

INFORMATION TO USERS

This reproduction was made from a copy of a manuscript sent to us for publication and microfilming. While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. Pages in any manuscript may have indistinct print. In all cases the best available copy has been filmed.

The following explanation of techniques is provided to help clarify notations which may appear on this reproduction.

1. Manuscripts may not always be complete. When it is not possible to obtain missing pages, a note appears to indicate this.
2. When copyrighted materials are removed from the manuscript, a note appears to indicate this.
3. Oversize materials (maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or in black and white paper format.*
4. Most photographs reproduce acceptably on positive microfilm or microfiche but lack clarity on xerographic copies made from the microfilm. For an additional charge, all photographs are available in black and white standard 35mm slide format.*

***For more information about black and white slides or enlarged paper reproductions, please contact the Dissertations Customer Services Department.**

UMI University
Microfilms
International

8603648

Hirschberg, Julia Bell

A THEORY OF SCALAR IMPLICATURE

University of Pennsylvania

Ph.D. 1985

**University
Microfilms
International** 300 N. Zeeb Road, Ann Arbor, MI 48106

Copyright 1985

by

Hirschberg, Julia Bell

All Rights Reserved

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark .

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages
15. Dissertation contains pages with print at a slant, filmed as received _____
16. Other _____

University
Microfilms
International

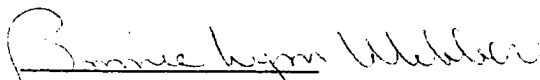
A THEORY OF SCALAR IMPLICATURE

Julia Bell Hirschberg

A DISSERTATION
in
Computer and Information Science

Presented to the faculties of the University of Pennsylvania in Partial
Fulfillment of the Requirements for the Degree of Doctor of Philosophy.

1985


Supervisor of Dissertation


Graduate Group Chairperson

**COPYRIGHT ©
Julia Bell Hirschberg
1985**

To D., C., and L.

ACKNOWLEDGEMENTS

I thank David Stemple for introducing me to computer science and persuading me to retrain in it and Peter Buneman for convincing me to go on for a PhD. Thanks also to Martha Palmer and Barbara Grosz for much valued advice throughout my graduate career. I am also grateful to IBM for generous financial support.

For help and encouragement during my thesis research, I thank my advisor, Bonnie Webber. I also thank my committee members: Ellen Prince, who introduced me to pragmatics and said 'scalar quantity implicature' at a crucial moment; Aravind Joshi, for questions that haunted me for days, and, in fact, still do; Tim Finin, whose help has ranged from theoretical to implementational; and Candy Sidner, for questioning my treatment of conversational implicature. A number of people have contributed naturally occurring tokens for the thesis, including Tim Finin, Dan Hirschberg, Kathy McCoy, Aravind Joshi, Martha Pollack, Ellen Prince, Ethel Schuster, Gregory Ward, and Bonnie Webber, without which the thesis would be much the poorer. Thanks also to Gopalan Nadathur for some crucial discussions and to Lloyd Komatsu and Sandy Waxman for guiding me through the classification literature. Special thanks to Ira Winston and Richard Billington, for making the machines that supported this research work -- even on weekends; they are the best. I am most grateful to Kathy McCoy, who has given me excellent comments and criticisms on every draft of every manuscript that has preceded this thesis; no mean feat. Thanks also to my office mate Ethel Schuster for her constant support. Finally, I am deeply indebted to Martha Pollack, whose least valuable but best-remembered contribution to this thesis is the proper hyphenation of 'natural-language processing', and to Gregory Ward, for telling me all about language: Better friends, collaborators, and critics I never hope to find.

Table of Contents

1. Introduction	1
2. Conversational Implicature in the Gricean Framework	6
2.1. A Problem for Semantics	6
2.2. The Gricean Framework	8
2.2.1. Classifying Meanings	8
2.2.2. Formalizing Conversational Implicature	9
2.3. Maxims of Cooperative Conversation	11
2.4. Conversational Implicature	16
2.4.1. Implicature and Speaker Intention	17
2.4.2. Other Necessary Conditions on Conversational Implicature	19
2.4.2.1. Cancellation	24
2.4.2.2. Nondetachability	33
2.4.2.3. Redundancy of Conjunction/Reinforceability	36
2.4.3. Defining Conversational Implicature	38
2.4.4. Inferring Conversational Implicature	39
3. Scalar Implicature	41
3.1. Generalized Conversational Implicature	42
3.1.1. Classes of Generalized Implicatures	42
3.1.1.1. Asymmetric 'And'	43
3.1.1.2. Conditionals	44
3.1.1.3. Principle of Extra Effort	46
3.2. Generalized Quantity Implicature	47
3.2.1. Scalar Predication	47
3.2.2. Quantity-Quality Implicature	51
3.2.3. Defining Classes of Quantity Implicature	52
3.2.4. Summary	56
3.3. Scalar Implicature	56
3.3.1. Additional Measures of Informativeness	57
3.3.2. Higher, Lower, or Alternate Values	59
3.3.3. Affirming, Denying, and Declaring Ignorance of Scalar Values	62
3.3.4. Conventions of Scalar Implicature	65
3.4. Summary	66
4. Representing Scalar Implicature	68
4.1. Representing Quantity Implicature	69
4.1.1. Representing Scale	69
4.1.2. Representing Utterances	70
4.1.2.1. Sentences	71
4.1.2.2. Expressions	71
4.1.3. Ranking Utterances	72
4.1.4. Speaker Commitment	74
4.1.5. Summary	75

4.2. Epistemic Force	75
4.2.1. Horn's 'Distance from Pole'	75
4.2.2. The Epistemic Force of Scalar Quantity Implicatures	77
4.2.3. The Epistemic Force of Scalar Implicatures	79
4.3. Representing Scalar Implicature Conventions	82
4.4. Summary	82
5. Quantifying Informativeness	83
5.1. Orderings Supporting Scalar Implicature	84
5.1.1. Quantifier Orderings	84
5.1.2. Modals	87
5.1.3. Logical Connectives	87
5.1.3.1. Disjunctive Assertions	89
5.1.3.2. Conjunctive Assertions	90
5.1.3.2.1. All-Implication	90
5.1.3.2.2. Specificity of Assertion	91
5.1.4. Numerical Orderings	92
5.1.4.1. Cardinal Predicates	92
5.1.4.1.1. Approximating Contexts	93
5.1.4.1.2. Lexicalized Cardinals	93
5.1.4.1.3. 'Reversed' orderings	94
5.1.4.2. Ordinal Predicates	95
5.1.5. Indefiniteness and Definiteness	96
5.1.6. Ranked Entities, States, Actions, and Attributes	97
5.1.6.1. Rankable Entities	97
5.1.6.2. Rankable Activities, States, or Attitudes	98
5.1.6.3. Rankable Attributes	101
5.1.7. Temporal Implicatures	102
5.1.8. Spatial Orderings	105
5.1.9. Process Stages and Prerequisites	106
5.1.10. Sets and Whole/ Part Relationships	108
5.1.11. Type/ Subtype, Instance-of, and Generalization/ Specialization Relationships	111
5.1.12. Entity/ Attribute Orderings	112
5.1.13. Summary	114
5.2. Defining Scale	115
5.2.1. Limits of an Entailment Condition	115
5.2.2. Defining Scales by Suspension	117
5.2.3. Gazdar's Definition of Scale	120
5.2.4. Other Definitions of Scale	121
5.3. Ranking Utterances via Posets	121
5.3.1. Relations Not Supporting Scalar Implicature	122
5.3.2. Poset Rankings	125
5.3.2.1. Defining Posets over Expressions	126
5.3.2.2. Higher, Lower, and Alternate Values	126
5.3.2.3. Representing Scalar Implicature Orderings as Posets	128
5.4. Summary	132

6. Computing Scalar Implicatures	133
6.1. Context and Scalar Implicature	134
6.2. Defining Posets in Discourse	135
6.2.1. Sources of Divergent Definition	136
6.2.2. Defining Poset Compatibility	137
6.3. Salience and Scalar Implicature	140
6.3.1. Identifying Salient Expressions	141
6.3.1.1. Syntactic Cues	142
6.3.1.2. Intonational Cues	143
6.3.1.2.1. FALL-RISE Intonation	143
6.3.1.2.2. Contrastive Stress	145
6.3.1.3. Semantic and Pragmatic Cues	147
6.3.1.3.1. Given and New Information	147
6.3.1.3.2. Centering and Discourse Coherence	148
6.3.1.3.3. Local Focus	151
6.3.1.4. Communicative Dynamism	152
6.3.1.5. Defining Salient Expressions Using Focus	153
6.3.2. Identifying Salient Orderings	155
6.3.2.1. Selecting Salient Classification Hierarchies	156
6.3.2.1.1. Basic and Entry Levels Defined	157
6.3.2.1.2. Bounding Hierarchies by Entry Level	160
6.3.2.1.3. Extending 'Entry Level' Beyond Classification	163
6.3.2.2. Salient Orderings from Speakers' Goals	164
6.3.2.3. Selecting Between Duals	165
6.4. Calculating Scalar Implicatures	167
6.4.1. Multiple Salient Expressions	171
6.5. Summary	174
7. Scalar Implicature and Question-Answering	175
7.1. Appropriate Responses to Yes-No Questions	175
7.1.1. Theoretical Accounts of Yes-No Questions	176
7.1.2. Computational Approaches to Yes-No Questions	178
7.1.2.1. Availability-Driven Responses	178
7.1.2.2. Goal Inference and Indirect Responses	180
7.1.3. Blocking Potential False Scalar Implicatures	182
7.2. The QUASI System	185
7.2.1. Overview of the System	185
7.2.2. Domain	187
7.2.3. Knowledge Representation	188
7.2.3.1. Epistemic Representation	188
7.2.4. Representing Posets	189
7.2.4.1. Canonical Orderings	189
7.2.4.2. Derived Posets	189
7.2.4.3. Storing Identified Posets	190
7.2.5. Input and Semantic Processing	190
7.2.6. Direct Response Retrieval	191
7.2.7. Identifying Salient Expressions and Posets	192
7.2.8. Calculating Cooperative Responses	193
7.2.8.1. Responses When Multiple Expressions are Salient	196

7.2.8.2. Changing Salient Expression Changes Output	201
7.2.8.3. When Several Responses are Cooperative	204
7.2.9. Outputs	206
7.2.10. Future Research	206
7.2.11. Summary	207
8. Conclusions	208
BIBLIOGRAPHY	210
INDEX	219

List of Tables

Table 5-1: Horn's Suspenders

118

List of Figures

Figure 2-1: The Gricean Framework	8
Figure 3-1: Affirming a Scalar Value	48
Figure 3-2: Denying Scalar Values	49
Figure 3-3: Affirming Alternate Values	61
Figure 3-4: Denying Alternate Values	62
Figure 3-5: Implicatures Arising Via Denial	63
Figure 3-6: Declaring Ignorance of Higher and Lower Values	64
Figure 3-7: Declaring Ignorance of Alternate Values	64
Figure 6-1: Metrical Stress Tree	144
Figure 7-1: The QUASI System	186

CHAPTER I

Introduction

Speakers may convey many sorts of 'meaning' via an utterance, including ENTAILMENTS,¹ PRESUPPOSITIONS,² and IMPLICATURES,³ as well as the ILLOCUTIONARY FORCE⁴ and PERLOCUTIONARY EFFECT(S),⁵ which may be associated with any utterance. While students of natural-language processing have looked at some of these phenomena in considerable detail,⁶ they have left others unexamined. In particular, little attention has been paid to CONVERSATIONAL IMPLICATURES -- those non-truth-functional, context-dependent meanings identified by Grice [Grice 75] as relying for their interpretation upon speaker and hearer recognition of certain maxims of cooperative conversation. In effect, since hearers take into account not only those propositions speakers commit themselves to, but also those they do not, speakers may anticipate this inferential capacity when framing an utterance.⁷

This thesis presents a formal account of one type of conversational implicature termed here SCALAR IMPLICATURE, which relies for its generation and interpretation upon the hypothesis that cooperative speakers will say as much as they truthfully can that is relevant to a conversational exchange. For example, B's utterance of (1a)

- (1) A: How was the party last night?
a. B: Some people left early.
b. Not all people left early.

¹Those meanings which must be true in every possible world or model in which the sentence uttered is true.

²Traditionally, those meanings which are entailed by both the sentence uttered by the speaker and by its negation. More pragmatic views of presupposition define it as those meanings which must be true in order for a speaker felicitously to utter the sentence in question. [Gazdar 79a]

³Non-truth-functional meanings.

⁴The type of illocutionary act a speaker performs via some utterance, as, *asserting, promising, questioning*, etc. Roughly, what the speaker intends the utterance to do. [Searle 69]

⁵The effect a speaker has on some hearer by uttering some sentence; so, *convincing, inspiring, and getting hearer to realize* are termed perlocutionary acts. [Searle 69]

⁶Such as illocutionary force in indirect speech acts [Cohen 79, Perrault 80] and presupposition [Kaplan 79, Mercer 84].

⁷Note throughout that I will consider only declarative utterances.

B knows, (1b) also holds -- even though the truth of (1b) clearly does not follow from the truth of (1a). A may reasonably conclude that, if B in fact believes that everyone left early, s/he would have said so -- and, in planning his/her response, B may anticipate this conclusion. In Grice's terms, B has conversationally implicated (1b). This meaning is representative of the class of scalar implicatures.

Scalar implicatures may be distinguished from other conversational implicatures in that their generation and interpretation is dependent upon the identification of some salient relation that orders a concept referred to in an utterance with other concepts. In 1, for example, the salience of an inclusion relation between 'some people' and 'all people' in the discourse is prerequisite to B's implicating that (1b) -- and to A's understanding that (1b) has in fact been implicated. Given that 'all people' can be said to include 'some people', we can account for the scalar implicature of (1b) intuitively as follows: B has said (1a). S/he might just as easily have said (1b). (1b) may be viewed as a 'more informative' response than (1a), since it conveys all the information that (1a) does -- and more. If B had said (1b), we would not only know that B believed that some people left early -- but that s/he believed it not the case that everyone left early. This ordering of an utterance spoken with other possible utterances via the ordering of concepts mentioned and not mentioned is at the heart of scalar implicature.

The description of scalar implicature I will present is based upon the examination of 'non-logical' inferences apparently licensed in a large corpus of naturally occurring data collected by me and by others from 1982 through 1985. The bulk of data examined were transcripts of a radio call-in program, 'The Harry Gross Show: Speaking of Your Money', recorded from February 1-5, 1982, in which an expert provides callers with financial advice [Pollack 82]. Other protocols examined include Kathy McKeown's student advising transcripts [McKeown 85], Philip Werner's medical protocols [Werner], and Ethel Schuster's cooking instruction transcripts [Schuster 82], as well as a number of individual tokens. Although, for the sake of brevity, many of the tokens presented in the thesis consist of question/ answer pairs, the other protocols permit the study of scalar implicature in a larger discourse context.

While the claim that speakers are in fact implicating some meaning must often be based upon post hoc intuitions, subsequent discourse often provides convincing evidence that one or both conversational participants has indeed considered the implicitly conveyed meaning to be part of what the speaker has conveyed. For example, in the following conversation between a caller (A) and a hospital clerk (B),

- (2) A: Do you have information on a patient?
 B: What's the name?
 A: K___ M___ for maternity.
 B: I don't think she's delivered yet.
 A: Then she HAS been admitted.
 B: Yes.

the caller requests confirmation of the implicated 'K___ M___ has been admitted'. Similarly, in 3, A makes B's implicature explicit

- (3) B: Do you know who's taking your place down here?
 A: You told me -- x is.
 B: Yeah.
 A: Are you upset about that.
 B: I'm not upset.
 A: Do you mean to imply that other people are?

by her question. Gazdar [Gazdar 79a] notes a similar phenomenon in his example (presented here in 4

- (4) A: Is your mother well and back?
 B: Well she's back, yes.
 A: She's not well then.

These and other examples presented below provide an empirical base for the account I will present of the sorts of meanings speakers can convey via scalar implicature and the linguistic and non-linguistic contexts in which they can convey them. This description in turn forms the basis of a computational account of the phenomenon. While it is difficult to avoid at least implicit claims to cognitive reality in such an enterprise, I have tried to limit such claims to the following: People are able to generate and interpret meanings of the sort I term scalar implicatures. While other characterizations of these meanings and accounts of how they might be communicated are indeed possible, that presented below does allow us to simulate this generation and interpretation computationally.

A computational account of one type of conversational implicature of necessity involves some re-examination of the broader phenomenon. In particular, no sufficient conditions have been proposed for conversational implicature -- even in non-computational work -- and those necessary conditions suggested in the literature are the subject of considerable debate. In fact, only one serious attempt at a formal account of conversational implicature has so far been made [Gazdar 79a] -- for a small subset of the class. So, in defining scalar implicature, I must first develop a formal account of conversational implicature. In Chapter 2 of this thesis, I examine some of the major controversies in the linguistic literature on conversational implicature and some of the obstacles to a satisfactory definition of the phenomenon. I propose that a conversational implicature p_j is licensed by a speaker's utterance of some u_i when a

speaker intends to convey p_j by saying u_i ; when speaker and hearer mutually believe that speaker is being cooperative and that belief in p_j is 'required' given u_i and speaker cooperativity; and when p_j is CANCELABLE, NONDETACHABLE, and REINFORCEABLE.

In Chapter 3, I examine previous attempts to identify sub-classes of conversational implicature, and, in particular, the work of Horn [Horn 72], Harnish [Harnish 79], and Gazdar [Gazdar 79a], which I draw upon in defining scalar implicature. Scalar implicature in effect subsumes the phenomenon identified by Horn and Gazdar, while including a wider variety of inferences which I argue should be seen as members of a single class. I then introduce scalar implicature and relate it to these previous descriptions. In particular, I present a new strategy for utterance ranking and propose a redefinition of the epistemic force of licensed implicatures; both of these contributions respond to problems long recognized but unresolved in the literature. From an intuitive description of scalar implicature, I propose a set of scalar implicature conventions which, for a given utterance and given salient ordering(s), permit the identification of licenseable scalar implicatures.

In Chapters 4-6, I show how these conventions may be used to calculate scalar implicatures. Chapter 4 defines a representation for utterances and implicatures and justifies the epistemic force I assign to scalar implicatures. In Chapter 5, I describe the types of ordering relation that can support scalar implicature and provide a definition of utterance ranking in terms of partial ordering relations.⁸ In Chapter 6, I relate the conventional aspects of scalar implicature to the interpretation of particular inferences in context. I discuss how current notions of discourse FOCUS/ TOPIC/ CENTERING can be adapted to the task of identifying salient expressions and salient partial orderings in a discourse. Finally, I extend the scalar implicature formalism developed in earlier chapters to accommodate more complex implicatures arising from utterances in which more than one expressions is salient.

While various existing systems -- COOP [Kaplan 79] and HAM-ANS [Hoepfner 84a], for example -- have implicitly recognized the need to accommodate some aspects of the behavior examined in this thesis, efforts to interpret or generate such meanings have had limited success. In Chapter 7, I argue that such failures have occurred in large part because of the lack of an independently motivated account of the phenomenon. I describe one application of the theory of scalar implicature developed in previous chapters to computer-human interaction -- a system that proposes alternate responses to yes-no questions based upon the potential scalar

⁸In previous work on scalar implicature I have termed such posets 'scales' after Horn and Gazdar [Hirschberg 84a, Hirschberg 84b, Hirschberg 84c, Hirschberg 84d]. However, confusion over intuitive definitions of scales as linear orderings has convinced me to abandon this terminology. Mea culpa.

implicatures which direct responses might license. I present sample sessions in which this system, QUASI, proposes responses to questions about switches in a computer mail domain. In the conclusion to the thesis, I suggest further ways in which the calculation of scalar implicature should be useful in natural-language processing. For natural-language processing, knowledge of meanings such as those conveyed via conversational implicature is essential to provide a theoretical basis for more sophisticated models of human behavior and to permit machine inference of all that natural-language input may convey -- and machine anticipation of all that generated text may communicate.

CHAPTER II

Conversational Implicature in the Gricean Framework

One day in class, a student asked the rabbi what he would rather have, five daughters or \$5000. 'Five daughters,' answered the rabbi. 'Oh, come now, Rabbi, that's not possible,' protested the student. 'Well,' explained the rabbi, 'in the first place, if I had \$5000, I would no doubt want more, and greed is a terrible sin. Secondly, should I desire \$5000, I still wouldn't get it--wishing doesn't make things so--and I would look like a fool. And, thirdly, young man, I would rather have five daughters because I actually have eight.'

In this chapter I will first motivate the study of conversational implicature for computational linguistics. Next, I will examine previous attempts to define the phenomenon and identify certain necessary conditions on the phenomenon. I will then claim that these conditions represent sufficient conditions for conversational implicature. Based upon this definition, I will propose a formal account of conversational implicature that will form the basis in subsequent chapters for the calculation of SCALAR IMPLICATURES.

2.1. A Problem for Semantics

Linguists and philosophers of language have long noted the inadequacy of a truth-functional semantics to capture certain aspects of utterance meaning. For example, each of the following utterances may, in some contexts, convey more than a speaker actually *says*:

- (5)
- a. George went to jail and became a criminal.
 - b. Ellen is twenty-one.

Truth-functional semantic theories cannot differentiate between the utterances in 5 and the corresponding utterances in 6, even though the members of each pair may license quite different additional 'meanings'.

- (6)
- a. George became a criminal and went to jail.
 - b. Ellen is at least twenty-one.

So, (5a) and (6a) will have the same truth conditions: Each is true just in case both conjuncts are true. But the asymmetric interpretation of conjunction in natural language (See Section 3.1.1.1.) permits the utterer of (5a) to convey a very different meaning than s/he can convey by uttering (6a): By (5a), a speaker may convey that George's going to jail resulted in his corruption, which cannot be conveyed via (6a); but (6a), and not (5a), can convey that George's criminality led to his imprisonment. The utterer of (5b) may convey that, as far as s/he knows, Ellen is no older than twenty-one, while the utterer of (6b) cannot. However, if I say (5a) and you later find out that George became a criminal before he went to jail, you will probably not say that I have lied, although you may feel I have misled you; similarly for (5b) and (6b).

Note also that these non-truth-functional meanings will not always be licensed by the utterance of the sentences in 5. For example, the utterances in (7a) and (7b) are not likely to license the meanings licensable by

- (7)
- a. *George went to jail and became a criminal*, though not in that order.
 - b. *Who in this crowd can buy us some beer? Ellen is twenty-one.*

by the utterances presented in 5. So, these meanings are both non-truth-functional and context-dependent.

Context-dependent meanings pose obvious problems for computational linguistics. The simple lexical solution -- redefining lexical entries for items such as *and* to include an additional temporal/causal meaning -- is impossible just because of the context dependency of this meaning. A more reasonable alternative -- adding new senses for each context-dependent meaning which the use of a lexical item may license -- is also unsatisfactory: First, since the number of items which can be used to license such meanings is large, this solution is undesirable from a practical point of view. Second, it entails that all the utterances in 5, 6, and 7 will always be treated as ambiguous, which seems both counter-intuitive and computationally expensive. While designating a 'default' meaning might override these objections, it is not clear how such a default might be chosen. (See Section 3.1.1.1.) And including non-truth-functional meanings in the lexicon is in general somewhat risky. Third, it leaves unanswered the central problem of how this proliferating ambiguity might be disambiguated.

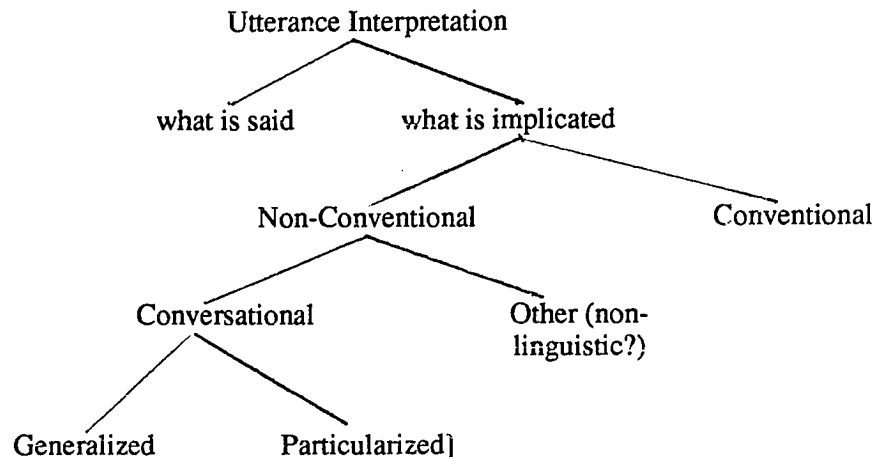
2.2. The Gricean Framework

As an alternative to such lexico-semantic approaches, Grice [Grice 75, Grice 78] proposes a solution within the framework of linguistic PRAGMATICS.⁹ Grice retains a simple lexicon, as well as a truth-functional semantics for the natural-language counterparts of the operators, quantifiers, and connectives of the first-order predicate calculus such as *and*. He accounts for discrepancies between natural language and its formal representation -- i.e., for non-truth-functional, context-dependent aspects of utterance interpretation -- as CONVERSATIONAL IMPLICATURES.¹⁰

2.2.1. Classifying Meanings

In his broad view of utterance interpretation, Grice distinguishes *what is said* from *what is implicated*, as illustrated in Figure 2-1.

