Therefore i_1 does not belong to I(P), contrary to the hypothesis. QED.

3. 3. 3. Pair-list questions

Consider sentence (3) again ("Each of the students read *Othello* or *King Lear*"). If (3) is understood as an answer to a pair-list question like *"Which students read which plays by Shakespeare?"*, exhaustification predicts an exclusive reading for *or*. Note that the translation of a certain natural language sentence into a sentence of propositional logic will yield different results for different underlying questions (see section **3.2.**). In the case of the above pair-list question, but not in other cases, atomic sentences represent elementary answers of the type 'x read y', and (3) will be translated as something like (3'):

(3') (A∨B)∧(C∨D) ∧(E∨F)∧…… ∧(G∨H)

Exhaustification of (3') yields the desired result (exclusive reading for all the disjunctions). This context-dependency explains why judgments are not uniform.

3. 4. Non-positive propositions

We have seen in **1.4** that negative answers are not exhaustified, but nevertheless trigger some implicatures. This is straightforwardly predicted if the alternative set of a negative proposition P consists in *the closure under disjunction and conjunction of all the literals that P favors.* The asymmetry between negative and positive answers then boils down to the fact that positive answers are compared to all positive answers, while negative answers are compared only to a proper subset of the negative answers.

Regarding answers that are neither positive nor negative, the data are quite complex, and judgments are not very robust. A good strategy is to look at the clearest cases, find which principles could account for them and then let these principles decide for the other cases:

(11) a. Among Peter, Mary and Jack, who came ?
b. Peter, but not Mary
>> No exhaustivity effect: we infer nothing regarding Jack

(12) a. Among the philosophers, the linguists and the chemists, who came?
b. Between two and five linguists
>> Exhaustivity effect : we infer that no chemist and no philosopher came.

One difference between (11) and (12) is that, even though both are neither positive nor negative, (12) is *quasi-positive* in the following sense:

Def 1: A proposition P ***strongly favors*** *a literal L if P favors L and P does not favor the negation of L*
Def 2: A proposition P is ***quasi-positive*** *if P does not strongly favor any negative literal.*

If we want to predict that only quasi-positive sentences lead to exhaustification, we may adopt the two following rules, which cover all the cases:
If P is quasi-positive, P's alternative set consists in the union of the set of positive propositions and {P} itself.
If P is not quasi-positive, then P's alternative set consists in the closure under union and intersection of all the literals that P favors.

These rules make the following predictions (assuming the question under discussion is the same as in (12)):
(13) Between two and five linguists and no philosopher came.
>> No exhaustivity effect: nothing should be implicated regarding chemists
(14) Between two and five linguists and three philosophers came
>>Exhaustivity effect: suggests that no chemist came
(15) Three philosophers but less than two chemists came
>>No-exhaustivity effect: nothing should be inferred regarding linguists.

Though judgments are not so clear, an informal inquiry seems to indicate that most people have the expected intuitions. More work needs to be done in order to understand what is really going on here[15].

4. Conclusion

I have offered a precise formalization of the Gricean reasoning that underlies scalar implicatures, and exhaustification of answers. I have shown that the facts regarding exhaustification can be directly derived from the Gricean reasoning[16]. The only stipulations that were needed concern the rules according to which alternative sets are built. Yet the original notion of "scalar alternatives" is also stipulative. It remains to be seen whether the role played by polarity (namely, the distinction between positive and non-positive answers) can be derived in a more principled way. I suspect that introducing a notion of *utility* in our model of information processing, as Nilsenova & Van Rooy (2002) do in order to account for the pragmatic effects of polar questions, could help explain why "positive" and "negative" answers pattern asymmetrically. Another topic that must be investigated is the following one: a sentence like "John will come or John and Mary will" expresses the same proposition as "John will come", but is not interpreted in the same way. This shows the limitations of any procedure that only takes into account the literal semantic values of sentences, and not their actual phonological and syntactic form. I will explore some possible solutions to this problem.

Finally, let me point out that while the procedure I have defined is context-dependent (since implicature computation depends on what the question under discussion is), it is possible to devise a very similar procedure that would not be context-dependent. These two procedures, taken together, can provide us with an analytic tool for investigating to what extent scalar implicatures are *generalized* rather than extremely sensitive to context.

Appendix

1. Language

Let L be the following propositional language:

1.1. Vocabulary

i) Atoms : $p_1, p_2, \ldots, p_n$ (N.B. : there is a finite number of atoms)
ii) ⊥, T, (,), [,], ¬, ∧, ∨

1.2. Syntax

i) **Literals**: if p is an atom, then [p] is a **literal** and [¬p] is a **literal**. ⊥, T are literals.

ii) **Formulae**:

- if L is a literal, then L is a formula
- For any two formulae F and G, '(F ∨ G)' is a formula and '(F ∧ G)' is a formula.

Terminology: A positive literal is a literal of the form [p]; a negative literal is a literal of the form [¬p]. ⊥ and T are neither positive nor negative.

Notation: Let L be a literal whose atom is p. Then –L (the *negation of L*) is defined as follows :

–L = '[¬p]' if L = '[p]' and –L = '[p]' if L = '[¬ p]'.

1.3. Semantics

A *valuation* V is a function from all formulae to {0, 1} such that:

1) V(⊥) = 0
2) V(T) = 1
3) For any atom p, V([¬p]) = 1 – V([p])
4) For any formula F and any formula G, V ((F ∨ G)) = max (V(F), V(G))
5) For any formula F and any formula G, V ((F ∧ G)) = min (V(F), V(G))

A valuation is uniquely defined by the values it assigns to the positive literals. Hereafter, we treat valuations as functions from atoms to {0, 1}.

2. Definitions

Def 1: For any valuation V and any literal L distinct from ⊥ and T, we define V_{-L} as the unique valuation that is exactly like V except over the atom of L:

- For any atom p not occurring in L , V_{-L}([p]) = V ([p])
- V_{-L}(L) = 1 – V(L) (= V(-L), see below)

Def 2 (*favoring)* : Let F be a formula and L be a literal distinct from ⊥, _. Then F *favors* L if there exists a valuation V such that V(F) = V(L) = 1 and such that $V_{-L}(F) = 0$

Def 3 (*essentially mentions)* : Let F be a formula and L be a literal distinct from ⊥ and _,. Then F *essentially mentions* L if L occurs in every formula F' equivalent to F.

Def 4 (*positive proposition*): a *positive proposition* is a proposition that favors no negative literal and is distinct from ⊥ and T.

3. Theorems

Theorem 1: For any formula F and any literal L distinct from ⊥ and T, F essentially mentions L if and only if F favors L, and there exists a formula F' equivalent to F in which all the literals that F favors occur, and only those.

The proof relies on three lemmas:

Lemma 1: If F favors L, then F essentially mentions L.

Lemma 1.a: If F favors L then L occurs in F (equivalently : if L does not occur in F, then F does not favor L)

Proof: Let L be a literal. We prove by induction that for any formula F, is F favors L, then L occurs in F:

i) Suppose F is itself a literal. If F is distinct form L, then F does not favor L. (In particular, if F = - L, then F does not favor L). Therefore, if F favors L, then F = L, and L occurs in F

ii) Suppose F and G are formulae such that if F favors L, then L occurs in F and if G favors L, then L occurs in G. We prove that if 'F∨G' favors L, then L occurs in 'F∨G', and that if 'F∧G' favors L, then L occurs in 'F∧G'.

- Assume '(F∧G)' favors L. Then there exists a valuation V such that V((F ∧ G)) = V(L) = 1 & V_{-L}((F∧G)) = 0, i.e. min ((V(F), V(G)) = V(L) = 1 & min ((V_{-L}(F), V_{-L}(G)) = 0. From which it follows that V(F) = V(G) = V(L) = 1 and either V_{-L}(F) = 0 or V_{-L}(G) = 0. Therefore F favors L or G favors L. Therefore L occurs in F or in G, hence in '(F ∨ G)'.