Figure 2-1: The Gricean Framework



What is said is, for Grice, the CONVENTIONAL FORCE of an utterance -- those context-independent meanings that follow from the conventional meaning of an utterance and which

⁹Earlier work on contextual or pragmatic inference is summarized in [Hungerland 60].

¹⁰Grice introduces the verb *implicate* to denote the licensing of these particular understandings as well as the nouns IMPLICATURE (the act of implicating) and IMPLICATUM (that which is implicated). These terms are intended primarily to distinguish these non-truth-functional inferences from those that may logically follow from an utterance via standard rules of deduction operating on its semantic representation. [Grice 75:41-42]

determine its truth conditions. *What is implicated* represents those additional aspects of utterance interpretation that do not affect its truth conditions. CONVENTIONAL IMPLICATURES are non-deductive inferences derived from the conventional meaning of utterances; thus, while conventional implicatures are non-truth-functional, they are context-independent. For example, for Grice, lexical items such as *but* and *therefore* license conventional implicatures, as in (8a) and (8b).¹¹

- (8)
- a. He is an Englishman; he is, therefore, brave.
 - b. He's a New Yorker but I like him.

Conventional implicatures form part of what is implicated rather than what is said chiefly because their meaning is not captured in a truth-functional semantic representation.¹² For example, in Grice's view, while the proposition in (8a) will be false if the male referent is, say, French, it will not, strictly speaking, be false if the consequential connection between being an Englishman and being brave fails to hold.

In this framework, meanings such as those discussed in Section 2.1 form part of the class of CONVERSATIONAL IMPLICATURES. Like conventional implicatures, conversational implicatures are non-truth-functional; however, unlike conventional implicatures, they are context-dependent.

2.2.2. Formalizing Conversational Implicature

In defining conversational implicature, I will employ a representation which assumes a modal logic which includes operators expressing intention, belief, and mutual belief. Here I will present a partial syntax for this representation, which will be augmented in this and subsequent chapters. In the following discussion, assume that p_i, p_j, \dots are variables ranging over well-formed formulas (wffs); u_i, u_j, \dots range over utterances; S denotes some speaker and H , some hearer. C_h, C_i, \dots range over contexts, which, for the moment, may be seen simply as temporally ordered sequences of utterances in the current discourse.¹³ So, for present purposes,

¹¹Note that, while it is now generally accepted that such items convey their meaning via conventional implicature, for most such items what that meaning is has not been established. So, while it might seem reasonable to paraphrase *but*'s contribution as *X and, surprisingly enough, Y*, no such paraphrase has been generally agreed upon in the literature. The standard interpretation of 'but' is that it conveys a sense that something is 'contrary to expectation'.

¹²Unsuccessful proposals have been made to incorporate conventional implicature in the realm of semantics, as, for example, [Karttunen 79].

¹³A complete representation of *context* will not be attempted in this thesis; presumably, the computation of other types of conversational implicature may require additional contextual information.

C_j will be represented as $[\dots, u_i, u_j]$, where u_j is the utterance immediately prior to the current utterance, u_k -- the utterance whose meanings are currently in question. u_i is the utterance immediately prior to u_j , and so on. Note that the index of the current context C_j is identical to the index of the immediately prior utterance u_j . Given this simplification, contexts will be updated simply by the concatenation of u_k to C_j to form C_k .

The following are wffs in this representation:

$\neg p_i$

$p_i \vee p_j$

SAY(S, H, u_i, C_h): S says u_i to H in C_h .

REALIZE(u_i, p_i): u_i realizes p_i .

CAUSE(p_i, p_j): p_i causes p_j .¹⁴

INTEND(S, p_i): S intends to do p_i or intends that p_i hold.¹⁵

KNOW(S, p_i): S knows that p_i .

BEL(S, p_i): p_i follows from S 's beliefs.

BMB(S, H, p_i): the proposition represented by p_i follows from S 's beliefs about what is mutually believed by S and H .

Below I will assume that $p_i \wedge p_j$ abbreviates $\neg(\neg p_i \vee \neg p_j)$, that $p_i \Rightarrow p_j$ corresponds to $\neg p_i \vee p_j$, and that $(p_i \leftrightarrow p_j)$ abbreviates $(p_i \Rightarrow p_j) \wedge (p_j \Rightarrow p_i)$.

Note that REALIZE identifies an utterance with the semantic representation of the proposition it realizes. For computational purposes, then, REALIZE can be seen as taking an utterance and returning its representation in some semantic theory (or vice versa, for generation). I have not distinguished here among propositions, events, and actions, although clearly a satisfactory definition of CAUSE and INTEND requires such a distinction. CAUSE should probably relate one action to another, while INTEND should probably relate an agent either to an action s/he intends to perform or to a proposition she intends to make hold.¹⁶ KNOW and BEL are epistemic operators axiomatized as for [Hintikka 62]'s K and B operators. In particular,

Knowledge as True Belief

$\text{KNOW}(S, p_i) \leftrightarrow \text{BEL}(S, p_i) \wedge p_i$.

¹⁴The use of propositions as objects of causality is discussed below.

¹⁵The use of propositions as objects of intention is discussed below.

¹⁶Despite this current simplification, there is a simple mapping between the use of CAUSE and INTEND below and a more satisfying definition of each: In every case presented below, INTEND will be associated with an agent and a proposition which might better be construed as an event, while CAUSE will relate two propositions which might also be represented as actions.

BMB is the one-sided mutual belief of [Clark 81, Perrault 81]: A speaker (S) believes that P is mutually believed between S and hearer (H) when

$$MB(S,H,P) = SB(P) \wedge SBHB(P) \wedge SBHBSB(P) \wedge \dots$$

Similarly

$$MB(H,S,P) = HB(P) \wedge HBSE(P) \wedge HBSBHB(P) \wedge \dots$$

A proposition P is mutually believed by both H and S (two-sided mutual belief) if both MB(S,H,P) and MB(H,S,P) hold. One-sided mutual belief (MB(S,H,P), or, here, BMB(S, H, p_i)) is equivalent to $SB(P) \wedge SB(MB(H,S,P))$; hence, BMB is defined equivalently as:

Mutual Belief:

$$BMB(S, H, p_i) \leftrightarrow BEL(S, p_i) \wedge BEL(S, BMB(H, S, p_i))$$

I will employ this syntax throughout the thesis, translating from that of other authors where possible to avoid proliferation of notational systems. Also note that, for the sake of simplicity, all variables will be assumed to be universally quantified unless otherwise specified.

2.3. Maxims of Cooperative Conversation

In Gricean theory, participants in conversation assume that their partners share knowledge of certain underlying domain- and exchange-independent conversational goals, which Grice subsumes under his COOPERATIVE PRINCIPLE (CP):

Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged. [Grice 75:45]

Because the CP is shared knowledge, speakers can license and hearers can interpret inferences beyond the conventional force of an utterance, in large part by comparing 'what is said' to 'what might be said' in the exchange.

Grice distinguishes four maxims of cooperative conversation, which further specify what it means to observe the CP, the MAXIMS OF QUANTITY, QUALITY, RELATION, and MANNER.¹⁷ These maxims are "standardly (though not invariably) observed by participants in a talk exchange." [Grice 78:113] While Grice does not claim that these four exhaust the notion of communicative cooperativity, so far the pragmatic literature has, by and large, accepted these as (collectively) sufficient conditions on the CP.¹⁸

¹⁷Grice [Grice 75:47] notes that there are other maxims of an aesthetic, social, or moral nature, such as 'Be polite'; he does not include these with his conversational maxims, since they are not so specifically connected with the particular goals of a talk exchange. These provide the basis for those nonconventional but nonconversational implicatures noted in the figure above.

¹⁸Attempts have been made to order these maxims, to identify one (Relation or Quality) as primary, subsuming the rest, and to define sub-maxims. See, for example, [Sperber 83, Kempson 75].

To permit formalization of speaker obedience of these maxims, I will introduce variables m_i, m_j, \dots which range over individual (labels for) maxims and M_i, M_j, \dots which range over sets of m_i . I will also define for each of the standard maxims what it means to observe that maxim via the predicate OBEY, where $OBEY(S, m_i, C_h)$ indicates that S is observing the maxim m_i in the context C_h . This predicate will later permit a general definition of 'speaker cooperativity' in the Gricean framework.

Grice's original MAXIM OF QUALITY,

- Try to make your contribution one that is true.
1. Do not say what you believe to be false.
 2. Do not say that for which you lack adequate evidence. [Grice 75:46]

enjoins speakers from providing information for which they have insufficient evidence and instructs them to consider the effect of their contribution in assessing its truthfulness. That is, S obeys the Maxim of Quality iff propositions not believed true are not stated. Or, equivalently, S is obeying this maxim iff everything S says s/he also believes to be true:

Obeying Grice's Maxim of Quality

$$((SAY(S, H, u_i, C_h) \wedge REALIZE(u_i, p_i)) \Rightarrow BEL(S, p_i)) \leftrightarrow OBEY(S, \text{Quality}, C_h)$$

However, Joshi's [Joshi 82:190] revision of this maxim presents a more difficult task. Joshi defines Quality in terms of S and H 's MUTUAL BELIEFS¹⁹ as follows:

- if S says p_i , then it must be that S believes p_i ;
- based on S 's assessment of S and H 's mutual beliefs, it should not be possible for H , from what has been said (p_i), to infer some other fact (say p_j) which S knows to be false;
- if there is such a possibility, then, after saying p_i , S should add further information (e.g., $\neg p_j$) to "square away" the mutual beliefs;
- otherwise, S is likely to mislead H by possibly making him continue to believe in p_j , which is false in the knowledge base.

In other words, to obey Joshi's Maxim of Quality, S should not only speak the truth and avoid the false, but should avoid truths which might lead H to draw false conclusions. This revision might be defined as:

Obeying Joshi's Revised Maxim of Quality

$$((SAY(S, H, u_i, C_h) \wedge REALIZE(u_i, p_i)) \Rightarrow BEL(S, p_i)) \wedge (BEL(S, CAUSE(SAY(S, H, u_i, C_h), BEL(H, p_j))) \wedge SAY(S, H, u_i, C_i) \wedge$$

¹⁹Here I have changed only the notation for S and H .

$$\text{BEL}(S, \neg p_j) \Rightarrow (\text{SAY}(S, u_j) \wedge \text{REALIZE}(u_j, \neg p_j))^{20} \leftrightarrow \text{OBEY}(S, \text{Quality}_{\text{Joshi}}, C_h)$$

If S is obeying $\text{Quality}_{\text{Joshi}}$, then not only will S believe what s/he says, but, also, if S believes that the saying of u_i will license belief in p_j -- and if S does indeed say u_i while believing p_j to be false -- then S will also say some u_j which realizes $\neg p_j$. For the remainder of this work, reference to the Maxim of Quality should be understood as reference to $\text{Quality}_{\text{Joshi}}$.

While the Maxim of Quality applies only to assertions,²¹ suggestions for extending it to queries, imperatives, and other speech acts have been made [Kempson 75]. For example, if a speaker observing this maxim asks a question, that speaker may be presumed not to know the answer to that question.

Grice's remaining maxims are less easily formalized. The MAXIM OF QUANTITY,

- a) Make your contribution as informative as is required (for the current purposes of the exchange).
- b) Do not make your contribution more informative than is required. [Grice 75:45]

requires that S select a level of informativeness appropriate to the exchange. [Fogelin 67] paraphrases the first part of this maxim (implicitly in terms of the Maxim of Quality as well) as "Make the strongest possible claim that you can legitimately defend." However, it is not clear how to quantify 'informativeness' or 'strength of claim' to permit comparison of potential utterances. Furthermore, specifying what the 'required' or 'appropriate' level of such informativeness is in some C_h seems impossible to do in any general way. While I will examine these questions in some detail in Chapters 3 and 5 and in Section 6.3.2.1, $\text{OBEY}(S, \text{Quantity}, C_h)$ must remain, for now, intuitively defined.

The same is true for a notion of obedience to Grice's MAXIM OF RELEVANCE, $\text{OBEY}(S, \text{Relation}, C_h)$. Relation is summarized as: Be relevant. Whatever information S provides should be relevant to the exchange. Grice himself [Grice 68] has suggested that this maxim might subsume part (b) of the Maxim of Quantity. Others have gone further to propose that this maxim may in fact subsume all Grice's other maxims; if information is *relevant*, they claim, then clearly it will also be truthful and appropriately detailed. [Horn 84, Sperber 83]

Finally, Grice's MAXIM OF MANNER,

²⁰REALIZE is actually too strong here, since u_j may simply convey $\neg p_j$ by means other than assertion. In Section 2.4.2, I will introduce the concept of LICENSING belief which will be substituted in this definition.

²¹Cf. [Gazdar 79a:48].

Be perspicuous.

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief (avoid unnecessary prolixity).
4. Be orderly.

requires that S provide information in a format appropriate to H and to the circumstances of the exchange. Gazdar [Gazdar 79a] notes that S 's observance of this maxim accounts for the implicatures associated with ASYMMETRIC *and*, as in (5a) and (6a) above.²² More specifically, Harnish [Harnish 79] suggests the following submaxims of Manner:

Super Submaxim: Be representational; in so far as possible, make your sayings "mirror" the world.

SubMaxim of Time: In so far as possible, make the order of saying reflect the order of events.

SubMaxim of Space: In so far as possible, if objects $a, b, c, \dots \phi$ together, put their names together when reporting this ϕ -ing.

Some attempts toward formal accounts of this maxim -- particularly, of its submaxims -- have been noted in [Gazdar 79a:43-44]. However, since these formulations lack precision and account only for parts of the behavior encompassed by this maxim, I will assume that OBEY(S, C_h) too is intuitively defined.

Given these notions of what it means for speakers to obey the Gricean maxims, we can now define Grice's more general concept of speaker cooperativity in terms of obedience to them. Because it will be useful to talk about speaker observance of particular maxims or sets of such maxims, I will define cooperativity as follows:

Speaker Cooperativity:

$$\forall m_i \in M_i \{ \text{OBEY}(S, m_i, C_h) \leftrightarrow \text{IS_COOP}(S, C_h, M_i) \}$$

That is, S is being cooperative with respect to maxims M_i iff S is obeying each $m_i \in M_i$. Given this definition, we can specify Grice's notion of speaker cooperativity as just speaker obedience to the Maxims of Quality, Quantity, Relation, and Manner: IS_COOP($S, C_h, \{ \text{QUALITY, QUANTITY, RELATION, MANNER} \}$).

Speakers may sometimes FLOUT, or ostensibly violate, one or more maxims, intending thereby to communicate some additional meaning to H . So, in Grice's classic example, a philosophy professor flouts the Maxim of Quantity to implicate a poor opinion of a pupil applying for a position in Philosophy:

²²I.e., the surface order of conjuncts reflects the temporal order of events mentioned.

- (9) Dear Sir,
 Mr. X's command of English is excellent, and his attendance at
 tutorials has been regular.
 Yours, etc.

It might seem that the Maxim of Quantity is being violated here, since letters of recommendation for philosophers might be expected to contain some (favorable) reference to a pupil's skills in that field. Yet Grice explains that *H* interprets *S*'s behavior as cooperative: If *S* is in fact observing the CP, then, in particular, *S* has said as much as he truthfully can (that is favorable) about his/her pupil's skills. If *S* says nothing favorable in this regard, then *H* may conclude that *S* is not able truthfully to comment favorably on his/her pupil *qua* philosopher. Since *S* can anticipate this conclusion on *H*'s part, *S* can in fact convey to *H* 'I can say nothing favorable about Mr. X as a philosopher' via 9. So, an apparent violation or flouting of the CP in fact represents the fulfilment of *S*'s obligation to observe it.²³ And *S* relies upon *H*'s belief that *S* is being cooperative in order to convey such implicit meanings.

Clearly, however, speakers do not always obey these maxims either implicitly or ostensively. *S* may *violate* the maxims, thus misleading *H*; for example, *S* may lie (violating the Maxim of Quality) or understate (violating the Maxim of Quantity). However, it is significant that the very notions of lying or understatement rely for their definition upon the notion that these maxims represent the 'norm'. *S* may also *opt out* of the maxims, making it plain that s/he is **not** being cooperative: Watergate-style 'stonewalling' exemplifies this option. *S* may also be faced with a *clash* between two or more maxims and be unable to fulfill all of them. Consider 10, where

- (10) A: Where does C live? (= [Grice 75]'s (3))
 B: Somewhere in the south of France.

A is planning with B an itinerary for a holiday in France. Both know A wants to visit his friend C. By his response, B implicates he does not know C's precise whereabouts. He is faced with a clash between the maxims of Quantity and Quality: to obey Quantity, he should supply a more

²³Although some have claimed [Harnish 79] that meanings licensed via the flouting of a maxim must be licensed differently from those licensed by its observance, this claim seems to miss Grice's notion that one may follow the CP less directly in these cases. [Harnish 79] terms conversational implicatures licensed via obedience to the maxims IMPLICATURE SIMPLICITER.

specific location, but to obey Quality, he should say only what he has adequate evidence for.²⁴

For Grice, the CP and its maxims are to be viewed as part of a speaker's communicative competence: Although our observance of them may be implicit, "...our linguistic practice is *as if* we accepted these rules and consciously followed them." [Grice 68] For Morgan, the CP and its maxims are among his CONVENTIONS OF LANGUAGE USE; they represent "rules for inferring the intentions behind speech acts, or, from the speaker's viewpoint, for selecting one's utterances so as to convey one's intentions, by exploiting the maxims." [Morgan 78:262] The 'meanings' speakers exploit these maxims to license are, then, Grice's conversational implicatures.

2.4. Conversational Implicature

While the notion of conversational implicature is not hard to grasp intuitively, it has proven difficult to define precisely. The generality of the CP and the conversational maxims makes it difficult to specify just what is being observed or flouted in a particular instance, and analyses of which maxims are involved in particular implicatures vary widely from observer to observer. While Grice proposed several characterizations of conversational implicature, each has been shown to have serious limitations. Subsequent efforts [Thomason 73, Walker 75, Katz 76] to formulate alternative definitions appear even less useful. So, it is not surprising that only one serious attempt at a computational approach to conversational implicature has been made [Gazdar 79a], and that this attempt eschews a definition of the general phenomenon.

For Griceans, conversational implicatures are those non-truth-functional, context dependent aspects of utterance interpretation which depend critically upon *S* and *H*'s (implicit) recognition of the CP and the conversational maxims for their conveyance. So, for example, if B says (11a) and assumes that A believes B is obeying the CP,

- (11)
- a. I have five dollars.
 - b. I have more than five dollars.

²⁴To explain *S*'s behavior when faced with such clashes, [Harnish 79] proposes a PRINCIPLE OF CHARITY: "Other things being equal, construe the speaker's remark so as to violate as few maxims as possible". He and other authors have suggested that the maxims in fact should be ordered by the degree to which they are important to successful communication; if such an ordering is correctly identified, then speakers' behavior when faced with clashes should be infringement of the lower-ranked maxims before the higher-ranked maxims. However, there is little agreement on how the maxims should be ordered. Some [Sperber 83, Horn 84] see Relation as primary: whatever *S* says must first of all be relevant to the conversation. Alternatively, Grice's [Grice 75] argument that Quality (or, truthfulness) is probably *S*'s first consideration is also persuasive, particularly in cases like 10.

B may license the inference ‘for all B knows $\neg(11b)$ ’.²⁵ This inference is licensed via the shared assumption that a cooperative speaker says as much (Maxim of Quantity) as s/he truthfully (Maxim of Quality) can that is relevant (Maxim of Relation) for the particular exchange. So, if B believes his/her total readily available cash is relevant, then by choosing to affirm *five* when s/he might, with equal brevity, affirm some higher value x entailing *five* (in the sense that, if one has six dollars then one also has five dollars), B licenses the conclusion that s/he is unable truthfully to affirm x . Either s/he believes s/he does not have x dollars or s/he does not know whether s/he has x dollars. Hence, in this instance, B implicates $\neg(11b)$ by asserting (11a).

A single utterance may license multiple implicatures. For example, B’s response in 12 may convey

- (12) A: I don’t have enough money for lunch.
B: I have five dollars.

that B is willing to lend A lunch money, as well as $\neg(11b)$.

2.4.1. Implicature and Speaker Intention

Although Grice’s general work on utterance meaning makes it clear that he intends conversational implicature, as all aspects of utterance meaning, to be defined in terms of speaker intention, confusion about this point is evident throughout the literature on implicature. That is, conversational implicatures are often spoken of as meanings which H infers in a certain way, rather than as meanings S seeks to communicate in a certain way. Of course, even if we define conversational implicature in terms of S ’s intentions, we can also talk about H ’s beliefs about those intentions as a condition on H ’s inference that S has implicated that p_j (See the definition of inferred conversational implicature in Section 2.4.4.); but this does not change the need to choose the point of view from which one will define the phenomenon itself.

Grice identifies conversational implicature as part of the *meaning_{nm}* (‘nonnatural meaning’) of an utterance [Grice 69]. *Meaning_{nm}* is distinct from ‘natural meaning’ in that it is not conveyed directly to H ,²⁶ who must consequently do a certain amount of inferencing to

²⁵See Section 5.1.4.1 for a fuller discussion of conversational implicatures licensed via mention of cardinals.

²⁶In response to criticism that his account of *meaning_{nm}* failed for audienceless utterances, Grice substituted ‘addressee’ for ‘hearer’. Although it is theoretically possible that S might implicate p_j to him/herself, this possibility will be ignored here, so I will continue to use ‘hearer’ instead of addressee.

interpret *S*'s utterance.²⁷ More usefully, *meaning_{nn}* differs from natural meaning in that it is a function of *S*'s intention and of *S*'s beliefs about *H*'s recognition of this intention, while intentions do not play this determining role in cases of natural meaning. For his utterance to convey some *meaning_{nn}*, *S* must intend to produce a certain effect in *H*, which will often be the recognition of some belief; Grice labels this intending M-INTENDING. *S* must also intend this effect to result (at least in part) because of *H*'s recognition of *S*'s intention. However, Grice notes that neither *S* nor *H* need be "aware of" these intentions *qua* intentions [Grice 69].

While the concept of *meaning_{nn}* has been the subject of considerable debate,²⁸ Grice's discussion does make clear his position that aspects of *meaning_{nn}*, in particular, conversational implicature, should be defined in terms of *S*'s intentions rather than *H*'s beliefs about those intentions. *S*'s implicating *p_j* is independent of *H*'s understanding that *S* has implicated *p_j* in the same way that *S*'s asserting *p_i* is independent of *H*'s understanding that *S* has asserted *p_i* -- or that *S*'s implying *p_j* by asserting *p_i* is independent of *H*'s recognizing that *p_i* logically implies *p_j*. In neither of these latter cases would we want to say that a meaning has been asserted or implied simply because *H* has understood it or because *H* believes that *S* has asserted or implied it. Similarly, then, if *H* infers some *p_j* that *S* has not intended, even if *H* believes that *S* has implicated *p_j*, we would not want to say that *S* has implicated *p_j*. So, *S* has not implicated *p_j* unless s/he has intended to implicate *p_j*.²⁹ Conversely, *S* implicates *p_j* just in case s/he intends to implicate *p_j* -- regardless of whether *H* understands it.

Curiously, even Grice does not always maintain this distinction between implicating and understanding what has been implicated, as is evident from his identification of conversational implicature with the lexical items involved its licensing (See Section 3.1.1) and from his

²⁷ [Wright 75] notes that this distinction is not dichotomous but should be seen on a continuum.

²⁸ Grice originally claimed that speaker intention was prior to utterance and even sentence meaning. However, as many [Searle 69, Kempson 75, Wright 75, Yu 79] have noted, his account is circular: Some independent aspect of sentence meaning must surely constrain the set of communicative intentions speakers may have when uttering that sentence, or we could not explain how speakers choose among possible utterances. That is, if I want to tell you the time, I am constrained in the ways I may communicate the time to you by more than merely my intention to tell you the time. I will not usually say 'It is raining' to tell you that it is five o'clock, for example, although, if speaker intention were prior to sentence meaning, there should be no way for me to choose between 'It is raining' and 'It is five o'clock'. So, most Griceans, including, apparently, Grice himself, have abandoned this claim, while holding to some weaker claim for the centrality of speaker intention in utterance meaning. [Kempson 75:138-141, Grice 78:120, Wright 75, Yu 79]

²⁹ This position appears to accord with [Walker 75:157]'s definition of conversational implicature:

S conversationally implicates ϕ by μ_i iff in uttering μ_i *S* M-intends to convey ϕ to his audience, and intends this intention to be recognised partly because of the audience's recognition of the sense of μ_i (together with its expectation that *S* also knows the sense of μ_i); but partly also because the audience expects *S* to be conforming to the Cooperative Principle, and expects *S* to anticipate this expectation and to act accordingly.

discussion of implicature CANCELLATION (See Section 2.4.2.1). However, while it is true that *S* will take (his/her beliefs about) *H*'s inferential capacity into account in planning to implicate p_j , and while it is certainly possible to define what it means for *H* to infer that *S* has implicated p_j , Grice's belief in the primacy of speaker intention to utterance meaning suggests that what it means for *S* to implicate p_j is best defined from *S*'s point of view. So, as a beginning to a definition of conversational implicature, I will specify that *S* IMPLICATES p_j to *H* by saying u_i only when:

Condition 1. *S* intends to convey p_j to *H* by saying u_i .

That is, *S* can be said to implicate p_j via u_i only if

Intention to Implicate:

INTEND(*S*, CAUSE(SAY(*S*, *H*, u_i , C_h), BEL(*H*, p_j)))

Note that it is not sufficient to say that *S* intends to convey p_j -- but that s/he must intend to convey p_j by saying u_i if we are to count p_j a conversational implicature. That is, conversational implicatures are confined to those meanings licensed via linguistic behavior. Note also that, while Condition 1 is a necessary condition on conversational implicature it is obviously not sufficient; for Grice, all utterance meaning must hinge upon *S* intention.³⁰

2.4.2. Other Necessary Conditions on Conversational Implicature

While Condition 1 is derived from Grice's more general work on utterance meaning, Grice himself proposes other necessary conditions on conversational implicature. However, these suggestions are flawed -- chiefly by Grice's failure to be specific about the point of view from which each is defined. [Grice 75:49-50] presents the following characterization:³¹

A man who, by (in, when) saying (or making as if to say) that p_i has implicated that p_j , may be said to have conversationally implicated that p_j PROVIDED THAT

G1. he is to be presumed to be observing the conversational maxims, or at least the cooperative principle;

G2. the supposition that he is aware that, or thinks that, p_j is required in order to make his saying or making as if to say p_i (or doing so in THOSE terms) consistent with this presumption; and

G3. *S* thinks (and would expect *H* to think that *S* thinks) that it is within the competence of *H* to work out, or grasp intuitively, that the supposition mentioned in G2 is required.

³⁰In fact, if we could assume this view to be universal, we could define it for all utterance meaning, not simply conversational implicature.

³¹I will distinguish Grice's proposed necessary conditions on conversational implicature from my own by prefacing his with 'G'.

So, Grice would say *S* implicates p_j by saying p_i to *H* when *H* presumes that *S* is observing the CP; when this presumption plus *S*'s saying that p_i require one to conclude that *S* believes that p_j ; and when *S* thinks (and would expect *H* to think *S* thinks) that *H* can figure all this out.

Note however that the belief spaces postulated for each of Grice's conditions -- or the lack of such specification -- make it difficult to determine the point of view from which conversational implicature is to be defined. Grice does not specify who must presume or suppose conditions G1 and G2, although it seems that he means G1 to be a condition on *H*'s beliefs. However, given that condition G3 is a condition on *S*'s beliefs, condition G1 as stated allows conversational implicature to be defined in terms of *H*'s beliefs, when *S*'s may be at variance. That is, suppose you believe I am obeying the CP but I believe that you do not believe this. If, say, I believe you believe I am not observing the Maxim of Quality, then, when I assert p_i , I will not believe that I can implicate p_j , even though condition G1 holds. So, under this condition, implicatures will be defined which *S* does not believe have been licensed by an utterance, violating Grice's belief in the primacy of speaker intention to meaning -- my Condition 1 above. So it seems that this condition should be revised to require that *S* believe that *H* believe in *S*'s cooperativity.³²

Suppose then that $BEL(S, BEL(H, 'S \text{ is obeying the CP}'))$ but that it is not the case that $BEL(S, BEL(H, BEL(S, BEL(H, 'S \text{ is obeying the CP}'))))$. Perhaps I believe you believe me cooperative but I also believe you do not realize I believe this -- $BEL(S, \neg BEL(H, BEL(S, BEL(H, 'S \text{ is obeying the CP}'))))$. Imagine that you and I have just quarreled. In the midst of the quarrel, you call me a liar and are instantly contrite. Since we have been long-time friends, I do not believe that you believe me a liar. But when I see how upset you are at your accusation, I realize that you believe I believe you must believe me a liar. At this point the postulated conditions hold with regard to the Maxim of Quality: I believe you believe me truthful but I also believe you do not believe this. Then I will not believe that I can convey to you some implicature via your belief in my adherence to Quality, since I will not believe that this assumption can play any role in your 'working out' of an implicature.³³ Because G1 is thus sensitive to nested belief contexts, not only must we specify that $BEL(S, BEL(H, 'S \text{ is obeying the CP}'))$ but we must require that $BEL(S, BMB(H, S, 'S \text{ is obeying the CP}'))$.

³²Note that this condition will not require that *H* actually believe in *S*'s cooperativity.

³³While it is also possible that *S* might license certain inferences when s/he believes *H* believes *S* is not obeying the maxims, it seems difficult to propose any regularity in this lack and, thus, difficult to propose any inferences derivable from it.

With this revision, Grice's G1 may be restated as:

Condition 2. *S* believes that it is mutually believed between *S* and *H* that *S* is being cooperative.

Although it may be difficult to tell when *S* is observing the CP and the conversational maxims, as I have noted above, we can represent this observance via IS_COOP(*S*, *C_h*, {QUALITY, QUANTITY, RELATION, MANNER}).³⁴ Then Condition 2 may be represented as:

Mutual Belief in Speaker Cooperativity:

BEL(*S*, BMB(*H*, *S*, IS_COOP(*S*, *C_h*, {QUALITY, QUANTITY, RELATION, MANNER}))).

Next, consider how Condition G2 encapsulates the notion that some (non-deductive) argument *Arg* exists -- independent of *S* or *H* apparently -- which, taking *S*'s cooperativity and utterance of *p_i* as premises, yields *p_j* as a conclusion. That is, in Grice's terms, conversational implicature is calculable: There exists some *Arg* that represents a way of 'working out' that a cooperative *S*'s saying that *p_i* must lead *H* to conclude that *p_j*. Unfortunately, as well as failing to specify who must believe that *Arg* yields *p_j*, Grice fails to make clear the sense in which he intends that a non-deductive conclusion such as *p_j* be 'required', given *Arg*.

>From Grice's examples, it seems that condition G2 is often satisfied via *S*'s belief about *H*'s belief that other choices of utterance *S* might make would violate the maxims, i.e., be less informative or less truthful or less relevant or less appropriately presented than the utterance *S* has actually made. That is, procedures by which conversational implicatures are 'worked out' seem often to rely upon a comparison of an utterance with those other utterances *S* might have produced. In Chapter 3 I will pursue this notion in more detail. I will propose that, at least for general classes of conversational implicature, it is possible to specify ways in which this notion of 'required' can be defined. For the moment, I will introduce a predicate, LICENSE(*S*, *H*, *p_j*, *u_i*, *C_h*, *M_i*), which denotes that some inference represented by *p_j* is licensed by SAY(*S*, *H*, *u_i*, *C_h*) given IS_COOP(*S*, *C_h*, *M_i*).³⁵

It should be clear that Grice's notion of the inference of *p_j* being 'required' relies upon at least *S*, *H*, and *u_i*. LICENSE must also depend upon context, even though Grice fails in G2 to

³⁴Note that this will allow future change to the set of maxims that define speaker cooperativity if desired. It is also possible that conversational implicature does not require mutual belief that *S* is obeying all conversational maxims. Theoretically, it seems possible that *S* might implicate that *p_j* via mutual belief that *S* is observing only a subset *M_i* of the Gricean maxims, providing that the working out of *p_j* relies only upon mutual belief in *S* observance of *M_i*. I will return to this possibility in Chapter 3. In the meantime note only that this representation allows for future revision of our theory in this direction.

³⁵For meanings other than conversational implicatures, of course, *M_i* may be null.

specify how *Arg* can be seen as context dependent.³⁶ Given that the context-dependence of implicature is the motivating factor in the introduction of the concept itself, this is a serious omission. Note also that, while condition G1 may best be defined in terms of speaker obedience to all of the maxims in the CP, condition G2 need not be. Although it is not always possible to reach agreement on which maxims are involved in some implicature, unless it is possible to specify these maxims, it is clearly impossible to talk about the process by which implicatures are to be ‘worked out’. For example, we would not say that *S*’s obedience to the Maxim of Manner plays a role in the implicature of $\neg(11b)$ via the utterance of (11a). So it will be useful to specify for any implicature the maxims that are involved in the ‘requirement’ that some conclusion be reached; hence, the M_i argument to LICENSE.

Given this notion of LICENSE we can proceed to examine another major weakness in Grice’s description of condition G2, namely, his failure to specify the belief space in which $LICENSE(S, H, p_j, u_i, C_h, M_i)$ will hold. Also, while some implicata may be of the form $BEL(S, p_j)$, others, such as $\neg(11b)$ will be of the form $\neg BEL(S, p_j)$. That is, *S* may license belief in *S*’s belief or belief in *S*’s lack of belief via implicature. Given these considerations, and assuming for the moment that speaker cooperativity is defined as in my Condition 2, we might then redefine Grice’s condition G2 as something like:

$$BEL(S, LICENSE(S, H, p_j, u_i, C_h, M_i))$$

Note however that, ‘ $BEL(S, LICENSE(S, H, p_j, u_i, C_h, M_i))$ ’ may be true while ‘ $BEL(S, BEL(H, LICENSE(S, H, p_j, u_i, C_h, M_i)))$ ’ is false. That is, I may believe that my belief in p_j is required if I am cooperative and have uttered p_i , but at the same time fail to believe that you share this belief. In this case, I will not believe that my saying that p_i will convey to you p_j . For example, if I believe you believe I am lying, then my belief in the inferences that my utterance will license given that I am perceived as obeying the CP will not be of much use. So, condition G2 might be met, while *S* might believe that no implicature had been licensed. Similarly, if ‘ $BEL(S, BEL(H, LICENSE(S, H, p_j, u_i, C_h, M_i)))$ ’ is true but ‘ $BEL(S, BEL(H, BEL(S, LICENSE(S, H, p_j, u_i, C_h, M_i))))$ ’ is false, then I may believe that you believe my belief in p_j follows from my cooperativity and my utterance, but I may believe that you believe I do *not* believe this. Say, I tell you ‘*I have five dollars*’ when I believe you believe I believe you only need five dollars. Then I will not imagine that you will conclude that I have no more than five dollars. In effect, then, condition G2 must be redefined in terms of *S*’s beliefs about *S* and *H*’s

³⁶Thomason’s [Thomason 73] definition of conversational implicature is the first actually to specify contextual dependence:

A sentence ϕ conversationally implicates ψ relative to a class C of contexts of utterance, if for all $c \in C$, such that ϕ ’s assertion in c does not violate the maxims of conversation, ψ is presumed in c .

See [Gazdar 79a:41] for a discussion of the limitations of this definition.

mutual beliefs: that is, for S to implicate p_j to H by saying u_i in a context C_h , it is necessary that:

Condition 3. S believes that it is mutually believed between S and H that p_j is 'required' (or licensed) given that S is obeying some particular M_i and S says u_i in C_h .

I will represent this condition as:

Licensing Conversational Implicature:
 $BMB(S, H, LICENSE(S, H, p_j, u_i, C_h, M_i))$

Grice explains his condition G3, or, how S might envision H 'working out' condition G2, in terms of the following S belief about H 's reasoning:

S has said that p_j ; there is no reason to suppose that he is not observing the maxims, or at least the CP; he could not be doing this unless he thought that p_j ; he knows (and knows that I know that he knows) that I can see the supposition that he thinks that p_j is required; he has done nothing to stop me thinking that p_j ; he intends me to think, or is at least willing to allow me to think, that p_j ; and so he has implicated that p_j [Grice 75:50].

While S 's implicating that p_j is independent of H 's understanding that S has implicated that p_j , this condition expresses the notion that part of what it means for S to implicate p_j is this anticipation of H 's ability to 'work out' the implicature. While the importance of this condition has led to some confusion over the point of view from which implicature is defined, note that in fact S need only believe that H can 'work out' the implicatum in order to implicate p_j . In the end, G3 does not require that H actually 'work out' p_j for S to have implicated p_j .

Note however that my revised condition G2, ($BMB(S, H, LICENSE(S, H, p_j, u_i, C_h, M_i))$) subsumes that 'BEL($S, BEL(H, BEL(S, H$ can 'work out' that p_j is required for IS_COOP(S, C_h, M_i) given SAY(S, H, u_i, C_h)))'. By 'work out' it seems then that what Grice really means is 'come to believe'. Thus, 'BEL($S, BEL(H, BEL(S, H$ can 'work out'....))' is equivalent modulo tense to 'BEL($S, BEL(H, BEL(S, BEL(H, LICENSE(S, H, p_j, u_i, C_h, M_i))))'$. But the latter is subsumed by the mutual belief context of the revised condition G2. So, conditions G2 and G3 can be collapsed to a single condition on S and H 's mutual beliefs, i.e., my Condition 3.

So far, then, I have proposed three necessary conditions on conversational implicature, one derived from Grice's early work on intention and utterance meaning and two representing revisions of (three) necessary conditions Grice himself has proposed. That is, S may implicate p_j to H by saying u_i in a context C_h when:

1. INTEND($S, CAUSE(SAY(S, H, u_i, C_h), BEL(H, p_j))$);
2. BEL($S, BMB(H, S, IS_COOP(S, C_h, \{QUALITY, QUANTITY, RELATION, MANNER\}))$);
3. BMB($S, H, LICENSE(S, H, p_j, u_i, C_h, M_i)$)

That is, *S* may implicate some p_j by saying p_i to *H* in some discourse context C_h when: *S* intends to convey p_j to *H* by saying that u_i ; *S* believes it is mutually believed between *S* and *H* that *S* is being cooperative; and *S* believes it is mutually believed between *S* and *H* that p_j is somehow 'required' in the circumstance in which the *S* has said u_i in C_h and that *S* is obeying the maxims specified in M_i .

However, while Conditions 1-3 do represent necessary conditions for implicature in general, they do **not** distinguish conversational implicature from other types of implicature. In particular, CONVENTIONAL IMPLICATURE (See Section 2.2.1.) also exhibits the same features. These conditions are also insufficient to exclude entailment from conversational implicature. So, they do not yet provide the definition desired. However, Grice [Grice 75] has also proposed certain additional features of conversational implicature, some of which can be incorporated into a more precise definition of the phenomenon:

G4. a conversational implicature is CANCELABLE, explicitly or contextually;

G5. a conversational implicature is NONDETACHABLE;

G6. a conversational implicature is *not* part of the conventional force of the expression that gives rise to it;

G7. a conversational implicature is not carried by what is said but by the saying of it;

G8. a conversational implicature is often a disjunction of several possible interpretations of an utterance and is often indeterminate.

Grice's conditions G6 and G7, both intended to exclude entailments and conventional implicatures, do so only by fiat and thus exhibit some circularity. As Sadock [Sadock 78:284-285] notes, if it were possible to decide intuitively between the conventional and nonconventional aspects of an utterance, or between what is said and what is licensed by the saying of it, no additional criteria would be necessary. Condition G8 is clearly not a necessary condition even in Grice's original formulation; just as clearly it is not a sufficient condition. Thus the value these last three conditions in a definition of conversational implicature is small. However, conditions G4 and G5 have been adopted as standard diagnostics for conversational implicature. Despite considerable controversy over their status, I will maintain that they do differentiate conversational implicature from other types of utterance meaning.

2.4.2.1. Cancellation

According to Grice, a "putative conversational implicature" p_j can be explicitly CANCELED by the addition of expressions such as *but not p_j* or *I do not mean to imply that p_j* [Grice 78]. That is, the denial of that which is implicated without the denial of that which has been asserted constitutes a successful cancellation, as in 13:

- (13) The Adelphi apartment hotel that they were going to tear down and evict me from, *although not in that order*, had been built back in the early twenties about the time that the claw-footed bathtub was beginning to disappear from the American scene.
Oliver Bleeck, *No Questions Asked*, p.16. (My italics)

In this example, the implicit temporal ordering licensed by the use of asymmetric *and* is denied. Note that, in 13, there is no sense that the denial of the implicatum is contradictory or inconsistent with the entailments of the utterance. Since conversational implicatures do not follow logically from the utterances by which they are licensed, the truth of what is said and the truth of what is conversationally implicated by *S* are independent. So, the latter may be denied without denying the former -- and without licensing a sense of inconsistency or contradiction. That is, if p_i is the semantic representation of an utterance u_i , and u_i licenses the meaning p_j , then p_j is cancelable just in case the conjunction ' p_i and $\neg p_j$ ', is not contradictory or inconsistent.

Following Gazdar [Gazdar 79a:107], I will define semantic entailment and semantic consistency as follows:

A sentence p_i is **semantically entailed** by a set of sentences T just in case p_i is true in every possible world in which all members of T are true.

A sentence p_i is **semantically consistent** with a (consistent) set of sentences T just in case p_i is true in some possible world in which all the members of T are true.

So, ENTAILED(p_i, C_h) will denote that p_i is entailed by the sentences realized by the utterances in (a consistent) C_h . And CONSISTENT(p_i, C_h) will be true iff p_i is consistent with the sentences realized by the utterances comprising C_h . That is, a representation p_i of an utterance is consistent with a context (or set of utterances) C_h just in case the sentence realized by p_i is consistent with the set of sentences realized by the members of C_h . Then the following will be true:

$$\text{ENTAILED}(p_i, C_h) \Rightarrow \text{CONSISTENT}(p_i, C_h)$$

That is, if p_i is true in every World in which the sentences realized by the utterances of C_h are true, it is true in some world. Also:

$$\text{ENTAILED}(p_i, C_h) \Rightarrow \neg \text{CONSISTENT}(\neg p_i, C_h)$$

If p_i is true in every world in which the sentences realized by the utterances of C_h are true, then there is no world in which these sentences and the negation of p_i are true.

Then, as a first attempt at defining cancelability, we might say that (the representation of) a meaning p_j licensed by an utterance u_i is cancelable iff both p_j and its negation are consistent with u_i .

$$\text{CANCELABLE}(u_i, p_j) \leftrightarrow (\text{CAUSE}(\text{SAY}(S, H, u_i, C_h), \text{BEL}(H, p_j)) \wedge \text{CONSISTENT}(p_j, C_i) \wedge \text{CONSISTENT}(\neg p_j, C_i))$$

where C_i is the concatenation of C_h and u_i .

Clearly, if p_i is an entailment of some utterance, it is not cancelable by this definition. Say some utterance u_i realizing p_i entails p_i . Then $\text{ENTAILED}(p_i, C_i)$. So, $\text{CONSISTENT}(p_i, C_i)$ -- but $\neg\text{CONSISTENT}(\neg p_i, C_i)$ -- by the axioms noted above. So, for any p_i entailed by u_i , $\neg\text{CANCELABLE}(u_i, p_i)$. For example, one cannot cancel the self-entailment of (14a), as evidenced by the infelicity of (14b).

(14)

- a. George likes apples.
- b. #George likes apples, although George does not like apples.³⁷

It is less clear whether (15b) should be counted as an entailment of (15a), since the proper semantic representation of this utterance has been the subject of some controversy [Russell 04, Strawson 50]. If not, then by a semantic definition of presupposition,³⁸ (15a) may be said to presuppose (15b).

(15)

- a. The King of France is bald.
- b. There is a King of France.
- c. #The King of France is bald but there is no King of France.

In either case, the attempted cancelation in (15c) is infelicitous.

If presupposition is defined semantically, then our current definition of cancelation will account for the infelicity of (15c). However, other sorts of meaning clearly cannot be defined in terms of entailment, such as pragmatic presupposition³⁹ and conventional implicature. For example, (16a), the conventional implicature of (8a) -- which is not entailed by (8a) -- cannot be felicitously canceled, as is attempted in (16b).

(16)

- a. His being brave follows from his being an Englishman.
- b. #He is an Englishman; he is, therefore, brave, but I don't mean to suggest any connection between the two.

³⁷'#' is used throughout to denote pragmatic infelicity.

³⁸ p_i semantically presupposes p_j iff whenever p_i is true, p_j is true, and whenever $\neg p_i$ is true, p_j is true.

³⁹'My neighbor has hurt herself' is said to pragmatically presuppose that the neighbor is female.

Of course, (14b), (15c), and (16b) are all possible as (self) REPAIRS,⁴⁰ where this contradiction is precisely what is intended.

So, while entailments and at least some presuppositions will not be cancelable by the above definition of CANCELABLE, conventional implicatures will be. Not surprisingly, if we define cancelability only in terms of semantic representations -- since we are defining consistency in terms of truth conditions -- meanings not captured semantically will not affect assessments of cancelability. I will return to this problem below.

Unlike other diagnostics discussed below, cancelability is not sensitive to surface order. A canceling clause may follow a putative conversational implicature, as in 13, or precede it, as in 17.

(17) As it turned out there was no connection, but George ate chicken soup and got sick.

Grice also notes that cancelation may be explicit, as in 13 and 17, or contextual. Suppose you have asked me for a five dollar loan, and I hold up a ten dollar bill while uttering '*I have five dollars*'; I will be unlikely to intend to convey '*I do not have more than five dollars*'. In Grice's terms, the putative implicature '*I do not have five dollars*' will be contextually canceled, even without an explicit canceling utterance.

Note, however, that Grice's account of cancelation suggests that conversational implicature is to be defined more in terms of the form of *S*'s utterance than of his/her intentions. That is, Grice seems to be saying that we can say that *S* has canceled a conversational implicature p_j if *S* has (felicitously) asserted $p_i \wedge \neg p_j$ where the assertion of p_i alone *might have* implicated p_j . However, my Condition 1 demands that *S* intend to convey p_j by saying p_i as a condition on conversational implicature. So, if Grice's identification of cancelation with conversational implicature holds, we are forced to say that, in cases like 17, *S* has asserted p_i intending to implicate p_j but also intending to cancel that implicature by the utterance of $\neg p_j$. And if *S* asserts p_i implicating some p_j which is contextually canceled, one must say that *S* intends to implicate p_j while believing that it will be canceled by the context of utterance. Neither of these explanations seems convincing.

Given these considerations, it seems useful to separate the cancelability test for conversational implicature from the use of cancelation in discourse. The identification of cancelations of conversational implicature in discourse must depend upon speaker intention. Simply because *S* realizes the conjunction $p_i \wedge \neg p_j$ in an utterance, such that, if *S* had chosen

⁴⁰Repairs include word recovery, self-editings, error-replacement, and so on.

instead to realize p_i alone in the same context, s/he might have licensed the conversational implicature p_j , one would not say that S has both implicated and canceled p_j . Unless S intends for p_j to be conveyed then p_j has not been implicated. It may be that 17, for example, represents a cancellation of conversational implicature, but only if S intends (for some particular effect) to license an implicature and cancel it in the same utterance.

In fact, the assertion of $\neg p_j$ in cases like 17 seems more naturally interpreted as the BLOCKING of inferences S anticipates H might draw from the assertion of p_i , if perhaps H were to 'work out' some implicature which S has not intended. Since phenomena such as ASYMMETRIC *and* for example, are familiar to conversational participants, S may sometimes anticipate that H will infer a meaning from S 's use of constructions or lexical items that can be used to license implicatures. So, S may utter $\neg p_i$ to block this undesired inference. In 18, for example, S blocks the possible inference that he performed no more than two experiments.

(18) I performed two experiments. Well, I performed FOUR, but I'm presenting TWO.

In a sense 18 has the flavor of a repair rather than a cancellation, although there is no sense that S has said something contradictory or inconsistent by the second utterance. Instances of 'cancellation' noted in the literature, then, may more properly be explained as the blocking of inferences H might otherwise be tempted to 'work out' as conversational implicatures.

However, cancelability can still serve as a test for conversational implicature as follows: If we believe that the utterance of some p_i can license p_j and wish to test whether that p_j is licensed via conversational implicature, then the question "Can p_i be felicitously uttered in a context in which p_j is denied or clearly false?" will sufficiently represent what we need from the cancelability test. One can still ask whether the assertion of p_i conversationally implicates p_j by asking (among other things) whether the assertion of $p_i \wedge \neg p_j$ would be felicitous -- without specifying that such an artificial construction itself represents an implicature and its cancellation.

Although cancellation has proven the most useful of Grice's tests, Grice and others have noted its weaknesses: First, it is easy to confuse with SUSPENSION, the 'calling into question' of an asserted proposition [Horn 72]. Suspension does **not** distinguish between conventional and conversational implicature or between conversational implicature and presupposition; even entailment may be suspended by some suitable hedge. Compare the felicitous suspension of that which is presupposed by *only* in (19a) with the infelicitous attempted cancellation in (19b).⁴¹

⁴¹'Only x P(x)' is generally analyzed as presupposing that P(x) and conventionally implicating that there are other y such that it might have been supposed that P(y) while, in fact, $\neg P(y)$.

- (19)
- a. Only Muriel voted for Humphrey, if even she did.
 - b. #Only Muriel voted for Humphrey and even she didn't.

While ease of confusion with suspension is no reason to throw out the cancelation test, it does make it less attractive. Especially since the definition of cancelation given above -- and commonly accepted -- is lamentably imprecise, this potential for confusion represents a significant, though perhaps tractable, problem.⁴² Much less clear, however, is what can constitute a CONTEXTUAL CANCELATION. Grice appears to mean that conversational implicatures that would be licensed in some contexts will not be in others; that is, the context will 'cancel' the implicature. But since conversational implicature is by definition context-dependent, it seems odd to propose an additional mechanism to explain this definitional criterion.