- Assume '(F∨G)' favors L, Then there exists a valuation V such that V((F∨G)) = V(L) = 1 & V_{-L}((F∨G)) = 0, i.e. max ((V(F), V(G)) = V(L) = 1 & max ((V_{-L}(F), V_{-L}(G)) = 0. From which it follows that V_{-L} (F) = V_{-L}(G) = 0 and either V(F) = V (L) = 1 or V(G) = V(L)= 1, i.e. F favors L or G favors L. Therefore L occurs in F or in G, hence in '(F ∨ G)'. QED

Proof of Lemma 1: Let F be a formula and L be a literal that F favors. Suppose F' is a formula that is equivalent to F. Then F' favors L, and therefore L occurs in F' (*Lemma 1.a*). Therefore if F favors L, then F essentially mentions L. QED

Lemma 2 : For any formula F, there exists a formula F' equivalent to F in which the only literals that occur are those favored by F.

Proof of Lemma 2: We consider two cases;

i) First case : F is a tautology or a contradiction. Then F favors no literal at all and is equivalent either to ⊥ or to T, which are formulae in which the only literals that occur are those favored by F (in that case, no literal at all is favored by F).

ii) Second case : F is contingent. We will construct a disjunctive normal form F' equivalent to F in which only the literals favored by F occur.

Let us say that two valuations V_1 and V_2 are *F-equivalent* if :

$V_1(F) = V_2(F) = 1$ and for any literal L favored by F, $V_1(L) = V_2(L)$.

This equivalence relation defines a partition over the set of the valuations making F true. Each equivalence class can be represented by a partial valuation V_i that gives a value only to the atoms occurring in the literals favored by F. Let us consider the set E = $\{V_1,\ldots, V_n\}$ consisting of all these partial valuations. Let K $=\{L_1, \ldots, L_m\}$ be the set of the literals favored by F. For any i between 1 and n, let K_i be the subset of K such that V_i makes each member of K_i true, i.e. $K_i = \{L: L \in K \ \& \ V_i(L)=1\}$. Let D_i be the conjunctive formula such that every member of K_i occurs once in D_i, in "increasing order", i.e. for any j between 1 and m - 1, L_j occurs immediately on the left of L_{j+1} . Let D be the set $\{D_i,\ldots, D_n\}$. Let F' be the formula consisting of the disjunction of all the members of D, in increasing order. F' = '$D_1 \vee \ldots \vee D_n$'.

<u>Claim</u> : F' is equivalent to F

i) F' entails F.

Let V be a valuation such that V(F') = 1. Then V makes at least one member of D true, call it D_i. Since D_i is a conjunctive formula, V makes each of the literals occurring in D_i true, i.e. V makes each member of K_i true. Therefore it belongs to one of the equivalent classes defined by the relation "equivalent for F", namely, the equivalence class represented by the partial valuation V_i. Therefore V(F) = 1.

ii) F entails F'.

Let V be a valuation such that V(F) = 1. Let V_i be the partial valuation that is identical to V with respect to all the atoms that occurs in the literals favored by F and does not assign a value to other atoms. Let D_i be the corresponding conjunctive formula, which necessarily occurs in F'. By construction, $V_i(D_i) = 1$ and therefore $V(D_i) = 1$. Therefore V makes true at least one of the conjunctive formulae occurring in F', from which it follows that V(F') = 1

F' is, by construction, a normal disjunctive form in which the only literals that occur are those favored by F. It follows from *Lemma 1* that all the literals favored by F occurs in F'. QED

Lemma 3: For any formula F and any literal L distinct from ⊥ and T, F favors L if and only if F essentially mentions L

a) if F favors L, then F essentially mentions L (*Lemma 1*)
b) If F essentially mentions L, then F favors L

By contraposition : if F does not favor L, then F does not essentially mention L

Suppose F does not favor L. Let F' be a formula equivalent to F in which only the literals favored by F occur (such a formula exists by *Lemma 2*). Then L does not occur in F', and therefore F does not essentially mention L. QED

Proof of Theorem 1: *Theorem 1* follows from *Lemma 1*, *Lemma 2* and *Lemma 3*.