Second, Grice himself notes [Grice 78] that cancelability does not constitute a sufficient test for conversational implicature. When people speak "loosely", the SENSE of an utterance (what is asserted) may be contextually canceled. For example, he claims that, when looking at a tie in different lights, *S* may say (20a) without intending that *H* believe the tie really is changing colors.

- (20)
- a. It is a light green now.
 - b. It is a light green now, although of course it hasn't really changed color.

So, part of the conventional force of the utterance is contextually canceled. [Walker 75:159] provides further arguments in this vein. Although one might salvage the cancelability condition by using only explicit cancelability to distinguish conventional from non-conventional, it seems that (20b) is also possible. A more telling counter is to note the ambiguity of the copula here; if interpreted as 'appears' then there is no need to postulate contextual cancelation of its other senses.

⁴²It does seem that explicit cancelations exhibit some interesting regularities. For example, Gregory Ward (personal communication) has noted that cancelations seem to require some conventional signaling. For example, (ia) is fine but (ib) is odd.

- (i)
- (a) Some people left early and *in fact* everyone did.
 - (b) #Some people left early and everyone did.

However, phrases such as *in fact*, *actually* and *indeed* are also associated with the suspension of presupposition, so these signals are not unambiguous indications of implicature cancelation [Allwood 72].

The argument that contextual disambiguation may be seen as cancelation -- and thus that cancelability does not provide a useful test for conversational implicature -- is made explicit by Sadock [Sadock 78]. He notes that cancelability does not distinguish cases of ambiguity from cases of univocality plus possible conversational implicature. One sense of an ambiguous utterance such as (21a) (which conveys either (21c) or (21d)), can always be denied, 'giving the impression of' a cancelation of one interpretation. So, the denial in (21b) of one interpretation of (21a) -- namely, (21c) -- appears similar to cancelation.

(21)

- a. Everyone speaks one language.
- b. Everyone speaks one language although no one language is spoken by everyone.
- c. One and the same language is spoken by everyone.
- d. Everyone speaks at least one language.

Again, though, this objection is much like the contention that cancelation is difficult to distinguish from suspension; if disambiguation is only difficult to distinguish from cancelation, this fact may make cancelability harder to determine, but it does not invalidate cancelation as a diagnostic for conversational implicature. If disambiguation is instead seen as a **form** of cancelation, then the constraint that conversational implicature be a part of *S*'s intentions (Condition 1) should serve to exclude unintended senses from the realm of conversational implicature. That is, if *S* does not **intend** to convey that (21c), then that cannot be deemed a conversational implicature. If *S* does intend to convey (21c) by uttering (21a) -- i.e., if *S* believes that (21c) -- then (21b) should be seen as contradictory.

Third, a more serious attack on the usefulness of the cancelation test for conversational implicature is Sadock's [Sadock 78:293] claim that cancelability is not even a necessary condition for conversational implicature. He contends that at least one type of conversational implicature, that licensed by the utterance of *almost p_i*, as in 22,

(22) Gertrude almost swam the English Channel.

is "just about uncancelable" -- but nonetheless constitutes a conversational implicature. Contra [Karttunen 75], Sadock proposes that the utterance of *almost p_i* can conversationally implicate 'not *p_i*' on the following bases [Sadock 78, Sadock 81]: First he claims that the additional meaning here is easily 'worked out': if *S* asserts *almost p_i*, s/he may anticipate that *H* can 'work out' that 'for all a (cooperative) *S* knows, not *p_i*', since, if *S* knows that *p_i*, it is *misleading* to say *almost p_i*. Sadock must of course maintain that this would be at most *misleading*, not *false*; for, if saying *almost p_i* when *p_i* is believed true is seen as falsehood, then the understanding 'not *p_i*' would be part of the conventional force of 'almost *p_i*'. Second, Sadock cites the high degree

of nondetachability of *almost p_i*.⁴³ That is, the construction passes the other classic test for conversational implicature. But, Sadock finds that it is almost impossible to cancel the 'not *p_i*' implicature, as in 23

(23) ??Gertrude not only almost swam the English Channel, in fact she swam it.⁴⁴

So, if utterances like 23 are indeed bizarre, but if *almost p_i* does pass the other test for conversational implicature, then Sadock would maintain that cancelability is not a necessary condition for conversational implicature.

Several responses can be made to Sadock's contention: One might claim that *almost* does license conversational implicatures but that these in fact are cancelable. This contention is supported by the fact that many speakers do find 23 and similar cancelations of *almost* acceptable.⁴⁵ For those who do not think 23 felicitous, the next question must be: is the additional meaning licensed by *almost p_i* part of the conventional force of the utterance? If so, then *almost* will not present a counter-example to Grice's cancelability diagnostic for conversational implicature.

The claim that 'almost' licenses conventional implicatures might be supported in several ways: First, if *almost* fails what has been the most reliable diagnostic for conversational implicature, it is difficult to say why the meaning it conveys should be termed conversational implicature. The nondetachability test, itself even harder to apply than cancelability, does not appear to produce such clear results as Sadock believes. For example, if *all but p_i* is included among the various 'ways of saying' *almost p_i*, as seems intuitively justifiable, then, while its utterance does clearly license *not p_i*, this understanding seems to be a conventional implicature of the utterance, apparently licensed in some way by the presence of *but* in the utterance.⁴⁶ That is, *all but p_i and in fact p_i* appears even more clearly infelicitous than *almost p_i and in fact p_i*. Consider (24a) and (24b).

So, at least one of the 'other ways of saying' *almost p_i*, appears to license conventional

⁴³See below, Section 2.4.2.2. Briefly, any other way of saying 'almost *p_i*', such as *just about p_i* or *come close to p_i*, may license the same understanding.

⁴⁴Sadock uses '??' to indicate oddness but not complete infelicity.

⁴⁵However, it is not clear that 23 would actually represent a simple cancelation of 22 even if it is felicitous. Placing the potential implicator 22 within the scope of either 'not' or 'only' may well have inhibiting effects on the licensing of implicature. However, speakers do find that a fairer test, ii, is even less acceptable.

(ii) Gertrude almost swam the English Channel, and, in fact, she swam it.

⁴⁶The compositionality of 'all but' is an interesting question for further analysis.

- (24)
- a. #Mark is all but engaged and in fact he is.
 - b. (?)Mark is almost engaged and in fact he is.

meanings, not conversational implicature; hence, ‘almost’ might better be seen as licensing conventional implicatures itself. Finally, Sadock himself notes that utterances such as 22 do not depend for their interpretation on context. Since conversational implicatures are just those meanings that arise from the interpretation of an utterance in context, this observation suggests that *almost* does not license conversational implicature.

In summary, Sadock’s argument against cancelability as a necessary condition for conversational implicature rests on shaky ground indeed. Better cases can be made both for *almost* licensing conversational implicatures which indeed are cancelable and for *almost* licensing conventional implicatures than for *almost* licensing noncancelable conversational implicatures. In neither case will *almost* represent a counter-example to cancelability as a necessary condition for conversational implicature. Thus, despite Sadock’s claim, cancelability does appear to remain a necessary condition for conversational implicature. So, we can augment our Conditions 1 through 3 above with:

Condition 4. p_j is cancelable.

That is, if S saying u_i (which realizes p_i) conversationally implicates p_j , then p_j can be felicitously denied without denying p_i .

Initially, I noted certain problems involved in a formal representation of cancelability. Since, at the moment, we have no way to represent conventional implicatures or pragmatic presuppositions -- or any meaning beyond truth-functional meaning -- it is impossible to exclude such meanings from the set of cancelable meanings without appealing to less ‘objective’ concepts. For now, we must simply stipulate that $FELICITOUS(p_i, C_h)$ holds whenever the semantic representation p_i of an utterance is felicitous in a context C_h . So, $FELICITOUS$ will represent a weaker condition on p_i and C_h than $CONSISTENT$. Then we can define cancelability as follows:

Cancelability:

$$\begin{aligned} CANCELABLE(u_i, p_j) \leftrightarrow & (CAUSE(SAY(S, H, u_i, C_h), BEL(H, p_j)) \wedge \\ & REALIZE(u_i, \neg p_j) \wedge FELICITOUS(u_i, \{C_h \cup u_i\})) \end{aligned}$$

That is, a meaning p_j which is licensed via the utterance of u_i is cancelable iff S may felicitously realize $\neg p_j$ in an utterance u_j in a context in which s/he has uttered u_i .

2.4.2.2. Nondetachability

According to [Grice 75], conversational implicatures other than those arising via the Maxim of Manner cannot be DETACHED from what is asserted (Grice's condition G5): if one assertion of p_i implicates p_j , then 'any other way' of asserting p_i will also implicate p_j . So, when speaker intentions and contexts are held constant, the substitution of utterances with the same conventional force preserves conversational implicature. For example, if S wishes to implicate (25c) in some context C_h , then s/he may choose among at least (25a) and (25b) to accomplish this goal.

- (25)
- a. Jack bought half a dozen bagels.
 - b. Jack bought six bagels.
 - c. \neg BEL(S , *Jack bought more than six bagels*)

The same does not hold for conventional implicatures, which are not nondetachable.⁴⁷ For example, the utterance of (26a) cannot license the additional meaning communicated by the utterance

- (26)
- a. He's a New Yorker and I like him.
 - b. He's a New York but I like him.

of (26b), even though these utterances have the same conventional force. Although nondetachability thus distinguishes between conventional and conversational implicature, it does not distinguish conversational implicature from other aspects of utterance interpretation, such as presupposition and entailment [Grice 78, Sadock 78]. So any way of asserting (27a) -- such as, (27b) -- will presuppose (27c).

- (27)
- a. He has stopped beating his wife.
 - b. He no longer beats his wife.
 - c. He used to beat his wife.

And no way of asserting (28a) -- say, (28b) -- fails to license (28c).

- (28)
- a. Bill and Harry left.
 - b. Harry and Bill left.
 - c. Harry left.

So, nondetachability, like cancelability, may at best represent a necessary condition on conversational implicature. And, of course, for conversational implicatures arising via the

⁴⁷Although it seems simpler to say then that they are DETACHABLE, this simplification has not generally been made in the literature.

Maxim of Manner it is not even this.⁴⁸

Still other limitations of this diagnostic arise from the difficulty of applying it. As [Walker 75, Sadock 78] point out, what it means for some u_i, u_j to represent ‘different ways’ of asserting p_j is not at all clear. For example, (29a) and (29b) might well be seen as ‘saying the same thing’.

- (29)
- a. I have one leg.
 - b. I have a single leg.
 - c. \neg BEL(S, I have more than one leg)

However, if, in a given context, the utterance of either (29a) or (29b) licenses (29c), it seems intuitively clearer that (29a) licenses (29c) via conversational implicature than that (29b) licenses (29c) via conversational implicature. Furthermore, (29b) is not cancelable. So, (29a) and (29b) must differ in their conventional force. But at this point, the nondetachability test becomes somewhat circular, unless an independent notion of conventional force can be established.

It seems that what Grice intends by conventional force is that part of utterance meaning which can be captured in a (truth-functional) semantic theory, whatever that theory might be. Then nondetachability can at least be defined with respect to a particular semantic theory for the purpose of applying the diagnostic to particular utterances. So, u_i and u_j might qualify as ‘different ways of saying they same thing’ just in case their surface forms differ but their ‘semantic representations’ in this theory are identical (modulo reference resolution). That is, u_i and u_j both realize (the proposition represented by) p_i . So we might define nondetachability as follows:

Nondetachability:

$$\begin{aligned} \text{NONDETACHABLE}(u_i, p_j) \leftrightarrow \\ (\text{REALIZE}(u_i, p_i) \wedge \text{CAUSE}(\text{SAY}(S, H, u_i, C_h), \text{BEL}(H, p_j)) \wedge \forall u_j \\ (\text{REALIZE}(u_j, p_i) \Rightarrow \text{CAUSE}(\text{SAY}(S, H, u_j, C_h), \text{BEL}(H, p_j)))) \end{aligned}$$

Although this solution seems to make the diagnostic feasible, it does not really provide the independent definition of conventional force required. The burden of determining what is conventional and what is not is simply pushed back upon semantic theory. If *single* and *one* have been considered semantically equivalent, then the nondetachability test will predict that (29b) licenses (29c) via conversational implicature; if not, then that meaning will be assigned the status of conventional implicature.

⁴⁸As, for example, the conversational implicatures arising via the Maxim of Manner from the utterances of (5a) and (6a) are not nondetachable.

In addition to this fundamental problem of defining equivalence in conventional force, as Grice himself noted, nondetachability is not necessary for all conversational implicature, since conversational implicatures arising via the Maxim of Manner are detachable. In implicatures such as that which can be licensed via asymmetric *and*, for example, it is a particular way of saying something that licenses the implicature. So, (30a) and (30b) will have the same

- (30)
- a. Mable got sick and had chicken soup.
 - b. Mable had chicken soup and got sick.

semantic representation but their utterance may license different implicatures. Of course, since the Maxim of Manner is, like the rest, a very general one, it may be tempting to say of any candidate implicature which fails the nondetachability test that it is somehow derived from this maxim.

[Walker 75, Sadock 78] have also claimed that nondetachability is not a necessary condition even for implicatures arising **other than** from the Maxim of Manner. They contend 1) that there may be no way other than a ‘longer-winded’ way to say what is said, which would itself violate the Maxim of Manner; or, 2) that there may be no other way at all to say what is said, so the nondetachability test cannot be applied. Although these contentions do support claims that nondetachability may be of limited utility as a diagnostic, they do not support the claim that it is not a necessary condition: If there is no other way to say what is said, then, trivially, ‘any other way’ will license the same implicature. So, although nondetachability may be of less practical use in identifying conversational implicature than cancelability, it is nonetheless a necessary condition on conversational implicature.

For $\text{NONDETACHABLE}(u_i, p_j)$ will be vacuously true if we compute conversational implicatures from the semantic representations of utterances. That is, we have assumed that, if u_i and u_j have the same semantic representation p_i , they will be deemed to have equivalent conventional force. But if we then calculate implicatures p_j from p_i , u_i and u_j must license the same p_j . However, since we not only want to be able to identify conversational implicatures but also to describe properties of these implicatures, then condition 5 should be added to our definition of conversational implicature.

Condition 5. p_j is nondetachable except when it arises via the Maxim of Manner.

That is, if S saying u_i (which realizes p_i) licenses p_j , then every u_j that also realize p_i will also license p_j -- unless u_i licenses p_j via the Maxim of Manner. Either p_j is nondetachable -- or those maxims involved in LICENSE do not include the Maxim of Manner.

2.4.2.3. Redundancy of Conjunction/Reinforceability

Horn [Horn 72:63] suggests another test for conversational implicature, which is intended to distinguish conversational implicature from presupposition and entailment. He notes that, in cases in which p_i presupposes or entails p_j , it is possible to conjoin p_j to the left of p_i but not to the right. So, (31a), where (31c) is presupposed by (31d), is fine, while (31b) appears odd.

- (31)
- a. John is a man and he is a bachelor.
 - b. #John is a bachelor and he is a man.
 - c. John is a man.
 - d. John is a bachelor.

However, explicit assertions of conversational implicatures can be felicitously conjoined either to the right or left of the assertion which might otherwise license them, as in (32a) and (32b).

- (32)
- a. Some people left early but not everybody did.
 - b. Not everybody left early but some people did.

[Sadock 78]'s notion that conversational implicatures are REINFORCEABLE builds upon Horn's diagnostic. Sadock notes that REINFORCEABILITY also distinguishes conventional from conversational implicature.⁴⁹ He proposes that, since conversational implicatures are not part of the conventional force of an utterance, one can make them explicit (by subsequent conjunction as above) without being guilty of REDUNDANCY, as in his example reproduced as (33a). But understandings that are part of that force are

- (33)
- a. Maggie ate some, but not all, of the cheddar.
 - b. #It's odd that dogs eat cheese and they do.
 - c. Dogs eat cheese and it's odd that they do.

not reinforceable without redundancy, as in (33b) where a conventional implicature is reinforced. Of course, (33c) is acceptable. Unlike cancellation, reinforcement occurs frequently in natural discourse, as in 34.

⁴⁹He finds that cases such as *almost* can be distinguished better by reinforceability than by cancellation. That is, where (he believes) the cancellation test fails for *almost* above (23), reinforcement succeeds, since iii is fine.

(iii) Gertrude almost, but didn't quite, swim the English Channel.

However, if *almost* is not seen as licensing conversational implicatures – but rather conventional ones, this test becomes suspect. In such a case, it might be that iii carries little sense of 'redundancy' simply because it helps to disambiguate between two possible interpretations of 22: Gertrude swam part way across the channel and Gertrude thought about swimming the channel but didn't get near the water; iii seems to favor the former reading. But note that iv seems as acceptable as iii.

(iv) Gertrude almost, but didn't quite, jump off the Golden Gate Bridge.

(Thanks to Gregory Ward for this final example.)

- (34) "Some (*but not all*) of the points made in sections 1-3 of this paper are implicit or explicit in Montague [1974], especially in PTQ, the paper 'The Proper Treatment of Quantification in Ordinary English'. (Some of the suggestions in 1-3 are also similar to suggestions other papers: Fenstad [1978], Peacocke [1979], e.g.)."
 (J. Barwise and R. Cooper, "Generalized quantifiers and natural language", Spring 1980, mss, p.1a. My italics.)

Note that, like cancelation, reinforcement also fails to distinguish conversational implicature from ambiguity, as in the following:

- (35) Everyone speaks one language and it is the same language.

However, the same arguments apply then as above (Section 2.4.2.1).

Redundancy of conjunction/reinforceability (which I will term REINFORCEABILITY for convenience) thus appears to be an additional necessary condition on conversational implicature. Although a notion of 'redundancy' for this diagnostic may be even harder to tease out than one of 'contradiction' for cancelability, since our goal is not only to identify conversational implicatures but to describe properties of these implicatures, the fullest possible description of their features is desirable. So, we can add 6 to our conditions on conversational implicature:

Condition 6. p_j is reinforceable.

That is, if S saying u_i (which realizes p_i) licenses p_j , then an utterance of the u_i immediately follow by an utterance of u_j (where u_j realizes p_j) is not redundant.

A definition of reinforceability is made difficult for just the same reasons that cancelation and nondetachability are difficult to formalize: we are working with an inadequate representation of the conventional force of an utterance. So, again we are forced to define a less than objective notion -- this time REDUNDANT(u_i, C_h), denoting that u_i is redundant in the context C_h ⁵⁰ -- to permit the following definition:

Reinforceability:

$$\text{REINFORCEABLE}(u_i, p_j) \leftrightarrow (\text{CAUSE}(\text{SAY}(S, H, u_i, C_h), \text{BEL}(H, p_j)) \wedge \text{REALIZE}(u_j, p_j) \wedge \neg \text{REDUNDANT}(u_j, \{...u_i\}))$$

⁵⁰Although a satisfactory axiomatization of REDUNDANT must at least await an account of conventional implicature, it seems likely that we might propose a number of conditions under which u_i might be termed redundant, e.g.:

$$\text{REDUNDANT}(u_i, \{...u_i\})$$

That is, u_i is redundant in a context in which it has already been uttered -- modulo some temporal processing restrictions perhaps. Also, we might say that u_j will be redundant in a context in which a sentence p_i entailing a p_j which u_j realizes has already been realized:

$$(\text{ENTAILED}(p_i, \{...u_i\}) \wedge \text{REALIZE}(u_j, p_j)) \Rightarrow \text{REDUNDANT}(u_j, \{...u_i\})$$

2.4.3. Defining Conversational Implicature

I have proposed a revised set of necessary conditions for conversational implicature which not only help to identify particular conversational implicatures but more general classes of conversational implicature. In this section I will claim that these conditions are, collectively, sufficient to define conversational implicature as follows:

Conversational Implicature:

CONVERSATIONAL_IMPLIC(S, H, p_j, u_i, C_h) \leftrightarrow

1. INTEND($S, CAUSE(SAY(S, H, u_i, C_h), BEL(H, p_j))$) \wedge
2. BEL($S, BMB(H, S, IS_COOP(S, C_h, \{QUALITY, QUANTITY, RELATION, MANNER\}))$) \wedge
3. BMB($S, H, LICENSE(S, H, p_j, u_i, C_h, MAXIMS)$) \wedge
4. CANCELABLE(u_i, p_j) \wedge
5. (NONDETACHABLE(u_i, p_j) \vee (MANNER \in MAXIMS)) \wedge
6. REINFORCEABLE(u_i, C_h)

That is, we can define what it means for S conversationally to implicate some p_j to H by saying u_i in a context C_h as follows: S intends to convey p_j to H via u_i ; S believes that it is mutually believed between S and H that S is being cooperative -- i.e., obeying the maxims of cooperative conversation; S believes that it is mutually believed between S and H that, given S 's u_i in a context C_h , and given S 's cooperativity (i.e., obedience to the maxims represented in MAXIMS), p_j 'follows'; p_j is cancelable, nondetachable except when MAXIMS includes the Maxim of Manner; and p_j is reinforceable.

The principal arguments against the proposal that conditions 1 through 6 represent necessary and sufficient conditions for conversational implicature would be Sadock's claims that not all conversational implicature is cancelable; and that even conversational implicatures not derived via the Maxim of Manner may not be nondetachable. Above I have argued against both of these claims, maintaining that Sadock's analysis of *almost* is not convincing enough to warrant abandoning cancelability as a diagnostic for conversational implicature and that Sadock's objections to nondetachability, while well-founded, dispute its practicality rather than its necessity. However, one might also wonder whether, since it has been claimed that diagnostics 4 and 6 fail to distinguish conversational implicature from ambiguous meanings which may be canceled or reinforced, conditions 1 through 6 collectively make this distinction. In fact, as I have argued above, condition 1 at least excludes ambiguous senses from conversational implicature, and a case might also be made that conditions 3 and 5 do so as well. More generally, of those aspects of utterance interpretation that we want to distinguish from conversational implicature via this definition, entailment is ruled out by conditions 4 and 6;

presupposition by 4; conventional implicature by 4 through 6; and ambiguous meanings by 1 and (possibly) 3 and 5. While other aspects of utterance interpretation might possibly be defined,⁵¹ while other taxonomies of utterance interpretation might restructure the problem, and while it might be possible to discover other necessary conditions for conversational implicature, for the moment, the conditions stated above represent both the established necessary conditions for conversational implicature and, collectively, sufficient conditions to distinguish conversational implicature from all other aspects of utterance interpretation identified within the Gricean framework.

2.4.4. Inferring Conversational Implicature

Note that, although I have defined conversational implicature from *S*'s point of view in this definition, this definition also makes it possible to identify conditions under which *H* may infer that some conversational implicature has been implicated. Simply by placing conditions 1 through 3 within *H*'s belief space, we can say what it means for *H* to be entitled to infer that *S* has implicated that p_j .

Inferred Conversational Implicature:

- $$\text{BEL}(H, \text{CONVERSATIONAL_IMPLIC}(S, H, p_j, u_i, C_h)) \leftrightarrow$$
1. $\text{BEL}(H, \text{INTEND}(S, \text{CAUSE}(\text{SAY}(S, H, u_i, C_h), \text{BEL}(H, p_j)))) \wedge$
 2. $\text{BEL}(H, \text{BMB}(S, H, \text{IS_COOP}(S, C_h, \{\text{QUALITY}, \text{QUANTITY}, \text{RELATION}, \text{MANNER}\}))) \wedge$
 3. $\text{BMB}(H, S, \text{LICENSE}(S, H, p_j, u_i, C_h, \text{MAXIMS})) \wedge$
 4. $\text{CANCELABLE}(u_i, p_j) \wedge$
 5. $(\text{NONDETACHABLE}(u_i, p_j) \vee (\text{MANNER} \in \text{MAXIMS})) \wedge$
 6. $\text{REINFORCEABLE}(u_i, p_j)$

The notion of what it means for *H* to believe that *S* has conversationally implicated p_j is important in that it allows us to specify the circumstances under which *S* may believe that *H* has inferred an implicature -- whether or not *S* intended one; hence, defining conversational implicature in terms of *S* will not prevent us from reasoning about *H*'s inference of conversational implicature.

Given a definition of conversational implicature and a specification of conditions under which *H* will believe that *S* has licensed such implicatures, we would like to use it both to

⁵¹In particular, one might attempt to define those non-conventional, non-conversational and perhaps 'non-linguistic' implicatures noted in Figure 2-0.

identify and to guide the generation of conversational implicature in computer-human interaction. We would also like to be able to define properties of this class of utterance meanings, so that, ultimately, we will be able to integrate them into a broader notion of utterance meaning.

The usefulness of the formalism provided here is limited by its inclusion of notions such as **FELICITOUS**, which must for now be stipulated. However, better ways of defining cancelability and reinforceability appear possible if some semantics can be given for conventional implicature. Nonetheless, this definition of conversational implicature does provide a basis on which certain classes of conversational implicature termed **GENERALIZED CONVERSATIONAL IMPLICATURES** may be defined. In the next chapter I will discuss previous attempts to define classes of generalized implicature based primarily on the assumption of speaker observance of the Maxim of Quantity. These attempts form the basis for the definition of a class of generalized quantity implicature which I will term **SCALAR IMPLICATURE**.

CHAPTER III

Scalar Implicature

A: What month has 28 days in it?

B: February?

A: Every month has 28 days in it.

Katie Finin

The difficulty of formulating a satisfactory definition of conversational implicature in the abstract has led some to attempt an alternative approach: the definition of classes of conversational implicature which, as classes, can be argued to meet the hard-to-formalize criteria for conversational implicature in general. Such GENERALIZED CONVERSATIONAL IMPLICATURES have been characterized in terms of some specific lexical item or some linguistic construction, like the asymmetric use of conjunction discussed in Chapter 2.

Among such classes, conversational implicatures which involve the Maxim of Quantity have often served as a case in point. Very similar classes of generalized quantity implicature have been described in [Harnish 79], [Horn 72], and [Gazdar 79b, Gazdar 79a, Gazdar 80]. While Harnish and Horn describe the phenomenon from a more intuitive point of view, Gazdar attempts to formalize these intuitions for computational purposes. In this chapter, I will first explore the basis for the current definition of generalized conversational implicature. I will then describe theories of quantity implicature presented by Harnish, Horn, and Gazdar and discuss the strengths and weakness of each; together, these works form the intuitive basis for the class of SCALAR IMPLICATURES which I will subsequently introduce.

3.1. Generalized Conversational Implicature

Within the class of conversational implicatures, Grice differentiates between what is **GENERALLY** conversationally implicated and what is **PARTICULARLY** conversationally implicated. Particularized implicatures are exemplified by Grice's classic:

- (36) A: I am out of petrol.(= [Grice 75]'s 1)
 B: There is a garage round the corner.

In this example, Grice claims B is implicating that, as far as s/he knows, (37a)

- (37)
 a. The garage is able to provide fuel.
 b. There is a garage round the corner but unfortunately it's closed.

holds. For B could not be observing the Maxim of Relation unless his/her mention of the garage were relevant to A's statement. If the garage were closed or out of fuel -- and if B knew this -- then a cooperative B would not mention its location unless s/he blocked the inference (37a), as by asserting (37b). So, either B knows (37a) or s/he does not know whether (37a). Given that B and A can both work this out, B can implicate (37a) by his/her response in 36.

Implicatures such as these Grice terms **PARTICULARIZED** because the process by which they seem to be worked out appears wholly dependent upon the particulars of a given situation. There seem to him to be no principles which one can abstract from the form of 36 or from the context in which it is uttered that might link (37a) to implicatures licensed by other utterances. **GENERALIZED** conversational implicatures however, can be analyzed in broader terms. One might, however, counter that a frame or script-based approach might indeed propose similarities between this exchange and one like 38.

- (38) A: I'm hungry.
 B: There's a restaurant next door.

Indeed, I will propose that the traditional distinction between generalized and particularized implicature is a false one, an artifact of the inventiveness of analysts -- or lack thereof.

3.1.1. Classes of Generalized Implicatures

To date, a number of classes of generalized implicature have been identified -- primarily on the basis of a speaker's use of certain lexical items or certain syntactic constructions. These classes include implicatures characterized by mention of 'ranked verbs' such as *like* and *love*, by use of productive causatives (*cause to X*), and by mention of the natural-language equivalents of the logical quantifiers and connectives. However, it has been difficult to determine how best to define generalized implicature as a class.

3.1.1.1. Asymmetric 'And'

It has been claimed [Lakoff 71, Schmerling 75] that the concept of ASYMMETRIC *and* described in Chapter 2 identifies a class of generalized conversational implicature. Recall that implicatures such as those licensed via (39a) and (39b) are closely tied to

- (39)
- a. Maude went to Las Vegas and got rich.
 - b. Maude got rich and went to Las Vegas.

S's use of conjunction. Note the similarity of these implicatures to each other and to the implicature which might be licensed by the utterance of 40. In each case, *S* may license a temporal/causal inference via the Maxim of Manner.⁵²

- (40) Harold got bored and went to church.

Like all conversational implicature, these are contextually determined: The utterance of (41a), for example, might convey a temporal ordering between conjuncts, although clearly it need not always license this meaning.

- (41)
- a. No, Officer, I was sure Ralph hadn't done anything unusual the night of the crime. I checked with his friends and, sure enough, that night *Ralph did his homework and called Nancy*.
 - b. Ralph never lets his social life interfere with school, and he always sticks to his schedule. For instance, he was really afraid once that Murray was going to ask Ralph's girl Nancy to the Prom. Ralph never asked her out more than a week before, so of course he hadn't got her sewn up for that. Now, you or I would've probably called her the minute we got out of basketball practice. But not our Ralph. Sure enough, that night *Ralph did his homework and called Nancy*.

But it seems more likely that such a temporal implicature is intended by the utterer of (41b). So, these generalized implicatures, while more context-independent than particularized implicatures like that licensed in 36, are still context-dependent.

The relative context-independence of generalized conversational implicatures has led to serious confusion over the distinction between this phenomenon and the conventional force of an utterance. Even Grice [Grice 75:56] describes generalized conversational implicatures as cases in which "the use of a certain form of words in an utterance would normally (in the

⁵²Schmerling [Schmerling 75] suggests that the additional meaning conveyed in such cases is better captured by the notion that the first conjunct has some sort of priority over the second. This priority may be temporal or causal; or, the first conjunct may be a necessary condition for the second, as in v.

(v) I left the door open and the cat got in.(= [Schmerling 75]'s 72)

ABSENCE of special circumstances) carry such-and-such an implicature or type of implicature." From this description Harnish [Harnish 79:353] is led to claim that, when certain lexical items which can license generalized implicatures are uttered, the implicature will hold unless it is explicitly or contextually canceled; others effectively adhere to this view.⁵³

Perhaps the claim might be made that the 'default' implicature carried by *and* (or by conjunction) is canceled by the context in (41a). However, under this characterization one would be forced to say that the 'normal' use of *and* is the asymmetric understanding conveyed in (39a), 40, and (41b). In (41a), some unspecified 'special circumstances' contextually cancel this 'normal' temporal/causal understanding. But why asymmetric *and* should be more 'normal' than symmetric *and* is unclear. And how these canceling circumstances may be identified is never discussed by those who assume their existence. Finally, the central role of speaker intention which Grice himself champions elsewhere (See Section 2.4.1.) is impossible to maintain under this account of generalized conversational implicature.

These observations should indicate why I have rejected a representation of conversational implicature in general and scalar implicature in particular in terms of some DEFAULT LOGIC such as Reiter's [Reiter 80]. While such a strategy might be feasible in an engineering sense, there is no principled basis upon which to assign defaults, as the case of asymmetric *and* indicates. In particular, note the difficulty of assigning a 'default interpretation' to mention of the cardinals, as noted in Sections 3.2.1 and 5.1.4.1.

A similar case against current definitions of generalized conversational implicature may be made for those generalized conversational implicatures defined by *S*'s use of conditionals and productive causatives, as well as phenomena such as disjunction, indefinites, quantifiers, verbs of incompletion, and certain related noun or verb pairs which will be discussed in Chapter 5. While 'association with particular lexical items' or 'association with particular constructions' is sufficient to accommodate some of these classes, it will not suffice for all.

3.1.1.2. Conditionals

Gazdar [Gazdar 79a] identifies a class of CLAUSAL QUANTITY IMPLICATURES with implicatures licensed via the utterance of compound or complex sentences which have some constituent s_1 whose affirmation or negation is not entailed or presupposed by the matrix *s*. For example, in (42a), neither s_1 's affirmation nor its negation is entailed or presupposed by *s*. According to Gazdar, if *S* utters a sentence like (42a) while knowing s_1 to be either true or false,

⁵³This tendency is also apparent in the work of Horn [Horn 72], and Gazdar [Gazdar 79a], discussed below.

- (42)
- a. $s[s_1$ [If John is convicted] $_{s_1}$ he will hang] $_s$
 - b. $s[s_1$ [Since John has been convicted] $_{s_1}$ he will hang] $_s$

s/he violates the Maxim of Quantity. A more informative but equally brief sentence might be produced which presupposes either s_1 or its negation, perhaps, (42b). So, H is entitled to infer that ‘for all S knows s_1 ’ and ‘for all S knows $\neg s_1$ ’ are clausal quantity implicatures of (42a). Given that S is believed cooperative, then, s/he may implicate ‘for all S knows {John has been convicted, John has not been convicted}’ via (42a).

Following Geis and Zwicky [Geis 71], Horn [Horn 72] claims that conditionals can convey other sorts of generalized implicatures as well. To offer a condition for some p_i to apply is to implicate that **only** this condition will do; i.e., if S says that p_i is a sufficient condition for p_j , then S may implicate that p_i is a necessary condition too. So, by saying ‘if p_i then p_j ’, S can license the implicature that ‘if $\neg p_i$ then $\neg p_j$ ’.⁵⁴ Thus, for Horn, the speaker of (42a) may implicate 43

(43) If John hangs, he will have been convicted.

as well as Gazdar’s clausal quantity implicatures.⁵⁵

To characterize clausal quantity implicatures as ‘associated with’ the lexical items *if...then* is clearly unsatisfactory. Recall that all conversational implicature must be nondetachable (See Section 2.4.2.2.), i.e., any other waying of lexicalizing conditionality must convey the same implicature. Obviously, conditionality may be lexicalized by items other than *if...then* -- as, ‘say p_i occurs, then p_j ’. Such constructions may also license the belief that S can neither affirm nor deny p_i . So, again, it seems wiser to define classes of generalized conversational implicature in terms of phenomena broader than the mention of particular lexical items -- perhaps some semantic regularity.

⁵⁴In effect, this explains the well-known fallacy of denying the antecedent in terms of conversational implicature. See also Prince’s suggestion, discussed in Section 5.1.3.

⁵⁵This serves as a nice illustration of the fact that a single utterance -- and, even, a single aspect of that utterance -- may convey multiple implicatures. It should also strengthen the case for the primacy of speaker intention to a definition of conversational implicature, since clearly S may say that (42a) without wishing to implicate that 43.

3.1.1.3. Principle of Extra Effort

However, other classes of generalized conversational implicature have been defined in terms of a 'principle of extra effort'; which is difficult to encompass even in the more abstract terms of semantic similarity. As McCawley [McCawley 78] notes, when *S* says (44a) instead of (44b), *s/he*

- (44)
- a. The tulip is pale red.
 - b. The tulip is pink.

expends greater effort than the alternative utterance (here, (44b)) would require. In this case this 'extra effort' manifests itself in an utterance containing more morphemes than the alternative. Here, *S* conveys to *H* that, for some reason, (44b) will not fully convey his/her meaning. In (44a), McCawley believes that *S* conveys that the tulip is some color between pink and red. Similarly, by choosing to utter (45a) instead of (45b), *S*

- (45)
- a. He caused the sheriff to die.
 - b. He killed the sheriff.

chooses a syntactically more complex utterance with more lexical/phonological material. The choice of productive causatives (*cause to p_i*) over lexicalized causatives (such as *kill*) may convey that there is something unusual in the causal relationship under discussion; in (45a), it seems unlikely that *he* actually shot the sheriff -- but more likely that *he* brought about the death in some indirect manner.⁵⁶

It seems likely that further study of particularized implicatures may identify new classes of generalized conversational implicature. In fact, the distinction heretofore made between generalized and particularized conversational implicature may turn out merely to have been an artifact of the limited nature of studies of conversational implicature to date. However, whether this distinction survives, it is clear that currently identified generalized conversational implicatures represent more likely candidates for computation than do those now termed particularized. Of these, the classes that rely primarily upon speaker observance of the Maxim of Quantity appear to be the most promising.

⁵⁶See [McCawley 78] for additional arguments in favor of seeing productive causatives as involving more "effort" than lexical causatives.

3.2. Generalized Quantity Implicature

Generalized quantity implicatures rely for their generation and interpretation upon comparisons of the relative 'strength' or 'informativeness' of utterances made with other utterances a cooperative speaker might have made in a given context. Horn's [Horn 72] SCALAR PREDICATION implicatures, Harnish's [Harnish 79] QUANTITY-QUALITY implicatures, and Gazdar's [Gazdar 79a] SCALAR QUANTITY IMPLICATURES all represent attempts to characterize those conversational implicatures which rely for their generation and interpretation on the notion that a cooperative speaker will say as much as s/he truthfully can (and, in Harnish's cases, that is relevant to an exchange). These classes form the basis for my definition of SCALAR IMPLICATURE. In this section I will describe Horn's and Harnish's intuitive identification of their classes of quantity implicature and then introduce the class of scalar implicatures.

3.2.1. Scalar Predication

In a study of the relationship between logical operators -- in particular, negation -- and their natural-language counterparts, Horn [Horn 72] examines those generalized quantity implicatures that can be licensed by *S*'s mention of what Horn terms SCALAR PREDICATES; these predicates include the natural-language counterparts of the logical operators, quantifiers, modals, and connectives, as well as cardinals, ordinals, and numerous miscellaneous modifiers.⁵⁷

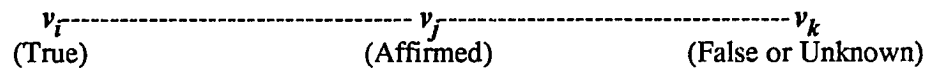
Following Grice's [Grice 65:451] statement that "One should not make a weaker statement rather than a stronger one unless there is a good reason for so doing", Horn defines SCALAR PREDICATION as follows: When a cooperative speaker refers to a value v_j on some SCALE Sc , where Sc is defined by SEMANTIC ENTAILMENT,⁵⁸ that v_j will represent the highest value on Sc that *S* can affirm while observing the Maxims of Quantity and Quality. If v_j entails v_i , then v_j can be seen as a HIGHER value on Sc than v_i . Then, any v_k that is higher on Sc than v_j

⁵⁷Horn's focus is the lexical incorporation of negation. Contending that languages contain only the lexical items they need, Horn proposes that, if the use of some lexical item licenses the implicature p_j , it is unlikely that p_j will itself be lexicalized. [Horn 72:205] For example, if the use of *some* conversationally implicates the negation of stronger predicates on a quantifier scale, such as *all*, then the negation of these stronger predicates, e.g., *not all* need not be lexicalized.

⁵⁸Horn's definition of semantic entailment is identical to Gazdar's one-sided entailment: v_j semantically entails v_i when v_i is true under every assignment of truth values (i.e., in every possible world) under which v_j is true. False values semantically entail nothing for both Horn and Gazdar, lest they be forced to define the invalidities as semantically entailing all sentences.

is implicitly marked (by S 's assertion of v_j) as either false or unknown, depending upon the 'distance' between v_k and v_j on Sc .⁵⁹ That is, *as far as S knows*, $\neg v_k$. And any v_i on Sc that is lower than (entailed by) v_j will be implicitly marked as true. In sum, for the scale depicted in Figure 3-1, if v_k entails v_j entails v_i , then v_k is a higher value

Figure 3-1: Affirming a Scalar Value



than v_j , and v_i is a lower value than v_j . S 's affirmation of v_j will implicate that higher values v_k are either false or unknown and will imply that lower values v_i are true. Implicatures arising via scalar predication, then, are always associated with corresponding implications.

For example, Horn would say that, by uttering (46a), S implicates $\neg(46b)$, based upon a

- (46)
- a. Some of the people left early.
 - b. All of the people left early.

quantifier scale *some/all*. That is, in saying (46a), S is saying as much (Maxim of Quantity) as s/he truthfully (Maxim of Quality) can. The 'working out' process may be understood by S and H as follows: When S affirms (46a), H may infer that, with equal brevity, S might have affirmed (46b). Surely, if a cooperative S could have truthfully affirmed (46b) s/he would have, since, by affirming (46b), S would have implicitly affirmed (46a), since (46b) semantically entails (46a). If S is indeed obeying Quantity and Quality, it must be the case that S is unable to affirm higher values in general and *all* in particular -- i.e., that, for S , higher values are either false or unknown. While values between *all* and the mentioned *some* may be marked as unknown, Horn claims that S 's failure to affirm *all* forces the inference that that value is known by S to be false, since *all* represents the positive poll of the quantifier scale (See Section 4.2.1.). Hence, by asserting (46a), S implicates $\text{KNOW}(S, \neg(46b))$.

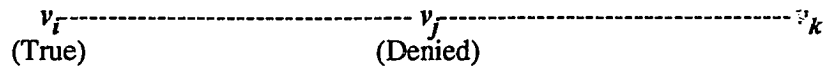
Denying scalar values will also license conversational implicatures. However, for Horn, where S may affirm a scalar value to establish an upper bound on some scale, s/he may deny a scalar value to establish a lower bound. For example, Horn claims that (47a) uttered with 'unmarked' intonation will license (47b).

⁵⁹See Section 4.2.1 for a discussion of Horn's notion of 'distance'.

- (47)
- a. I don't have three friends.
 - b. I have fewer than three friends.
 - c. I have more than three friends.

That is, for the scale illustrated in Figure 3-2, the denial of v_j will implicate that lower values v_i are true.⁶⁰

Figure 3-2: Denying Scalar Values



However, Horn claims that, when (47a) is uttered with stress on *three*, it will license (47c). [Horn 72:70] He is apparently referring to the utterance of (47a) with FALL-RISE intonation.⁶¹ Clearly, simple 'contrastive' stress on *three* does not necessarily produce a lower-bound interpretation (See (48a).), although, such an interpretation is

- (48) A: So, you say you don't have two friends to cosign this loan?
 a. B: No, I don't have THREE friends.
 b. B: Well, I don't have \three/ friends.

clearly possible, as in 49; in this example, B does convey that Gopalan smokes more than three cigarettes a day by stressing *three*.

- (49) A: I thought Gopalan said he smokes about three cigarettes a day.
 B: It's not THREE cigarettes a day.

However, even with FALL-RISE on *three* ((48b)) one can get an upper or a lower-bound reading. So, contra Horn, intonation does not appear to disambiguate negation of cardinals. Thus, denial of scalar values poses a problem for Horn's theory.

⁶⁰Horn does not predict that speakers may license their ignorance of such lower values by denial, which would complement his discussion of implicatures arising via affirmation. Also note that, since Horn's definition of semantic entailment is one-sided, he cannot predict that the denial of v_j will semantically entail the falsity of v_k – i.e., that *S* does not have *four* friends, or *five*. However, this sense certainly seems to be licensed in some way by the utterance.

⁶¹FALL-RISE is a type of FALLING-RISING intonational contour, describable in Pierrehumbert's [Pierrehumbert 80] system as L*+H L H%. It is distinguished from other falling-rising contours in that it is a SCOOPED contour (Cf. Ladd 1980), that is, one in which pitch peak is reached late in the accented syllable (SYL). (In examples below, SYL will be marked as 'SYL'.) There may be more than one such SYL within the FALL-RISE contour. For each SYL, a relatively abrupt drop in pitch must occur within the following two syllables. In addition, FALL-RISE is characterized by a sentence-final rise in pitch. See [Ward 85a] for a detailed study of this contour and its meaning.

Further problems are presented by the inadequacy of Horn's entailment definition of scale to encompass all the quantity implicatures he finds intuitively similar. Although most of the 'canonical' scales Horn identifies (which include the quantifiers and cardinals, as well as the natural-language counterparts of the logical modals (epistemic and deontic) and connectives, and the ordinals)⁶² can be defined in terms of his semantic entailment, he notes that implicatures licensed via adjectival scales such as *cold/ cool/ tepid/ warm/ hot* or *ugly /plain /pretty /beautiful* are difficult to account for. A hot soup may also be a warm soup, but is it also a tepid one? Surely it is not cold. Yet such scales **do** appear to support some sort of quantity implicature, as when B licenses the inference that he believes in some value less than warm in 50:

- (50) A: It's not cool outside.
B: Well, it's not warm.

And Horn notes that other rankings clearly cannot be defined in terms of his semantic entailment, but **can** also support quantity implicatures. He claims that a scale *tort/ misdemeanor/ felony/ capital crime* which cannot be defined by entailment (if a crime is a felony it is neither a misdemeanor nor a tort) does permit quantity implicature, as in 51, in which utterance of (51a) can implicate 'for all *S* knows $\neg(51b)$ '.⁶³

- (51)
a. Smoking marijuana is (at least) a misdemeanor in most states.
b. Smoking marijuana is (at least) a felony in most states.
c. #?Smoking marijuana is a misdemeanor and in fact it's a felony in most states.

However, while (51a) indeed appears to this meaning, it is not clear that this meaning is conveyed by conversational implicature. For $\neg(51b)$ (assuming 'at least' is not part of the utterance under consideration) does **not** appear to be cancelable (See (51c).). So, (51a) may license $\neg(51b)$ by means other than conversational implicature; logical implication from exclusive disjunction of crime types is a plausible way to model this. However, other orderings impossible to define via entailment more clearly convey conversational implicatures, such as the ordering inferrable as salient in 52.

- (52) A: Did you get Paul Newman's autograph?
B: I got Joanne Woodward's.

⁶²These will be described in Chapter 5.

⁶³I believe that '(at least)' in these examples is intended only to identify the sense of the utterance -- not to represent part of it.

3.2.2. Quantity-Quality Implicature

An apparently independent proposal of Harnish's [Harnish 79] considers additional metrics that support implicatures quite similar to those described by Horn. Harnish proposes a submaxim of the maxims of Quantity and Quality as the basis for these implicatures:⁶⁴

Maxim of Quantity-Quality: Make the strongest relevant claim justifiable by your evidence.

He claims that *S* may license quantity-quality implicatures if s/he is assumed by *H* to be observing this maxim and if *S* and *H* can compare the 'strength' of an utterance with the 'strength' of possible alternatives. By committing him/herself to some u_i , *S* can implicate that s/he is **not** able to commit him/herself to any stronger u_j .

So, in Chomsky's [Chomsky 72] example (illustrated in 53),⁶⁵ where Horn might explain the implicature ($\neg(53b)$) licensed by the utterance of (53a) as *S*'s assertion of the highest

- (53)
- a. One of my five children is in elementary school.
 - b. Two of my five children are in elementary school.
 - c. #Two of my five children but not one are in elementary school.

value on a cardinal scale that s/he can truthfully assert, Harnish compares 'strength of claim': If *S* asserts (53a), and if the number of children *S* has in elementary school is RELEVANT in the discourse, the strength of *S*'s claim can be measured by the number of children about which s/he predicates *be-in-elementary-school*. (53b) is a stronger assertion than (53a) because, "in simple quantitative terms", Harnish claims, *two* is 'stronger' than *one*. If *S* were to assert (53a) while believing (53b) to be true, s/he would **not** be obeying the Maxim of Quantity-Quality. So, if *S* assumes *H* believes him/her to be cooperative, *S* may implicate $\neg(53b)$ by saying (53a). In Horn's terminology, Harnish would presumably say that, not only must scalar values be mentioned, but the scale on which they appear must be relevant.

While, like Horn, Harnish finds that one way to measure strength of claim is via utterance entailment,⁶⁶ he proposes numerous other metrics as well. So, the implicature licensed via (53a) may be explained in terms of entailment: Having two children in elementary school

⁶⁴Although it would seem that Harnish's maxim is dependent upon Relation as well, he does not include it.

⁶⁵I have changed the order of the sentences.

⁶⁶His treatment of entailment is less satisfactory than Horn's and Gazdar's one-sided entailment -- and is not, in fact, an actual definition: For Harnish, "if an utterance u_j entails another u_i , then u_j and $\neg u_i$ is a contradiction. If u_j entails u_i and u_i does not entail u_j , then u_j may be seen as stronger or more informative than u_i ." Invalidities are excluded separately from any semantic entailment relationship.

entails having one child in elementary school -- but having one child in elementary school does not entail having two; so (53b) is a stronger assertion than (53a). However, other quantity-quality implicatures, rely upon different measurements of 'strength of claim'. For example, Harnish notes that the denial of some p_i for a specified time period may license the belief that p_i held for some prior time period. So, the utterance of (54a) can implicate (54b).

- (54)
- a. I haven't played tennis for ten years.
 - b. I played tennis before (ten years ago).

However, Harnish does not go beyond suggesting numerous additional metrics which appear to license related conversational implicatures. Furthermore, all the metrics he suggests appear to represent linear orderings.

Like Horn, Harnish notes in passing that quantity implicatures may be licensed via denial. For example, he claims that the denial of (55a) (i.e., "Not all of the tables are round.") can implicate that (55b):

- (55)
- a. All the tables are round.
 - b. Most of the tables are round.
 - c. Some of the tables are round.

By Harnish's definition of the relative 'strength' of utterances, the denial of (55b) is stronger than the denial of (55a), since $\neg(55b)$ entails $\neg(55a)$.⁶⁷ So, uttering $\neg(55a)$ can implicate $\neg\neg(55b)$, or, (55b). So, Harnish provides a more principled and somewhat neater explanation for implicatures resulting from denial than does Horn. However, Harnish does not explain why $\neg(55a)$ does not convey (55c) instead of (55b). Nor does he consider problems of intonational variation and scope ambiguity.

3.2.3. Defining Classes of Quantity Implicature

As I suggested in Chapter 2, it appears simpler from a computational point of view to identify classes of generalized conversational implicature which, as classes, can be shown to satisfy the definition of conversational implicature presented in that chapter -- and then to test candidate instances for membership in these classes -- than it is to test each candidate instance by a general definition of conversational implicature. For example, if we can define classes of generalized quantity implicature from Horn's and Harnish's descriptions of scalar predication implicatures and quantity-quality implicatures, and if these classes can be shown to satisfy the

⁶⁷Note that this holds only because Harnish's definition of entailment is two-sided -- i.e., if p_i entails p_j then $\neg p_j$ entails $\neg p_i$.

conditions on conversational implicature proposed in Chapter 2, then any identifiable instance of these classes will also represent a conversational implicature and exhibit the features of the more general category.