Theorem 2: if F favors only positive literals, then there exists a formula F' equivalent to F that contains no negation.

Proof: obvious from theorem 1.

Theorem 2 amounts to saying that any positive proposition can be expressed by a formula that contains no negation. Conversely, any formula that contains no negation and in which neither ⊥ nor T occurs expresses a positive proposition.

Theorem 3: for any non negative-proposition P, there is a unique positive proposition Q such that P entails Q and Q entails all the other positive propositions that P entails (i.e. Q is the strongest positive proposition that P entails).

We now represent a valuation. as the set of atoms it makes true. (A valuation defined as a function from atoms to truth-values is simply the characteristic function of a set of atoms). Propositions are sets of valuations, i.e. sets of sets of atoms.

We will show that for any non-negative P, *Pos(P)* as defined below is the unique positive proposition that P entails and that entails all the other positive propositions entailed by P.

Def 5 (*positive extension of a non-negative proposition*):

For any non-negative proposition P, Pos(P) = {V: V is a valuation such that there exists V'∈P such that V' ⊆ V}

Lemma 4: for any positive proposition Q, if a valuation V belongs to Q, then any superset V' of V belongs to Q.

Proof of lemma 4: Let Q be a positive proposition, and V be a member of Q. We first show that for any atom p that is not a member of V, V ∪ {p} belongs to Q. Suppose V ∪ {p} does not belong to Q. Then V is a valuation making both Q and [¬p] true and such that $V_{-[\neg p]}$, i.e. V∪{p}, makes Q false. Therefore Q favors a negative literal, namely [¬p], contrary to the hypothesis. Hence for any valuation V belonging to Q, every valuation V' obtained from V by adding one atom to V is also in Q. By repeating the same reasoning to all the valuations obtained from V by adding one atom, we conclude that any valuation V'' obtained from V by adding two atoms is also in Q, and so on for valuations obtained from V by adding a finite number of atoms. Since valuations are finite sets, it follows that all the supersets of V are in Q. QED

Proof of theorem 3: Let P be a non-negative proposition. Recall that P is viewed as a set of valuations. Let Pos(P) ={V : there exists V' ∈ P such that V' ⊆ V}. In other terms Pos (P) is the set consisting of all the supersets of the members of P.

We prove that Pos(P) is the unique positive proposition entailed by P and entailing all the propositions entailed by P. Keep in mind that 'A entails B' now means the same as 'A is included in B'.

We show that a) P entails Pos(P), b) Pos(P) is a positive proposition, and c) Pos(P) entails all the positive propositions entailed by P

a) P entails Pos (P) : obvious from the definition of Pos(P)

b) Pos (P) is a positive proposition.

i) if a valuation V belongs to Pos (P), then any superset of V belongs to Pos (P). Indeed, if V belongs to Pos(P), then V is a superset of a member of P. Therefore any superset of V is also a superset of a member of P, and hence belongs to Pos(P).

ii) *Ad absurdum*: Suppose Pos (P) favors a negative literal L = [¬p]. Then Pos(P) contains a valuation V such that V makes L true, i.e. such that p does not belong to V, and such that V_{-L} does not belong to P. But V_{-L} is the valuation identical to V except over p, i.e. V_{-L} = V ∪ {p}. V_{-L} is therefore a superset of V, and thus also belongs to Pos(p), given i) >> contradiction

c) Pos(P) entails all the positive propositions entailed by P.