It is possible to make some headway with such a definition from Horn's and Harnish's characterizations: Although Condition 1 on speaker intention is stronger than anything Horn and Harnish propose, it will apply to scalar predication implicature and quantity-quality implicature as to any conversational implicature. Condition 2 (mutual belief that *S* is observing the maxims) is also stronger than Horn's and Harnish's specification that *S* must be presumed cooperative; however, since the latter is itself equivalent to the Gricean condition which 2 subsumes (See Section 2.4.2.), this presents no difficulty. More interestingly, however, is the way that Horn's and Harnish's intuitions provide a basis for specifying circumstances under which Condition 3 (mutual belief that some inference is licensed -- given speaker obedience to certain maxims M_i and utterance of some u_i in a context C_h) may be satisfied.

Condition 3 can be satisfied in Horn's terms for some inference p_j as follows: *H*'s belief in p_j will be licensed when *S* and *H* mutually believe that *S* is obeying the Maxims of Quantity and Quality in affirming some scalar predicate v_i (via an utterance that realizes p_i) when p_j includes a higher value v_j than v_i (but is otherwise identical to p_i). To formalize this, we might, for the moment, make the following simplifying assumptions: Identify v_i and v_j with the semantic representations p_i and p_j of propositions which affirm these values, ranking p_i and p_j indirectly via the ranking of v_i and v_j . So, when v_j is higher on some scale Sc than v_i (i.e., when v_j entails v_i) we will say that p_j is higher than (entails) p_i , or, $HIGHER(p_i, p_j, Sc)$. Then, for Horn's class of scalar predication implicatures, Condition 3 may be satisfied if scalar predication occurs: That is,

Conversational Implicature via Scalar Predication:

$$(IS_COOP(S, C_h, \{QUANTITY, QUALITY\}) \wedge SCALAR_PRED(S, H, u_i, p_j)) \\ \Rightarrow LICENSE(S, H, u_i, p_j, C_h, \{QUANTITY, QUALITY\}).$$

Then we can define SCALAR_PRED as follows (again, assume universal quantification with widest scope over unquantified variables):

Scalar Predication:

$$\exists Sc \forall p_j \{ (REALIZE(u_i, p_i) \wedge HIGHER(p_j, p_i, Sc)) \Rightarrow SCALAR_PRED(S, \\ H, u_i, KNOW(S, \neg p_j) \vee \neg KNOW(S, p_j)) \}$$

That is, for all speakers, hearers, and utterances, if, for some u_i realizing p_i , there is some Sc ranking p_j higher than p_i , then, for all such p_j , *S* may license that s/he knows p_j to be false or that s/he does not know p_j to be true -- by saying u_i . Since Horn does not really take context into account at this level, this lack is reflected in the definition of SCALAR_PRED. Note that implicatures licensed via the denial of scalar predicates will require a separate axiom.

Quantity-quality implicatures might be represented similarly: Let $\sigma_i, \sigma_j, \dots$ range over measurements of strength of utterance which include but need not be limited to those mentioned by Harnish. Let $\text{STRONGER}(u_j, u_i, \sigma_i)$ denote that u_j is stronger than u_i when measured by some metric σ_i ; STRONGER might be at least partially defined in terms of Harnish's entailment. Then specify the relevance of σ_i in a context C_i by $\text{RELEVANT}(\sigma_i, C_i)$.⁶⁸ Then we can describe sufficient conditions for Condition 3 via Harnish's theory similarly to those defined for Horn's:

Conversational Implicature via Quantity-Quality Implicature:

$$(\text{IS_COOP}(S, C_h, \{\text{QUANTITY-QUALITY}\}) \wedge \\ \text{QUANTITY_QUALITY_IMPLIC}(S, H, u_i, p_j, C_h)) \Rightarrow \text{LICENSE}(S, H, u_i, \\ p_j, C_h, \{\text{QUANTITY, QUALITY}\}).$$

And, like SCALAR_PRED , $\text{QUANTITY_QUALITY_IMPLIC}$ can be defined as:

Quantity-Quality Implicature:

$$\exists \sigma_i \forall u_j \{(\text{RELEVANT}(\sigma_i, C_h) \wedge \text{REALIZE}(u_j, p_j) \wedge \text{STRONGER}(u_j, u_i, \\ \sigma_i)) \Rightarrow \\ \text{QUANTITY_QUALITY_IMPLIC}(S, H, u_i, \neg p_j, C_h)\}$$

That is, if u_j is stronger than u_i with respect to some σ_i which is relevant in C_h , and if u_j realizes p_j , then S may license $\neg p_j$ by saying u_i to H in C_h .

Note that the weaknesses of Horn's and Harnish's theories pointed out above are reflected in these axiomatizations of Condition 3, including Horn's lack of incorporation of context or relevance into his specification of how quantity implicature is licensed and Harnish's vagueness about measures of strength and relevance. These issues must be resolved before a successful definition of quantity implicature is possible.

It is difficult to test Horn's and Harnish's classes of implicatures for the remaining conditions identified in Chapter 2 -- cancelability, nondetachability, and reinforceability -- unless we assume entailment definitions of scale/measure of strength. If so, it seems that these classes, *qua* classes, can be shown either to be cancelable or to represent conventional implicatures as follows: Recall that cancelability rests upon the felicity of the conjunction of some (utterance that realizes) p_i to the denial of a p_j which it might otherwise license; if the implicature $\neg p_j$ is cancelable, then ' $p_i \wedge \neg p_j$ ' -- or, ' $p_i \wedge p_j$ ' -- is felicitous. So, for some quantity-quality implicature $\neg p_j$ licensed by an utterance u_i , $\text{CANCELABLE}(u_i, \neg p_j) \leftrightarrow \text{CAUSE}(\text{SAY}(S, H, u_i, C_h), \text{BEL}(H, \neg p_j)) \wedge \text{FELICITOUS}(p_j, \{C_h \cup u_i\})$ (See Section 2.4.2.1.).

⁶⁸Since Harnish nowhere attempts to define his notion of 'relevance', I will employ this simple predicate.

For Harnish's quantity-quality implicatures that are defined by an entailment metric, $\text{CAUSE}(\text{SAY}(S, H, u_i, C_h), \text{BEL}(H, \neg p_j))$ will generally be true.⁶⁹ So we must argue the felicity of asserting p_j when (in effect) ' $p_j \wedge p_i$ ' may be felicitously affirmed.

First we will argue that ' p_j and p_i ' does not represent a contradiction: For implicatures licensed via an entailment metric, Harnish states that, if u_j entails u_i , then ' u_j and $\neg u_i$ ' (or, its representation ' $p_j \wedge \neg p_i$ ') represents a contradiction. Furthermore, for these implicatures, he demands that u_j itself not realize a logical invalidity. But, if ' $p_j \wedge \neg p_i$ ' is contradictory, then ' $p_j \wedge p_i$ ' is not.

Of course, even if the utterance of ' $p_i \wedge p_j$ ' is not contradictory, it may nonetheless be infelicitous if ' $\neg p_j$ ' is conveyed via conventional implicature. Recall that, if ' $p_i \wedge p_j$ ' is infelicitous but not contradictory and, if ' $\neg p_j$ ' does not represent a conversational implicature, then ' $\neg p_j$ ' will represent a conventional implicature. So, meanings licensed via entailment are cancelable -- or they represent conventional implicatures. Similarly, Horn's implicatures can be shown to be either cancelable or conventional when scales are defined by entailment.

We could eliminate the second conjunct -- conventionality -- by showing that these meanings must be nondetachable (Condition 5) -- since conventional implicature is not nondetachable. While Harnish does not provide enough information to permit testing this condition, Horn does. So, we can rule out the possibility that implicatures arising via scalar predication are conventional by arguing that they will all be nondetachable: Horn specifies that his scalar predicates are not lexical items but rather semantic concepts. Since, as noted in Section 2.4.2.2, nondetachability may be seen as just semantic identity, it would seem that implicatures conveyed via scalar predication must be nondetachable. So, Conditions 4 (cancelability) and 5 (nondetachability) appear to be satisfied for those implicatures licensed via entailment scales.

Finally, entailment-defined scalar predication implicatures and quantity-quality implicatures can be shown to satisfy Condition 6 (reinforceability) in much the same way they can be shown to satisfy Condition 4: Since, critically, p_j entails that p_i but p_i does not entail that p_j , the affirmation of p_j when p_i has just been affirmed should not be redundant -- unless p_j represents a conventional implicature which -- at least for Horn's class of implicatures -- it cannot. Thus, $p_i \wedge \text{KNOW}(S, \neg p_j)$ or $p_i \wedge \neg \text{KNOW}(S, p_j)$ should not be (internally) redundant, and implicature arising via entailment-defined scalar predication will be reinforceable.

⁶⁹See Section 5.1 for criticisms of some of the particular implicatures Harnish claims.

In sum, it is possible to define a class of generalized quantity implicature in the terms described in Chapter 2 from Horn's and Harnish's descriptions -- if we assume an entailment definition of scale/measure of strength and if we accept the notion of 'relevant' as primitive. The problems associated with these assumptions, of course, are just those weaknesses of Horn's and Harnish's theories identified above: the need to incorporate a notion of relevance in context and of context itself into the identification of quantity implicatures and the need to identify an adequate way to compare the relative informativeness of utterances.

3.2.4. Summary

Horn's and Harnish's accounts represent independent attempts to define comparative informativeness in the Gricean framework at an intuitive level. While Horn's scalar predication is a simpler and better defined phenomenon -- and also better known -- Harnish's quantity-quality implicature encompasses a broader notion of what appears to be the same class of quantity implicatures. But, although Harnish suggests that a much broader range of relations may license these implicatures than Horn's semantic entailment, neither Harnish nor Horn is able to identify a metric or set of metrics which account for all the implicatures in the class they are defining. Furthermore, all the metrics suggested define linear orderings, when, in fact, implicatures intuitively similar to those they discuss may be supported by other sorts of orderings. Both Horn and Harnish recognize that context will play a role in the licensing of quantity implicatures, but only Harnish suggests how that role might be specified for the class in question. His quantity-quality implicatures will be licensed only when the metric used to compare strength of claim is relevant in the context. However, he does not discuss how such relevance might be determined. Although both scalar predication implicature and quantity-quality implicature suggest ways in which quantity implicature might be defined in terms of the definition of conversational implicature presented in Chapter 2, neither is sufficiently precise or inclusive to permit a comprehensive definition.

3.3. Scalar Implicature

While scalar predication and quantity-quality implicature capture some important generalizations about an intuitively coherent class of quantity implicatures, a study of naturally occurring discourse suggests that Horn's and Harnish's intuitions may be extended to account for a still broader class of implicatures, which I term SCALAR IMPLICATURE. Scalar implicature subsumes those quantity implicatures described by Horn and Harnish, and also includes subclasses which differ from those described by Horn and Harnish in three important ways: First, they rely upon orderings that are not linear, as well as upon metrics other than entailment.

Second, they involve inferences about utterances that reference -- in Horn's terminology -- not only the higher and lower values Horn and Harnish allow, but what I will term ALTERNATE values as well. Roughly speaking, these are values which are neither higher nor lower than one another but which share a common higher or lower value. Third, they arise not only from affirmation or negation of values (as allowed by Horn and Harnish) -- but also from a speaker's commitment to ignorance of some value. In the remainder of this section I will describe the extensions to scalar predication implicature and quantity-quality implicature that define the larger class of scalar implicatures.

To avoid present and future terminological confusion, I will denote the relationships which support scalar implicature (which Horn terms 'scales' and Harnish 'measures of strength of claim') ORDERINGS (represented by O), and will call the measuring principles which define such orderings ORDERING METRICS.⁷⁰ Since Horn, Harnish, and Gazdar require these notions only to rank utterances via values contained in them as 'higher' or 'lower' -- or 'stronger' or 'weaker' -- than other utterances, I will specify for now only that O be such that it will support these intuitive distinctions. I will adopt Horn's nomenclature for items ranked by these metrics in such orderings, which I too will call VALUES. Variables σ_i, σ_j ranging over metrics have already been defined, as have variables denoting values. However, I will also introduce new variables O_i, O_j ranging over orderings, which I will define more precisely in Chapter 5.

3.3.1. Additional Measures of Informativeness

Many utterances that appear to license implicatures similar to those identified by Horn and Harnish refer to entities, attributes, events, or states that cannot easily be viewed as values in linear orderings defined by entailment. Such items may be viewed as ordered hierarchically, as by some whole/part, type/subtype, entity/attribute, or set/subset relationship.

For example, in 56, B

- (56) A: Did you manage to read that section I gave you?
 B: I read the first couple of pages.

mentions a part (*first couple of pages*) of a whole (*section*) to implicate that he has not read the whole. In Horn's terms, he affirms the highest value he can -- but, here, it is a 'value' in a part/whole hierarchy.

57 and 58 illustrate implicatures that may be licensed via recognition of some

⁷⁰In previous work on scalar implicature I have termed such posets 'scales' after Horn and Gazdar. However, confusion over intuitive definitions of scales as linear orderings has convinced me to abandon this terminology.

type/subtype ordering among referenced items: In 57, B's denial of having made a particular type of

- (57) A: Have you made fondue in this pot yet?
B: Not chocolate fondue.

fondue conveys that she has made other types of fondue in the pot in question; in fact, as B later explained, A had given her the pot as a wedding gift and she was embarrassed to admit to not having used it before. So, she implicated having previously used the pot.⁷¹ In 58, B licenses the implicature that Leo does not like the other flavor⁷²

- (58) A: So, Leo likes Bonkers?
B: She likes liver flavor.

of the queried cat treat -- chicken. Note that, even if A does not know that there are several flavors of Bonkers, B's response will convey that other flavors exist. Type/subtype relations such as these may be seen as ordered by entailment from type to subtype, so that subtypes represent higher values in that they entail lower supertypes.

Set/subset relationships and set/member relationships will also license quantity implicatures. In 59

- (59) A: Have you gotten the letters yet?
B: I've gotten the letter from X.

B conveys that he has not received all the letters to which A refers -- by 'saying as much as he truthfully can' in affirming receipt of the *letter from X*. Set/subset orderings are seen as sets that entail all their subsets.

Finally, exchanges like 60 may rely upon several different metrics

- (60) A: Do you have Lana Moro or Bernat yarns?
B: We have Cassino.
A: Do you have Cassino in sweet violet?
B: We have Cassino.

to convey implicit information: In a telephone call to a yarn shop, A asks whether the shop sells two brands of yarn, *Lana Moro* and *Bernat*. The clerk replies with a type of Bernat yarn, *Cassino* -- implicating that she does not carry Lana Moro at all and that the only type of Bernat that she carries is Cassino. When A further inquires about a particular color of Cassino, B

⁷¹In the end, a compulsively honest B finally admitted her misleading implicature to A.

⁷²Cats have little choice.

conveys that she does *not* have the desired color.⁷³ So, in this exchange, A and B have utilized a set/subset ordering (*Lana Moro* and *Bernat* form a set of brands of yarn), another set/member ordering (*Cassino* is a member of the set of *Bernat* yarns), and a type/subtype or entity/attribute ordering (*sweet violet* is a type of *Cassino* yarn or it is the value of an attribute of this yarn, namely, its color.) Such entity/attribute orderings may be ordered from attributes to entity, such that an entity entails its (definitional) attributes.

It should be noted that the responses in 56 through 60 are appropriate not simply by virtue of their closeness of match to the topic of the preceding query. It is the asymmetry of the relationship between mentioned values -- or between mentioned values and some salient third value -- that is essential to explain scalar implicatures. So, to the linear metrics identified by Harnish and Horn as supporting their quantity implicatures, I will add these hierarchical relationships.

3.3.2. Higher, Lower, or Alternate Values

While Harnish and Horn concentrate on quantity implicatures generated regarding 'higher' value or 'stronger claims', scalar implicatures include inferences licensed about lower values/'weaker claims', and alternate values/claims. In 56, 59, and 60, for example, speakers convey information about higher values by affirming a lower value -- i.e., they convey the implicit information that they are unable to affirm a higher value. In 57 and in 61, speakers convey that they are unable

- (61) A: Was he cute?
B: He wasn't stunning.

to falsify lower values. In the latter, for example, B licenses the inference that, while he cannot deny *cute*, he cannot affirm *stunning*.

Note that conveying that higher values are false by denying a lower value, as in 62 and 63, does not follow from conversational implicature, but from logical implication. In 62, A and B are discussing who will pay for a trip A will be taking:

- (62) A: Maybe she thinks that X should be paying for all of it.
B: She shouldn't be paying for you to go to COLING.

In this exchange, B implies that, since X shouldn't be paying for part of the trip, she shouldn't

⁷³Some readers have found this interpretation hard to credit. With FALL-RISE over *Cassino*, A might have interpreted this response as a declaration of ignorance about available colors. But with *Cassino* deaccented, the interpretation presented here was clear. B indeed knew that she had no 'sweet violet'. It seemed that she imagined A might purchase some other color.

be paying for all of it. In 63, similarly, where A inquires about the whereabouts of several cakes, B conveys that she *did* not eat all of them, since she did not eat the

- (63) A: So, did you snarf all the cake down?
B: I didn't eat the chocolate one.

chocolate cake. However, in both cases B also licenses *H*'s belief that she cannot deny unmentioned values, i.e., the rest of the trip and the rest of the cakes. So, as far as B knows *she* should be paying for the rest of the trip (62) and she *did* eat the rest of the cake (63). Similarly, when speakers convey that lower values are true by affirming a higher value, as in 64 or 65, they convey this meaning too by logical

- (64) A: Dan thinks I'm stupid.
B: No, he knows you're stupid.

implication. In 64, B conveys that Dan not only thinks A is stupid -- he knows it. And, in 65, B implies that he'll be

- (65) A: So you'll be here all semester?
B: I'll be here all year.

here all semester, since, in fact, he'll be here all year.

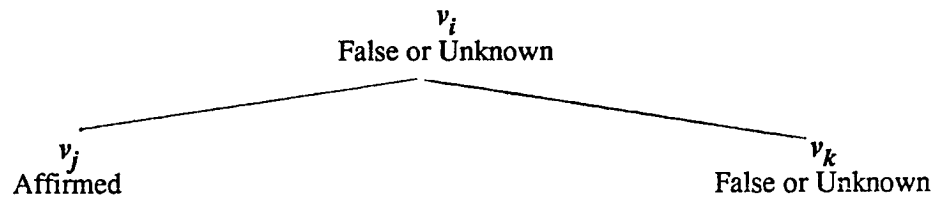
It should be noted that, although scalar implicature does not play a role in the licensing of implications via such utterances, the anticipation of possible scalar implicature does provide an account of a speaker's decision about which utterance to make in such cases: In all the above exchanges, a simple *yes* or *no* would be truthful. Recall the specification in Section 2.4.3 of conditions under which *H* will be entitled to infer that *S* has implicated that p_j . If B says only *yes* in 65, for example, he would (in Horn's terms) be affirming *all semester*; if a temporal ordering is salient in the exchange, as seems reasonable given the query, then A may reasonably be expected to infer that B has implicated that it is not the case that he will be here longer than one semester. That is, A may infer that B has indeed affirmed the highest value he is able to affirm. So, in contemplating making a simple, direct response, B may anticipate that A will infer that B cannot affirm *all year*. So, B may block this inference by providing the indirect response of 65.

Speakers may license implicatures about values that are neither higher nor lower than a queried value, but that share a common subordinate or superordinate. I term these ALTERNATE VALUES; together with the notions of *higher* and *lower* values, they will be defined more precisely in Chapter 5. In 66, for example, B affirms an

- (66) A: So you speak Sephardic?
B: Huh?
A: Do you speak Ladino?
B: I speak Spanish.

alternate to two queried values *Sephardic* and *Ladino*, all from the set of Iberian languages. B implicates that these values are false or unknown by affirming the alternate value, *Spanish*. Figure 3-3 shows implicatures that may be licensed via the affirmation of v_j for higher and alternate values, v_i and v_k :

Figure 3-3: Affirming Alternate Values



In 67, B's denial of *Fundamental Algorithms* licenses the implicature that B is unable to deny the similarly queried *Data Structures*.

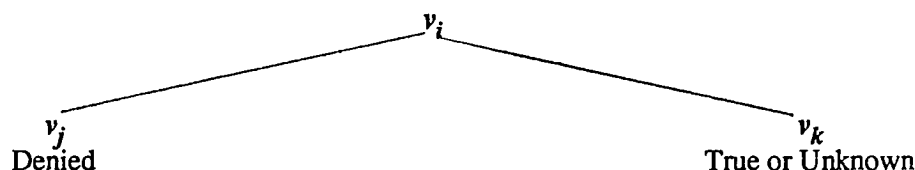
- (67) A: Let me just check whether you have all the prerequisites. You have Calc 1 and 2?
 B: Uh-hmm.
 A: You have Introduction to Programming?
 B: Oh, yeah.
 A: You have Data Structures, Fundamental Algorithms.
 B: No, I don't have Fundamentals.
 A: O.K.

So, as far as B knows B **has** taken Data Structures. Figure 3-4 illustrates the implicatures licensed with regard to alternate values v_k by denying v_j .⁷⁴

In sum, scalar implicatures may be licensed about items intuitively related to a mentioned item as higher/stronger than, lower/weaker than, or alternate to the mentioned item. Additionally, mentioned items may be affirmed or denied to license scalar implicature -- or speakers may license scalar implicature by declaring themselves ignorant of some item.

⁷⁴Note that no scalar implicatures will be licensed with regard to higher values v_i here.

Figure 3-4: Denying Alternate Values



3.3.3. Affirming, Denying, and Declaring Ignorance of Scalar Values

Recall that Horn mentions that quantity implicatures can be licensed by the denial as well as by the affirmation of scalar values; Harnish ranks similar negations in terms of strength of claim. In this section I will propose that scalar implicature may be licensed when speakers declare their ignorance of scalar values -- as well as when they affirm or deny them. First, however, I will propose a hybrid explanation of scalar implicature licensed via denial.

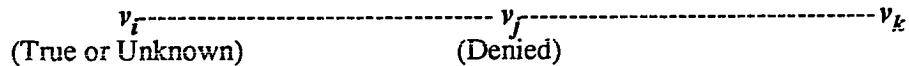
Where Horn explains inferences licensed via denial by postulating that negation of scalar values can set a lower bound, Harnish uses his two-sided definition of entailment to define $\neg p_i$ as stronger than $\neg p_j$ where p_j is stronger than p_i . While this definition is inadequate for non-entailment relations, Harnish's explanation can justify -- by analogy -- the definition of the dual to Horn's notion that cooperative speakers affirm the highest value they truthfully can: Cooperative speakers will deny the lowest value in some salient ordering that they can truthfully deny. By such denials, they affirm or convey their ignorance of lower values. If an ordering $v_i/v_j/v_k$ is defined by entailment, then, by denying v_j , S may convey s/he is unable to deny lower values v_i : If a cooperative speaker could deny v_i , then s/he should do so since, by denying an entailed v_j s/he could also deny the entailing v_i . For example, in 68, B denies *getting a mortgage*, a stage in the process of *purchasing a home*, to

- (68) A: Did you buy a house?
 B: We haven't gotten the mortgage yet.

indicate that earlier stages in this process, such as *making a bid*, are true or unknown. If the latter were false, then, since one cannot attempt to get a mortgage without having made a bid on a house, B might convey this information by denying the lower value *making a bid*. Since B does not deny this lower value, A can 'work out' that, as far as B knows, *making a bid* is true -- or at least not false. So, denying v_j can implicate that S cannot deny a lower value v_i . Figure 3-5 illustrates the implicatures licensed by the denial of a lowest deniable value -- given the

above account.⁷⁵

Figure 3-5: Implicatures Arising Via Denial



Note that, while this account is similar to Horn's (illustrated in Figure 3-2), the two are not identical. Denial of a scalar value will license the scalar implicature that lower values are true or unknown -- not simply true as in Horn's account.

Consider also the following exchange (69) between caller and hospital information clerk:

- (69) A: Do you have information on a patient?
 B: What's the name?
 A: Kathy M. for maternity.
 B: I don't think she's delivered yet.
 A: Then she HAS been admitted.
 B: Yes.

If the *process of having a baby* includes as ordered stages ... / *being admitted* / ... / *delivering a baby* / ..., then A's inference that Kathy M. has been admitted can be explained in terms of scalar implicature: B has denied the lowest value (earliest stage) she truthfully can. Therefore, as far as B knows, prior stages are true, and, so, Kathy M. has been admitted.

Speakers may also license scalar implicatures by declaring their ignorance of scalar values, although regularities here are more difficult to determine. It is unclear what the Gricean maxims might predict to be a speaker's obligations regarding assertion of ignorance about scalar values -- beyond the obligation not to feign ignorance. On the one hand, it seems plausible that *S* assert ignorance about the lowest value in some ordering that s/he is unable to affirm or deny in order to localize the extent of his/her ignorance precisely. For example, in 70 and 71,

- (70) A: Is it warm in Antarctica in the summer?
 B: I don't know if it gets above freezing.

if the orderings *freezing/ cold/ ... / warm/ ...* and *set of shots cats require* are salient, then, in

⁷⁵Note that no scalar implicatures will be licensed about higher values v_k .