Suppose Q is a positive proposition entailed by P, i.e. such that P is included in Q. By Lemma 4 and the fact that P entails Q, all the supersets of the valuations belonging to P also belong to Q. Since all the elements of Pos(P) are supersets of the elements of P, they all belong to Q. QED

Def 6 (*exhaustification*): the function *Exhaust* is defined as follows; for any non-negative proposition P, Exhaust(P) = {V | V∈ P and for any valuation V' in P, if V'⊆ V, then V'=V}

In other words, Exhaust(P) is the set of all the *minimal* members of P (where minimal is defined w.r.t. the ordering relation ⊆)

Facts to be proved: for any non-negative P,

1. If P is positive, Pos(P) = P
2. Pos (Exhaust (P)) = Pos (P)
3. If P is a positive proposition and V a minimal member of P, then P– {V} is a positive proposition.

Proof of Fact 1*:*

Obvious : if P is positive, then P is the strongest positive proposition entailed by P

Proof of Fact 2:

We first prove the following lemma:

Lemma 5: Any member V of P is a superset of some member of Exhaust(P).

Ad absurdum: Suppose that there is a valuationV_1 in P that is not a superset of a member of Exhaust(P). Then V_1 does not belong to Exhaust(P), and there is therefore (by definition of *Exhaust*) a valuation V_2 that is properly included in V_1 and belongs to P. But this valuation itself is not a member of Exhaust(P) (otherwise V_1 would in fact be a superset of a member of Exhaust(P)). Consequently, there is a valuation V_3 properly included in V_2 that is a member of P but not of Exhaust(P). By iteration of this reasoning, there is an infinite sequence $(V_i)_{i\in N}$ such that each of the members of the sequence is properly included in its predecessor and all of them belong to P. Since valuations are finite sets, however, there is an integer n such that V_n is the empty set (since each valuation has strictly less members than its predecessors); but then V_{n+1} cannot be properly included in V_n, which is contradictory. QED

In order to prove Fact 2, we prove first that a) Pos (Exhaust (P)) ⊆ Pos (P) and b) that Pos(P) ⊆Pos(Exhaust(P))

a) Pos (Exhaust (P)) ⊆ Pos (P)

Let V be a member of Pos(Exhaust (P)). Then V is a superset of some member of Exhaust (P). Call V' this member of Exhaust (P). Since Exhaust(P) ⊆ P, V' is also a member of P and therefore V is a superset of a member of P. Consequently, V ∈ Pos(P). QED

b) Pos(P) ⊆ Pos(Exhaust(P))

Let V be a member of Pos(P). Then there is a V' in P such that V'⊆V. By Lemma 5, there is a valuation V'' in Exhaust (P) such that V' is a superset of V''. Therefore V' belongs to Pos(Exhaust(P)). Since V is a superset of V', by lemma 4, V also belongs to Pos(Exhaust(P)). QED

Proof of Fact 3:

Let P be a positive proposition and V be a minimal member of P, i.e. V ∈ Exhaust(P).

We prove that P –{V} = Pos (P – {V}), from which it follows that P – {V} is positive.

Assume V_1 ∈ P –{V}. Then V_1 ∈ P, and, by Lemma 4 and the fact that P is positive, any superset of V_1 also belongs to P. Let V_2 be a superset of V_1 (which therefore belongs to P). We show that V_2∈P –{V}. We consider two cases.

Case i): $V_2 = V_1$. Then $V_2 \in P - \{V\}$

Case ii): V_1 is properly included in V_2. Then necessarily $V_2 \neq V$; indeed, if $V_2 = V$, then V would not be a minimal member of P, since V_1 would be a proper subset of V belonging to P. Since $V_2 \neq V$ and $V_2 \in P$, it is also the case that $V_2 \in P - \{V\}$.