- (71) A: So, does Leo need shots this spring?
B: I don't know about rabies.

each case, B may be seen as declaring ignorance of the lowest value s/he truthfully can. In 70, B implicates that, not only is he unaware of whether it is warm in Antarctica in the summer -- but he doesn't even know if it gets above freezing. In 71, B implicates that he is ignorant of only one member of the set of required shots -- rabies -- and intends to convey that he 'knows about' all the other *set of shots she requires*.⁷⁶ On the other hand, it seems equally plausible that a cooperative speaker assert ignorance about the most inclusive value s/he is ignorant of. In 72, for example, B chooses to

- (72) A: Did they have a boy or a girl?
B: I didn't know they'd had it yet.

mention her more general ignorance of the birth of any baby in response to A's query. In either case, the assertion of ignorance about some value v_j appears to license S's belief that higher values v_k in some salient ordering are false or unknown and that lower values v_i are either true or unknown, as illustrated below:

Figure 3-6: Declaring Ignorance of Higher and Lower Values

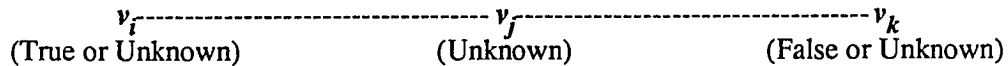
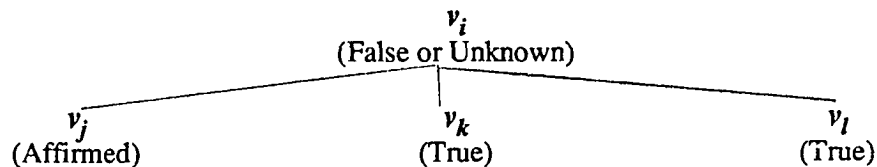


Figure 3-7: Declaring Ignorance of Alternate Values



⁷⁶In fact, B intended to convey his belief that Leo *did not* need the other shots.

3.3.4. Conventions of Scalar Implicature

The intuitive characterization of scalar implicature presented above can be summarized in a set of conventions, Imp_{1-3} , which capture the different subclasses of scalar implicatures licensed via affirmation, denial, and declaration of ignorance -- for higher, lower, and alternate values. Given an ordering O that is salient in a context C_i , a value v_j in O , and an utterance u_j affirming, denying, or asserting ignorance of v_j :

- **Imp₁**: If u_j affirms v_j , then for all v_k such that v_k is higher in O than v_j , S may license the scalar implicature that v_k is false or unknown; and, for all v_l such that v_l and v_j are alternate values in O , S may license the scalar implicature that v_l is false or unknown.
- **Imp₂**: If u_j is a denial of v_j , then for all v_i such that v_i is lower in O than v_j , S may license the scalar implicature that v_i is true or unknown; and, for all v_l such that v_l and v_j are alternate values in O , S may license the scalar implicature that v_l is true or unknown.
- **Imp₃**: If u_j is an assertion of ignorance of v_j , then for all v_i on O , such that v_i is lower than v_j , S may license the scalar implicature that v_i is true or unknown; for all v_k on O such that v_k is higher than v_j , S may license the scalar implicature that v_k is false or unknown; and, for all v_l such that v_l and v_j are alternate values in O , S may license the scalar implicature that v_l is true.

Note that, instead of Harnish's notion of 'relevance', I have employed an equally vague notion of 'salience', to denote that which is being attended to. Salience will be defined here as 'that which is most likely to be attended to'. If a scale is salient it is not only *relevant* in a particular context but it is that scale whose relevance a speaker/hearer is most likely to recognize. To license a scalar implicature, I will maintain that S must believe that some ordering is mutually believed salient by S and H , or, $\text{BMB}(S, H, \text{SALIENT}(O, C_h))$. I will discuss the incorporation of 'salience' into the calculation of scalar implicature in detail in Chapter 6.

Following the strategy introduced above in Section 3.2.3, this intuitive characterization of scalar implicature may be used to define a class of quantity implicature in the formalism proposed in Chapter 2 much as Horn's and Harnish's characterizations may. In particular, we can say that

$$(\text{IS_COOP}(S, C_h, \{\text{QUANTITY, QUALITY}\}) \wedge \text{SCALAR_IMP}(S, H, u_i, p_j, C_h)) \Rightarrow \text{LICENSE}(S, H, u_i, p_j, C_h, \{\text{QUANTITY, QUALITY}\})$$

Then Imp_{1-3} can be axiomatized as follows, where p_i and p_k are identified with the values v_i and v_k they make reference to, and variables are universally quantified as before:⁷⁷

⁷⁷The epistemic force assigned implicatures in the following conventions will be justified in Section 4.2.2.

- Imp₁: $\exists O \{(\text{BMB}(S, H, \text{SALIENT}(O, C_h)) \wedge \text{REALIZE}(u_i, \text{AFFIRM}(S, e_i, p_i)) \wedge$
 $(\text{HIGHER_SENT}(p_i, p_j, O) \vee \text{ALT_SENT}(p_i, p_j, O)))$
 $\Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, p_j), C_h)\}$
- Imp₂: $\exists O (\text{BMB}(S, H, \text{SALIENT}(O, C_h)) \wedge \text{REALIZE}(u_i, \text{DENIAL}(S, e_i, p_i)) \wedge$
 $(\text{LOWER_SENT}(p_i, p_j, O) \vee \text{ALT_SENT}(p_i, p_j, O)))$
 $\Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, \neg p_j), C_h)$
- Imp₃: $\exists O (\text{BMB}(S, H, \text{SALIENT}(O, C_h)) \wedge \text{REALIZE}(u_i, \text{IGN}(S, e_i, p_i))) \Rightarrow$
 $(\text{LOWER_SENT}(p_j, p_i, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, \neg p_j), C_i))$
 \vee
 $(\text{HIGHER_SENT}(p_j, p_i, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, p_j), u_i,$
 $C_i)) \vee$
 $(\text{ALT_SENT}(p_i, p_j, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \text{BEL}(S, p_j), C_i))$

For scalar implicature, however, as for scalar predication and quantity-quality implicature, until we can specify the relationship between utterances and scalar values formally, until a satisfactory definition of ‘ordering’ provides a semantics for HIGHER, LOWER, and ALTERNATE predicates; until SALIENT can be specified in some way; until the meaning of ‘affirming’, ‘denying’, and ‘declaring ignorance of’ values is specified; and until the role of context can be incorporated in these axioms, such a definition will be of limited use.

3.4. Summary

In this chapter I have described the major characterizations of quantity implicature in the literature and have discussed some of their deficiencies. I have also sketched out a new form of quantity implicature which I term SCALAR IMPLICATURE and have described some of the problems that will be involved in its definition.

It should be clear from this discussion that the chief obstacles to a computational treatment of generalized quantity implicature are: the lack of an acceptable measurement of ‘strength of claim’ or ‘quantification of informativeness’; the problem of determining whether a particular ordering is salient in a particular exchange; and the incorporation of context into the calculation of implicatures. In addition to these, certain representational problems must also be considered: In particular, Horn is less than clear and Harnish silent on the question of a speaker’s degree of commitment to their quantity implicatures. And neither specifies a formal representation for these implicatures. Both of these problems must be resolved to permit a computational approach to scalar implicature.

In Chapter 4, I will discuss problems of representation and will justify the epistemic force I have associated with scalar implicatures in the presentation of scalar implicature conventions

above; in Chapter 5, I will propose a new way of 'measuring informativeness'; and in Chapter 6, I will discuss how the salience of such a measurement in a discourse context might be assessed.

CHAPTER IV

Representing Scalar Implicature

From a discussion of the local pubs and eccentric things overheard therein on Islay, an island in the Inner Hebrides off the coast of Scotland: "Just two things overheard recently: 'How many tractors have you got now, Angus?' 'I have two -- and another one.'" New York Times Sunday Travel Section, 10 March 1985

The only serious formal account of conversational implicature in general and of quantity implicature in particular is found in [Gazdar 79a]. In his computational solution to the PROJECTION PROBLEM,⁷⁸ Gazdar formalizes Horn's scalar predication as SCALAR QUANTITY IMPLICATURE. Aspects of this formalism can be adapted for the representation of scalar implicature.

In this chapter I will propose a representation for scalar implicature within the more general representation of conversational implicature laid out in Chapter 2 -- based in part upon Gazdar's formalization of Horn. In particular, I will adopt Gazdar's approach to utterance representation and his syntactic approach to the derivation of quantity implicatures. However, contra Horn and Gazdar, I will propose a representation of the epistemic force of scalar implicature which differs critically from representations of scalar predication and scalar quantity implicature. Since I will propose a new definition of the orderings that support scalar implicature in Chapter 5, I will continue to use 'ordering' where Gazdar employs 'scale' and to substitute *O* for Gazdar's variables denoting scales except where discussing Gazdar's work on scales in particular.

⁷⁸The problem of how the various meanings licensed by the conjuncts of conjoined sentences are inherited by the conjunction.

4.1. Representing Quantity Implicature

Gazdar describes a method of calculating utterance meanings by defining functions which, for any utterance, will produce sets of entailments, potential presuppositions, and potential implicatures which the speaker of that utterance might license.⁷⁹ The actual meanings an utterance licenses are then calculated by incrementing the current context first with the current utterance's entailments and then with its potential meanings (in a particular but not explicitly motivated order). Meanings that cannot be added without making the context inconsistent are deemed to have been canceled by the context.

To demonstrate his general scheme, Gazdar chooses Horn's notion of scalar predication as one example of the conversational implicatures he will calculate; he terms Horn's phenomenon SCALAR QUANTITY IMPLICATURE. However, in adapting scalar predication for computational purposes, he is forced to make certain simplifying assumptions about the phenomenon. In particular, he implicitly adopts Horn's initial entailment definition of scale (which, as we have seen in Chapter 3, Horn himself recognizes to be inadequate), although he claims to have abandoned any 'semantic' definition of scale in favor of assuming that scales are 'given' [Gazdar 79a:58].

4.1.1. Representing Scale

To support the computation of scalar quantity implicature, Gazdar must formulate a more precise account of scales and values on them. His first task in this enterprise is to formalize Horn's notion of scale. After unsuccessful attempts to define a more satisfactory means of ranking utterances,⁸⁰ Gazdar in effect adopts Horn's entailment definition of scale -- augmented only by the constraint that scales be formed from values drawn from a single domain. So, for $n > 1$, an n -tuple of EXPRESSIONS⁸¹ $\langle e_1, e_2, \dots, e_n \rangle = Sc$ is a quantitative scale for Gazdar iff each member of Sc has the same DOMAIN OF SORTAL APPLICABILITY⁸² as every other member, and e_i, e_j are ranked by entailment.⁸³ Gazdar does not suggest how these scales may be derived or

⁷⁹Gazdar initially termed potential presuppositions *pre-suppositions* and potential implicatures *im-plicatures* to distinguish them from actual presuppositions and implicatures.

⁸⁰See Chapter 5 for a fuller discussion of Gazdar's attempts at ranking.

⁸¹See Section 4.1.2.2.

⁸²Two expressions have the same domain of sortal applicability if they are subject to the same set of selection restrictions [Thomason 72].

⁸³Gazdar curiously defines scales from stronger to weaker, e.g., all/ some.

how they may be identified from utterances. In effect, he is forced to assume that they are just 'given to us'. Below, I will substitute $\langle e_i, e_j, \dots, e_n \rangle = O$ for Gazdar's scales.

4.1.2. Representing Utterances

Gazdar argues convincingly that conversational implicatures should be calculated from an utterance's semantic representation. Reading implicatures from lexical items in surface structure is inappropriate from both a theoretical and a practical point of view, since to do so equates conversational implicature with conventional meaning -- as I discussed in Chapter 2. Particularly when we extend quantity implicature beyond Horn's canonical orderings, this approach is inappropriate for the derivation of quantity implicatures, since values in such orderings are conceptual rather than lexical. In addition, even if it were possible to anticipate all the lexical items that might be employed to license scalar implicature, such anticipation would require redundant listing of synonyms for lexical items referencing a common value. So, the scalar implicature licensed in (73a), that

- (73) A: I think you would have to get it from the instructor for the course...
 B: For which course?
 a. A: Possibly from both courses.
 b. A: Maybe from both courses.

'for all A knows not *certainly from both courses*' might also be licensed via the response in (73b). Reading implicatures from lexical items would involve a distinct analysis for (73a) and (73b), failing to capture the obvious generalization.

Also, Gazdar demonstrates that implicatures cannot just be read directly from the concept or proposition realized by an utterance -- i.e., from the utterance's semantic interpretation. Different utterances may realize the same proposition (have the same truth-conditions) but convey quite different implicatures, as we saw examples of asymmetric *and* in Section 3.1.1.1. Given that conversational implicatures are nondetachable (Condition 5 from Chapter 2), it may not be immediately clear how conversational implicatures that do **not** rely for their interpretation upon *S*'s observance of the Maxim of Manner will present this possibility. Since 'any other way' of saying u_i must also license p_j -- if p_j is a conversational implicature licensed by the saying of u_i -- how can it be, as Gazdar claims, that "many different sentences can express a given proposition and many of these will not contain the scalar item and thus not carry the implicature" [Gazdar 79a:56]? However, it is clear that a general representation of conversational implicature must accommodate conversational implicatures that rely upon the Maxim of Manner for their interpretation. So, Gazdar's general point seems well-taken even if his specific claim about scalar quantity implicature is unclear.

4.1.2.1. Sentences

Gazdar defines a level of representation intermediate between surface form and semantic interpretation, which he terms a SENTENCE, or, "any member of the set of proposition-denoting wffs defined by the formation rules of the language employed for semantic representation." At this level, expressions such as {*perhaps, maybe, possibly*} all referring to a single value in an ordering of epistemic modals (See Section 5.1.2.) will each be represented by the same item. Although Gazdar does not specify a particular semantic representation -- in fact, for illustrative purposes he simply employs English sentences -- he does note that it should be 'surfacey' enough to capture the surface order of conjunction, for example. So, while 'logically equivalent' statements will have the same semantic interpretation, they need not have the same semantic representation.

For the semantic representations of utterances licensing scalar implicatures, I will employ wffs of the representation introduced in Chapter 2. Since the calculation of scalar implicatures does not seem to require access to surface form, these wffs will simply be taken as Gazdar's sentences. If this work were to be extended to other types of generalized conversational implicature, of course, this equivalence would not be appropriate.

4.1.2.2. Expressions

These sentences may be ranked with regard to one another via values associated with 'subparts', or EXPRESSIONS, of sentences, which Gazdar does not further define. Sample expressions are '*possible*' in the sentence '*possible(exists(x)(person(x) and left(x,early)))*', as well as '*exists*', '*exists(x)*', '*person(x)*', and '*early*'. The variables e_i , e_j range over such expressions.

Since Gazdar does not constrain expressions except that they be subparts of sentences, presumably, any substring of a sentence may form an expression -- e.g., '*exists(x)(person*' or even ')'. Clearly, not every substring will correspond to a value -- although every value must be representable by some expression. Gazdar is not concerned with the problem of constraining the set of legal expressions, since, for him, orderings are simply given, apparently as lists of expressions, which, in turn, may simply be identified in semantic representations. However, once we extend scalar quantity implicature to scalar implicature, and allow for rankings among utterances that are not simply given, the problem of identifying the class of expressions which may denote values in an ordering becomes important. Given the representation described in Chapter 2, I will define a subexpression of a sentence p_i as any constant, predicate, logical operator (including the epistemic operators), connective, or quantifier symbol of p_i , or any wff that is a subformula of p_i .

In summary, the approach Gazdar takes to the derivation of implicatures in general and quantity implicatures in particular -- and the approach I will adopt for scalar implicature -- is a syntactic one. By manipulating the semantic representation of (a proposition realized by) an utterance, we can derive the semantic representations of propositions which may be licensed by that utterance -- i.e., the semantic representations of conversational implicatures. The semantic representations p_j of possible scalar implicatures licensed by some utterance u_i with semantic representation p_i , can be calculated by substituting for some expression e_i in p_i values e_j which appear in a common ordering O . p_i and p_j can then be ranked indirectly, via the ranking of their subexpressions in O .

4.1.3. Ranking Utterances

With a representation of scale (ordering) and the concepts of sentence and expression, Gazdar proceeds to describe how utterances may be ranked via the expressions in their semantic representations. To do this, he defines the notion of EXPRESSION ALTERNATIVE as follows:

Sentences p_i and p_j are **expression alternatives** with respect to e_i and e_j iff p_i is identical to p_j except that in ONE place where p_i has e_i , p_j has e_j .

This definition formalizes Horn's and Hamish's notion of utterance comparison. I will adopt the predicate $\text{EXP_ALT}(p_i, p_j, e_i, e_j)$ to denote that p_i and p_j are expression alternatives with respect to e_i and e_j .

Next, Gazdar defines a notion of sentence SIMPLICITY:

A sentence p_i is **simple** with respect to an occurrence of a component expression e_i iff p_i contains no logical functors having wider scope than e_i .

This definition is intended to allow Gazdar to avoid what he claims is a serious flaw in Horn's theory: failure to allow for the effect of logical functors (among which Gazdar includes negation, quantifiers, connectives, and modal operators) with scope over scalar values on predicted implicatures.

Gazdar notes that, by Horn's definition of scalar predication⁸⁴, the utterance of (74a) could implicate that (74b).

- (74)
- a. It is not the case that Paul ate some of the eggs.
 - b. Paul ate all of the eggs.
 - c. Paul ate a few of the eggs.

⁸⁴Actually, one of several. This particular definition is reproduced on page 75.

Gazdar claims Horn fails to recognize that, when such logical functors as ‘not’ have scope over mentioned values like ‘some’, quantity implicatures will not be licensed. That is, utterances such as (74a) (\neg which are not ‘simple’ with respect to some scalar under consideration) should be excluded from among those which may license scalar quantity implicatures. However, while it is clear that (74a) will not license (74b), it is not clear that Horn’s intuitive description of scalar predication would predict that it should. That is, Horn’s claim that the denial of values will set a lower bound for implicatures appears to cover just this point. With this interpretation, the utterance of (74a) should convey (for Horn) that lower values are true, as, (74c).

And Gazdar’s contention that other ‘logical functors’ such as modals, quantifiers, and connectives will block quantity implicatures does not appear well-founded. For example, it does seem that the utterance of (75a) licenses \neg (75b), that (75c) licenses \neg (75d), and

- (75)
- a. It is possible that Paul ate some of the eggs.
 - b. It is possible that Paul ate all of the eggs.
 - c. Paul ate some of the eggs or Paul is a liar.
 - d. Paul ate all of the eggs or Paul is a liar.
 - e. Some people think Paul ate some of the eggs.
 - f. Some people think Paul ate all of the eggs.

that (75e) can license \neg (75f).

So, Gazdar’s ‘set of logical functors’ should be confined to negation alone -- and he should account for scalar quantity implicatures that will be licensed for sentences that are not ‘simple’ as defined above. I will redefine sentence simplicity then as:

A sentence p_i is simple with respect to an occurrence of a component expression e_i iff p_i contains no instances of negation with wider scope than e_i .

Note that p_i may still include the negation operator and be simple with respect to some e_i so long as e_i is not within the scope of this negation.

Assuming the predicate $\text{SIMPLE}(p_i, e_i)$ to stand for ‘ p_i is simple with respect to e_i ’, then $\text{SIMPLE EXPRESSION ALTERNATIVES}$ can be defined as follows:

$$\text{SIMPLE_EXP_ALT}(p_i, p_j, e_i, e_j) \leftrightarrow \text{SIMPLE}(p_i, e_i) \wedge \text{SIMPLE}(p_j, e_j) \wedge \text{EXP_ALT}(p_i, p_j, e_i, e_j)$$

With these definitions, we can specify how expressions may be ranked via values and how they in turn may be used to rank sentences.

- For O a quantitative ordering such that $O = \langle e_i, e_j, \dots, e_n \rangle$ where $n > 1$ (i.e., a linear ordering as for Horn, Harnish, and Gazdar) $\wedge \text{SIMPLE_EXP_ALT}(p_i, p_j, e_i, e_j)$:
 - p_i is lower (or weaker) than p_j with respect to O ;
 - p_j is higher (or stronger) than p_i with respect to O ;

- For O a quantitative ordering such that e_i and e_j are alternate values in O (as described in Section 3.3.2) \wedge $\text{SIMPLE_EXP_ALT}(p_i, p_j, e_i, e_j)$:
 p_i and p_j are alternate sentences with respect to O .

Notions of higher/stronger, lower/weaker, and alternate will for now be understood as described in Section 3.3.2. So, for any quantitative ordering, if some expression precedes another in that ordering, then a sentence containing the first expression will be ranked lower (less informative) than a sentence containing the second, so long as there is no negation in the sentences external to the expressions in question. Similarly, the notion of a sentence being ranked higher, or more informative than another can be explained via the corresponding ranking of component expressions. Lastly, the notion of a value being alternate to another value can define a similar notion of sentence alternates. These definitions are captured in the following predicates:

Higher Sentences:

$$\exists O \text{HIGHER_SENT}(p_i, p_j, O) \leftrightarrow \exists e_i \exists e_j (\text{HIGHER}(e_i, e_j, O) \wedge \text{SIMPLE_EXP_ALT}(p_i, p_j, e_i, e_j))$$

Lower Sentences:

$$\exists O \text{LOWER_SENT}(p_i, p_j, O) \leftrightarrow \exists e_i \exists e_j (\text{LOWER}(e_i, e_j, O) \wedge \text{SIMPLE_EXP_ALT}(p_i, p_j, e_i, e_j))$$

Alternate Sentences:

$$\exists O \text{ALT_SENT}(p_i, p_j, O) \leftrightarrow \exists e_i \exists e_j (\text{ALTERNATE}(e_i, e_j, O) \wedge \text{SIMPLE_EXP_ALT}(p_i, p_j, e_i, e_j))$$

4.1.4. Speaker Commitment

Finally, in view of the characterization of the different types of utterances licensing scalar implicatures which I have made in Section 3.3.3, I will now propose yet another distinction among (declarative) sentences. First, assume that every utterance can be represented as S 's commitment to belief in some proposition or to lack of such belief; this restriction is necessary to represent declaration of ignorance and will be justified on theoretical grounds below in Section 4.2.3. Then

- a sentence p_i represents a DENIAL of a subexpression e_i iff p_i is of the form $\text{BEL}(S, \neg p_j)$ where p_j is simple with respect to e_i ;
- p_i represents an ASSERTION OF IGNORANCE of a subexpression e_i iff p_i is of the form $\neg \text{BEL}(S, p_j)$, and p_j is simple with respect to e_i ; and,
- p_i represents an AFFIRMATION of a subexpression e_i iff p_i is of the form $\text{BEL}(S, p_j)$ and p_j is simple with respect to e_i .

These definitions can be represented as follows:

Denying a Value:

$$\text{DENIAL}(S, e_i, p_i) \leftrightarrow (p_i = \text{BEL}(S, \neg p_j) \wedge \text{SIMPLE}(p_j, e_i))$$

Declaring Ignorance of a Value:

$$\text{IGN}(S, e_i, p_i) \leftrightarrow (p_i = \neg \text{BEL}(S, p_j) \wedge \neg \text{BEL}(S, \neg p_j) \wedge \text{SIMPLE}(p_j, e_i))$$

Affirming a Value:

$$\text{AFFIRM}(S, e_i, p_i) \leftrightarrow (p_i = \text{BEL}(S, p_j) \wedge \text{SIMPLE}(p_j, e_i))$$

4.1.5. Summary

Gazdar's formal treatment of scalar quantity implicature thus provides the basis for a formal account of scalar implicature. An utterance u_i can be identified by its semantic representation p_i . A component expression e_i of p_i can be associated with a value v_i in some ordering \mathcal{O} to permit the ranking of p_i vis a vis other p_j which are expression alternatives to p_i via some e_j associated with a higher, lower, or alternate value in \mathcal{O} . A revised definition of Gazdar's sentence simplicity can be used to specify whether p_i constitutes an affirmation, denial, or assertion of ignorance with respect to e_i .

With these definitions, we are close to a satisfactory representation of scalar implicature. However, the epistemic force of these implicatures must first be determined.

4.2. Epistemic Force

There is no agreement in the literature as to the epistemic force -- the appropriate characterization of speaker knowledge or belief -- which should characterize conversational implicature in general or quantity implicature in particular. For quantity implicature, Harnish in effect ignores the question, while Horn is inconsistent and unconvincing. Neither Gazdar nor his critics, in their simplifying proposals, come up with an intuitively satisfying solution.