Therefore any valuation V_1 belonging to $P - \{V\}$ is such that all its supersets also belong to $P - \{V\}$, i.e. $Pos(P - \{V\}) \subseteq P - \{V\}$; since $P - \{V\} \subseteq Pos(P - \{V\})$, $P - \{V\} = Pos(P - \{V\})$. QED

REFERENCES

Chierchia, G.: 2002, "Scalar Implicatures, polarity phenomena, and the syntax/pragmatics interface", in A. Belletti (ed.), *Structures and beyond*, Oxford University Press, Oxford.

Gazdar, G.: 1979 *Pragmatics: Implicature, Presupposition, and Logical Form*, Academic Press, New York.

Groenendijk, J. & Stokhof, M.: 1984, *Studies in the Semantics of Questions and the Pragmatics of Answers*, PhD Thesis, University of Amsterdam.

Nilsenova, M. & Van Rooy, R.: 2002, "No's no good alternatives", Ms., talk given at the Stuttgart Workshop on Information Structure.

Sauerland, U.: 2004, "Scalar Implicatures in Complex Sentences", *Linguistics and Philosophy*, **27**, 367--391**.**

Spector, B.: 2003, "Scalar implicatures: exhaustivity and Gricean Reasoning", in B. ten Caten (ed.), *Proceedings of the ESSLLI 2003 student session*, Vienna

Van Rooy, R.: 2002, "Relevance Implicatures", Ms., ILLC, Amsterdam, available at http://semanticsarchive.net/Archive/WIyOWUyO/Implicfinal.pdf

Van Rooy, R, & Schulz, K.: 2004, "Exhaustive interpretation of complex sentences", to appear in *Journal for Logic, Language and Information*.

Stechow, A. von & Zimmermann, T. E.: 1984, "term answers and contextual change", *Linguistics,* **22**, 3-40

Zeevat, H., 1994, "Questions and Exhaustivity in update Semantics", in Bunt & al. (ed.), *Proceedings of the International Workshop in Computational Semantics*, Institute for Language technology and Artificial Intelligence, Tilburg.

[1] Gazdar (1979)

[2] Sauerland (2002)

[3] In section **3**.**3**.3, I account for the fact that this inference is not systematic.

[4] It has been observed long ago, for instance in Gazdar (1979), that a disjunctive statement is generally taken as indicating that the speaker is uncertain regarding the disjuncts' truth-values.

[5] Sauerland's paper has appeared in *Linguistics and Philosophy* in 2004, but was already available to me in 2003.

[6] If all the sentences were alternatives of each other, then no sentence could ever be interpreted as conveying more information than what it explicitly says: a speaker obeying the maxim of quantity should always chose to use a sentence whose literal meaning contains all the information he wants to express.

[7] I reformulate Groenendijk & Stockhof's exhaustivity operator in more simple terms, but the difference is immaterial.

[8] Since the first draft of this paper was written, Robert van Rooy and Katrin Schulz have also attempted to derive the exhaustivity facts from Gricean maxims, in a very simial way. See van Rooy & Schulz (2004)

[9] As my formulation makes clear, I am now adopting the simplifying view that what the hearer compares are *propositions*.

[10] The exact definition of alternative sets is the topic of section III.

[11] Assuming that A and B are logically independent.

[12] As the reader will have noticed, I treat sentences both as sentences of the object-language and as names (in the meta-language) of propositions, i.e. names of sets of worlds, in which case conjunction and disjunction are understood as intersection and union.

[13] I do not give the proof here, due to lack of space.

[14] It should be clear that the alternative set is dependent on the question under discussion, since "positivity" is defined in terms of the propositional language derived from the question under discussion *via* the translation procedure defined above.

[15] I do not give the proof that my two rules achieve the results I claim they do, due to lack of space.

[16] As an anonymous reviewer noticed, I have not addressed all the cases that Chierchia pointed out as problematic for the standard neo-Gricean procedure. Once again, limitation of space prevents me from doing so. Let me mention that a more sophisticated version of my proposal is able to predict the phenomenon of *conditional strengthening* (inference from "if A, then B" to "B if and only if A").