4.2.1. Horn's 'Distance from Pole'

Initially, Horn glosses S 's implicit commitment to an implicature p_i licensed via scalar predication as 'for all S knows p_i ' or 'it is consistent with what S knows that p_i ', following Hintikka's explanation of his possibility operator, \mathbf{P} [Hintikka 62]. Part way through the thesis, however, Horn declares that a quantity implicature's epistemic force can be determined by measuring the 'distance' of the value involved from the positive pole on its scale. Horn's formalism runs as follows, where S_a^b denotes the substitution of b for all occurrences of a in S :

... given a quantitative ordering of n elements p_1, p_2, \dots, p_n and a speaker uttering a statement S which contains an element p_i on this scale, then

(i) the listener can infer $\neg S_{p_i}^{p_j}$ for all $p_j, p_i (j \neq i)$

(ii) the listener must infer $\neg S_{p_i}^{p_n}$

(iii) if $p_k > p_j > p_i$, then $\neg S_{p_i}^k \Rightarrow \neg S_{p_i}^j \dots$ [Horn 72:90]

That is, when S affirms p_i , H must infer that S 's highest value p_n is false; s/he may infer that intermediate values between p_i and p_n are false.⁸⁵ So, for Horn, (presumably in the absence of cancelation or blocking), the assertion of p_i forces the implicature that p_n is false but only invites the inference that values higher than p_i but lower than p_n are false. In these terms, then, Horn would claim that the assertion of (76a) would force the inference \neg (76d) but would only invite the inferences \neg (76b) and \neg (76c).

(76)

- a. Some of the people left early.
- b. Many of the people left early.
- c. Most of the people left early.
- d. All of the people left early.

Thus, for Horn, it is the structure of the ordering itself that determines the epistemic force of implicature.

Horn provides little by way of motivation for this claim. Its obvious flaw is essentially the issue discussed in Section 3.1: While generalized implicatures may be associated with particular lexical items or linguistic constructions, these phenomena are not sufficient conditions for conversational implicature. Context and speaker intention play critical roles. If we interpret Horn's account simply as characterizing the meanings speakers may license via conversational implicature, rather than as defining conversational implicature in terms of hearer interpretation, the role of speaker intention in conversational implicature which I view as primary emerges unscathed. However, Horn's account cannot be squared with the equally important contextual dependence of conversational implicature.

In (77a) and (77b), for example,

(77)

- a. A: Are all mushrooms poisonous?
 - b. A: Are many mushrooms poisonous?
- B: Some are.

it seems odd to say that B necessarily implicates $\text{KNOW}(B, \neg \text{all mushrooms are poisonous})$ -- in all contexts -- and may at the same time convey $\neg \text{KNOW}(B, \text{many mushrooms are poisonous})$. If B knows little about mushrooms, s/he will be equally uncertain about whether *many* or *all* mushrooms are poisonous, so the asymmetry of implicature appears odd here. Alternatively, if B knows a great deal about mushrooms, s/he may implicate $\text{KNOW}(B, \neg \text{many mushrooms are poisonous})$ in response to (76b) and $\text{KNOW}(B, \neg \text{all mushrooms are poisonous})$

⁸⁵In fact, as Gazdar notes, Horn fails to specify an additional condition, $i < j$, but this is clearly intended.

in response to (77a) as well. It seems more reasonable to say that the epistemic force of the implicature B licenses is a function of B's domain knowledge, rather than of the structure of the ordering evoked.

In addition, since it is difficult to identify endpoints of orderings in general, assigning epistemic force via distance from positive pole will be impossible in many cases. For example, does (78a) force the inference $\neg(78b)$, or does it instead force the inference $\neg(78c)$ and only invite the inference $\neg(78b)$?

- (78)
- a. The soup is warm.
 - b. The soup is hot.
 - c. The soup is very hot.

Finally, Horn does not explain why, throughout his thesis, he defines implicatures in terms of speaker knowledge rather than speaker belief.

4.2.2. The Epistemic Force of Scalar Quantity Implicatures

Gazdar's view of Grice's Maxim of Quality simplifies that injunction to "Say only that which you know." [Gazdar 79a:45-47] So, for Gazdar, the utterance of p_i by S implicates $\text{KNOW}(S, p_i)$. Rejecting the complexities of a logic of belief, Gazdar argues that, whether or not S has knowledge of p_i -- i.e., whether or not p_i represents S 's true belief -- by asserting p_i , S commits him/herself to knowledge of p_i . The important point for Gazdar is not whether S actually knows that p_i (i.e., that S believes that p_i and that in fact p_i holds) -- but that, by realizing p_i in an utterance, S has committed him/herself to knowing that p_i . In support, he cites Sacks' [Sacks68] argument that the following exchange is odd:

- (79) A: She KNOWS you're crazy.
 B: No, she THINKS I'm crazy. She happens to be right.

That is, speakers will be credited with knowledge whenever propositions they espouse are correct, whether or not they are aware of the truth of their beliefs. However, since I have argued that conversational implicature is defined from a speaker's point of view -- not from the inferences his/her hearer may draw -- this argument is only applicable in the sense that speakers might anticipate such inferences in making their implicatures. Furthermore, the possibility of cancelation indicates that propositions conveyed via conversational implicature should not in fact be accorded the same epistemic status in the discourse as those conveyed via assertion, so argument by analogy from assertion is not compelling. Furthermore, as we will see below, even those (like Gazdar and Soames) who represent implicatures as commitments to knowledge often describe implicatures as commitments to belief. For these reasons, it seems preferable to claim a weaker sort of speaker commitment for conversational implicature in general and scalar implicature in particular, a commitment to belief in and not knowledge of propositions.

Employing Hintikka's [Hintikka 62] epistemic logic, Gazdar identifies scalar quantity implicatures licensed by an utterance with the semantic representation p_i as $K_S(\neg p_i)$, or $KNOW(S, \neg p_i)$ in my notation.⁸⁶ Thus, Gazdar simplifies Horn's two-tiered notion of the epistemic force of quantity implicature.⁸⁷ All scalar quantity implicatures will exhibit the same epistemic force -- speaker commitment to knowledge of the falsity of some proposition. So, for Gazdar, affirming v_i marks higher values v_j as known to be false.

While Horn's distinction between invited and forced implicatures is untenable, Gazdar's simplification seems unwarranted. Contra Gazdar, Soames [Soames 82] argues that it is much too strong to claim that a speaker who affirms a lower value knows higher values to be false. He suggests that higher values are marked as false only when S can be expected to know their truth-value -- otherwise they are marked as unknown. Using Hintikka's P operator,⁸⁸ Soames proposes that sentences p_j referring to higher values v_j may be implicitly marked either as $K_S\neg(p_j)$ or as $P_S(p_j)$ by the utterance of p_i -- depending on H 's belief about S 's knowledge. Note here that it seems likely that Soames is implicitly employing belief rather than knowledge to identify implicatures -- although, for Soames, it is hearer belief that defines the epistemic force of implicatures.

Soames' criticism seems well taken. Given Gazdar's account, in fact, a speaker asserting v_i but ignorant of v_j would have to make that ignorance explicit (e.g., (80a)) lest s/he risk falsely implicating a knowledge s/he lacks (e.g., $\neg(80b)$).

⁸⁶Hintikka's knowledge operator is glossed as *S knows that* p_i .

⁸⁷More precisely, Gazdar defines a function f_s which, given a sentence ψ , returns a set of POTENTIAL scalar quantity implicatures as its value:

$$f_s(\psi) = \{\chi : \chi = K\neg\phi_{\alpha_i}\} \text{ for all } \phi_{\alpha_i} \text{ such that for some quantitative scale } Q, \alpha_i, \alpha_{i+1} \in Q$$

$$(i) \psi = X \phi_{\alpha_{i+1}} Y \text{ where } X \text{ and } Y \text{ are any expressions, possibly null}$$

$$(ii) [\psi] \subseteq [\phi_{\alpha_{i+1}}] \text{ where } \phi_{\alpha_i} \text{ and } \phi_{\alpha_{i+1}} \text{ are simple expression alternatives with respect to } \alpha_i \text{ and } \alpha_{i+1}$$

That is, the utterance of some u_i with the semantic representation ψ scalar quantity implicates that S knows it is not the case that ϕ iff there is some sentence ϕ' which

- is just like ϕ except that it contains a weaker scalar expression
- is entailed by ψ
- is either identical to ψ or forms part of it (e.g., it is a conjunct of ϕ)

so long as the scalar expressions in question are not within the scope of any logical functors in ϕ or ϕ' . (I have corrected some typographical errors in Gazdar's account here.)

⁸⁸Glossed as *for all the speaker knows* or *it is compatible with all the speaker knows that*. So, $P_S(p_i)$ would represent *for all S knows* p_i . Note that, if $P_S(\text{it is raining})$ is consistent with S 's knowledge, $P_S(\neg(\text{it is raining}))$ is also consistent. Or, where $P_S(x)$, $P_S(\neg x)$ must also be true.

- (80)
- a. Some of the people left early but I don't know if they all did.
 - b. All of the people left early.

However, Soames' solution presents its own difficulties: Soames claims that any given implicature will be taken **either** as indicative of *S*'s belief in the falsity of some proposition -- or, as *S*'s lack of knowledge about that proposition's truth or falsity. That is, *S* may implicate $K_S(\neg p_j)$ or $P_S(p_j)$. However, it is not clear that implicatures are this precisely delineated. First, in question-answer exchanges it is reasonable to assume that, in the general case, if a questioner asks a question, s/he has some reason to believe the speaker capable of answering it. But, in such cases, Soames' account would predict that **all** implicatures licensed by responses will license *S* knowledge of falsity, i.e., $KNOW(S, \neg p_j)$. But the same objections raised above against Gazdar must then apply. Second, if each scalar quantity implicature must be represented as either $K_S(\neg p_j)$ or $P_S(p_j)$, then the 'working out' of any implicature (both by *H* and by *S* in anticipation of this) must involve a decision about whether *S* might be expected to know the truth or falsity of p_j or not. While *H*'s belief about what *S* may be expected to know may indeed play an important role in the inferences s/he draws, to require that s/he always make this distinction -- as Soames does -- seems too strong. It often seems enough for a hearer simply to know that a speaker **will not** affirm some value; the source of this failure may be only of marginal interest. Elsewhere, Gazdar [Gazdar 80:7] himself appears (implicitly) to recognize such weaker implicatures as he describes the 'working out' of the implicature licensed via (81b):

- (81)
- a. John is patriotic and quixotic. (= [Gazdar 80]'s 40)
 - b. John is either patriotic or quixotic. (= [Gazdar 80]'s 41)

"So if my addressee hears me reply with (81b), and if he assumes that I am conforming to the maxims which govern conversations, then he can infer *at least that I do not know (81a) to be true, or even, more strongly, that I believe (81a) to be false.*" (my italics) Note here also that Gazdar himself seems to recognize that, whatever the computational difficulties presented by a logic of belief, it is belief and not knowledge that best characterizes these implicatures.

4.2.3. The Epistemic Force of Scalar Implicatures

Although none of the solutions discussed above appears satisfactory, parts of Horn's and Soames' discussions point to the solution I propose for scalar implicature: the use of disjunction to represent quantity implicature. Horn's informal glosses for his examples are presented as implicit epistemic disjunctions, i.e., 'for all *S* knows...' and Soames' claim that *S* may implicate either knowledge of the falsity of some p_j or lack of knowledge of p_j suggests

that scalar implicatures might be seen as epistemic disjunctions. However, for Horn and Soames, speakers may license **either** disjunct of such disjunctions. For scalar implicature, I propose that speakers license the disjunction itself.

Justification for the representation of scalar implicature as disjunction comes from Grice himself. Recall from Section 2.4.2 that Grice [Grice 75] characterizes conversational implicature as "often a disjunction of several possible interpretations" and "often indeterminate." As noted in that section, part of the 'working out' of any conversational implicature involves *S*'s belief that *H* can calculate some p_j which *H* must suppose *S* to believe in order to interpret *S*'s utterance of p_i as cooperative. According to Grice, such calculation often results in multiple possible beliefs q_i , which will be reflected in a disjunctive implicature, $q_1 \vee q_2 \vee \dots \vee q_n$. Consider Grice's:

(82) You are the cream in my coffee.

Here he proposes that, when used ironically, *S* intends *H* to reach first the metaphoric interpretation and then the ironic.

However, in other cases, while it seems reasonable to suppose that *H* identifies a disjunctive interpretation, one is less comfortable imputing the intentional conveyance of this disjunction to *S*. For example, in Grice's interpretation of 36, reproduced in 83, *B* implicates that either

(83) A: I am out of petrol.(= [Grice 75]'s 1)
B: There is a garage round the corner.

B knows the garage in question can supply fuel or *B* does not know whether it can do so. While it seems reasonable that *A* may infer this disjunction, it seems counter-intuitive to say that *B* has implicated it. For, if so, then, given Condition 3 on conversational implicature (*S* belief that *S* and *H* mutually believe that some p_j is 'required' given *S* cooperativity and an utterance u_i), *A* must believe that s/he and *B* mutually believe that 'KNOW(*B*, (garage can supply petrol) \vee \neg KNOW(*B*, \neg (garage can supply petrol)))' is required to make *B*'s saying '*There is a garage round the corner*' consistent with *B*'s observing CP and the maxims. But clearly the disjunction is not 'required' here in any intuitive sense,⁸⁹ since the supposition of either disjunct alone would suffice to permit the interpretation of *B*'s utterance as cooperative. And since *B* obviously knows which of the disjuncts holds -- i.e., the state of his/her own knowledge -- it seems odd to suppose that *B* will implicate a weaker (i.e., a disjunctive) proposition than s/he knows to be the case. In fact, such behavior might well constitute a violation of the Maxim of Quantity.

⁸⁹Recall that this is about the only sense in which we are allowed to understand 'required' -- since it cannot mean logically 'necessary' (See Section 2.4.2.)

But perhaps this oddness/uncooperativity is only apparent. Perhaps S does in fact choose to convey ' $p_i \vee p_j$ ' that is weaker than, although consistent with, his/her actual belief that p_i . S may anticipate that H will be able to disambiguate the implicature by figuring out the appropriate disjunct; at least, s/he can anticipate that H will be able to infer ' $p_i \vee \bar{p}_j$ '. That is, in 83 A might reason: "If I say 'There is a garage round the corner' now, B will believe either that I know the garage is open or that I don't know that the garage is closed. I know that the garage is open. However, whichever disjunct is true should make little difference to B; in either case, s/he should seek petrol at the garage in question, although with more or less confidence in finding it. So, the disjunction will suffice. To disambiguate explicitly would involve unnecessary effort." (Of course, as Grice notes, neither S nor H may recognize S 's communicative intention as such; that is, neither need 'work out' the implicature so literally.) There may be many such cases where there is no need for H to disambiguate disjunctive implicatures: when simply the belief that S cannot falsify some proposition -- or that S cannot affirm some proposition -- is enough. For these reasons, scalar implicatures will be represented as disjunctions -- which may be disambiguated by hearers if necessary according to their world knowledge or beliefs about a speaker.

Further, as noted above, I propose that these disjunctions will be disjunctions with regard to speaker belief about propositions rather than speaker knowledge of them.

So, scalar implicatures licensed by S 's affirmation of some value may be characterized as ' S believes higher p_j are false or S does not know whether higher p_j are true or false'. Such disjunctions may be represented simply by ' $\neg\text{BEL}(S, p_j)$ ', which will be true except when S believes p_j true. In effect, the denial ($\neg T$) of one of three logical possibilities (T,F,#) in a three-valued logic is equivalent to the disjunction of the other two ($F \vee \#$).

The three-valued logic assumed here does pose problems for the actual computation of scalar implicatures, since no theorem provers exist for such a logic. While I will maintain a three-way distinction among speaker beliefs throughout the remainder of the thesis, note that ' $\neg\text{BEL}(S, p_j)$ ' will be true in a two-valued logic just in case ' $\text{BEL}(S, \neg p_j)$ ' and, similarly, ' $\neg\text{BEL}(S, \neg p_j)$ ' will be true just in case ' $\text{BEL}(S, p_j)$ '. So, where logical systems do not permit representation of ignorance -- e.g., where a 'closed world' assumption is made -- scalar implicatures may still be represented as simplified by the assumptions of these systems in the same way that ignorance is simplified by them.

4.3. Representing Scalar Implicature Conventions

With solutions to problems of representing and ranking utterances, with a means of specifying the affirmation, denial, and declaration of ignorance of values, and with the epistemic force of scalar implicature identified, we can now improve upon the description of scalar implicature provided in Section 3.3. In Section 3.3.4, I proposed certain conventions of scalar implicature Imp_{1-3} to satisfy Condition 3 of my definition of conversational implicature. These conventions can now be modified to incorporate the specifications of speaker commitment to utterances and to licensed implicatures discussed above:

$$\begin{aligned}
 \text{Imp}_1: & \quad \exists O (\text{BMB}(S, H, \text{SALIENT}(O, C_h)) \wedge \text{REALIZE}(u_i, \text{AFFIRM}(S, e_i, \\
 & \quad \text{BEL}(S, p_i))) \wedge (\text{HIGHER_SENT}(p_i, p_j, O) \vee \text{ALT_SENT}(p_i, p_j, O))) \\
 & \quad \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, p_j), C_h) \\
 \text{Imp}_2: & \quad \exists O (\text{BMB}(S, H, \text{SALIENT}(O, C_h)) \wedge \text{REALIZE}(u_i, \text{DENIAL}(S, e_i, \text{BEL}(S, \\
 & \quad \neg p_i))) \wedge (\text{LOWER_SENT}(p_i, p_j, O) \vee \text{ALT_SENT}(p_i, p_j, O))) \\
 & \quad \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, \neg p_j), C_h) \\
 \text{Imp}_3: & \quad \exists O (\text{BMB}(S, H, \text{SALIENT}(O, C_h)) \wedge \text{REALIZE}(u_i, \text{IGN}(S, e_i, \neg\text{BEL}(S, \\
 & \quad p_j))) \Rightarrow \\
 & \quad ((\text{LOWER_SENT}(p_i, p_j, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, \neg p_i), C_i)) \\
 & \quad \vee \\
 & \quad (\text{HIGHER_SENT}(p_i, p_j, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, p_i), u_i, \\
 & \quad C_i)) \vee \\
 & \quad (\text{ALT_SENT}(p_i, p_j, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \text{BEL}(S, p_i), C_i)))
 \end{aligned}$$

So, for example, we can now represent that S may license the scalar implicature ‘ $\neg\text{BEL}(S, p_i)$ ’ by affirming his/her belief in p_j when p_i is ranked higher than or alternate to p_j and when the ordering O used to rank these sentences is salient. It remains still to provide a semantics for the notions of ‘higher’, ‘lower’, and ‘alternate’ which I have used to define these rankings, and to provide some understanding of how ‘salience’ will be represented in these definitions.

4.4. Summary

In this chapter I have continued the description of the class of scalar implicatures I introduced in Section 3.3, providing a way to specify utterance ranking and various aspects of speaker commitment to an utterance and to licensed implicatures. In the next chapter I will propose a new means of ranking expressions, which, in turn, will provide definitions for ‘higher’, ‘lower’, and ‘alternate’ sentences.

CHAPTER V

Quantifying Informativeness

Herman: What do you think, Grandpa. Will it be a boy or a girl?

Grandpa: Probably.

The Munsters

In previous chapters, I have shown that scalar implicature, like quantity implicature in general, relies upon *S* and *H*'s perception of the ranking of *S*'s utterance with other utterances s/he might have selected instead. Horn, Harnish, and Gazdar have based this ranking upon a notion of scale or 'strength of claim', which they have basically defined by some form of entailment. However, all three accept that entailment cannot in fact account for all and only those meanings they have identified as quantity implicatures.

In this chapter, I will first identify the orderings that support scalar implicature -- some derived from previous work on quantity implicature, some representing modifications of this work, and some new. Then, I will examine relations that do **not** support scalar implicature. From a comparison of these two classes, I will propose a new characterization of these orderings, as PARTIALLY ORDERED SETS and claim that any poset can support scalar implicature. I will demonstrate how utterances can be ranked via these partial orderings, and how this means of 'quantifying informativeness' accommodates just the class of scalar implicatures. I will then propose a more precise semantics for the notions of HIGHER, LOWER, and ALTERNATE sentences and expressions introduced in Chapters 3.3 and 4 -- based upon a poset condition on scalar implicature.⁹⁰

⁹⁰I note again that, in previous work on scalar implicature, I have termed such posets 'scales', following Horn's and Gazdar's usage. Confusion over intuitive definitions of 'scale' as 'linear ordering' have persuaded me to abandon this usage.

5.1. Orderings Supporting Scalar Implicature

In Section 3.2, I introduced some of the ways in which Horn and Harnish compare utterance informativeness; in Section 3.3 I proposed that hierarchical rankings too support scalar implicature. In this section I discuss these and other orderings, which, I claim, will all permit the licensing of scalar implicature.

A number of the orderings supporting scalar implicature are derived from Horn's 'canonical' entailment-defined orderings, which include the natural language counterparts of the logical quantifiers, modals (epistemic and deontic) and connectives, as well as numerical orderings. Harnish, Kempson, and Prince, have noted that the logical connectives support other implicatures as well. Still other rankings have been identified from definites and indefinites, spatial and temporal orderings, epistemic verbs, verbs of incompletion,⁹¹ and a host of noun pairs and miscellaneous modifiers. To these, I will add that metrics such as set/subset, whole/part, type/ subtype, and entity/ attribute relations, as well as generalization/ specialization hierarchies and instance-of relations also define orderings that support scalar implicature. These orderings include domain-dependent as well as domain-independent, and linear as well as hierarchical rankings. While many might be defined by some notion of entailment, a number cannot.

5.1.1. Quantifier Orderings

Since Aristotle, linguists and philosophers have debated the 'two *some*' problem [Horn 73] -- whether *some* means *some and not all* or *some and perhaps all*. So, in 84, B interprets A's '*some*' as conveying '*not all*'.

- (84) A: Well, some of it you can charge to your grant.
 B: Some?
 A: Oh, all.

But clearly *some* will not always convey *not all*, as in the suspension in 85.

- (85) I just received a letter from X describing their new project at Y, now that Z is finishing up. They plan to build a NL consultation system dealing with private investments. He asked if we could send him a copy of the Harry Gross transcripts. I believe *some (if not all) of them* were in a file at one time. What's the state? Can I get him a copy?

⁹¹Such as 'try to', 'intend to'.

As noted in Section 3.2.1, quantity implicatures based upon Horn's quantifier ordering *some/ all* provide a Gricean solution to this problem. Horn derives this ordering from the natural-language equivalents of the logical quantifiers, '∃' and '∀', which he orders by entailment. While this scale cannot be justified in model-theoretic terms -- $\forall x P(x)$ does not entail $\exists x P(x)$ -- the entailment of *some x* by *all of the x* may be maintained via Russell's contention [Russell 04] that universally quantified statements such as (86a) entail (86b): That is, such statements should be modeled as ' $\exists x P(x) \wedge \forall x P(x)$ ', which, of course, entails ' $\exists x P(x)$ '.

- (86)
- a. All of the unicorns like grapes.
 - b. Some of the unicorns like grapes.

Note that this argument will **not** hold for statements '*All X are Y*', which are standardly represented with simple universal quantification. However, Horn's argument that, in general, the use of *some* can implicate *not all* -- presumably, from his examples, even without the definite reference -- does seem intuitively plausible. That is, statements of the form '*Some X Y*' seem intuitively 'weaker' than statements of the form '*All X Y*', and statements such as (87a) do seem able to convey $\neg \text{BEL}(S, (87b))$.

- (87)
- a. Some unicorns like grapes.
 - b. All unicorns like grapes.

The point here is, however, that while quantifiers do form an intuitive scale, and while some such ordering can be justified in model-theoretic terms, that justification does not extend to all uses of quantification which seem to support quantity implicature.

Other intuitive notions of quantification permit the composition of finer grained orderings by analogy from *some/ all*, orderings such as *few/ some/ many/ most/ all*. So, in 88, A conveys that lower values in such a ranking are true or unknown, e.g., $\neg \text{BEL}(A, \neg(\text{I have some comments}))$.

- (88) A: Well, I don't have very many comments, Ethel.

These values may be lexicalized in a variety of ways, as in 89 and 90:

- (89) A: And you took some courses...
B: I've taken a lot of computer courses in high school.

- (90) A: The heaviest rain for us is over.
B: Not all the rain though?

Additional quantifier orderings can be derived from the base quantifier ordering by quantifying over places, times, persons, or other entities:

somewhere/ everywhere
sometimes/ often/ usually/ always
someone{body}/ everyone{body}

Each of these permits scalar implicatures similar to those discussed above: *S* may assert *sometimes* to implicate \neg *often*, \neg *usually*, and \neg *always* -- or *practically everybody* to implicate \neg *everybody*, as in 91.

- (91) Schuyler and Betty Chapin know practically everybody in the performing arts. Isaac and Vera Stern know *everybody* in the performing arts.
The New Yorker, 20 December 1982

And, *S* may deny higher values, such as *everywhere*, to implicate that lower values, such as *somewhere*, are true or unknown, as in 92.

- (92) Yet gains in agricultural productivity are not welcomed everywhere.
 NYT, 23 February 1983

Here, \neg BEL(NYT, \neg (*gains in agricultural productivity are welcomed somewhere*)) appears to be licensed; the reader assumes the Times **cannot** truthfully deny that productivity gains are welcomed somewhere -- i.e., that either gains are welcomed somewhere or the Times doesn't know whether they are or not.

Harnish's observation that the utterance of (93a) may convey that \neg (93b) relies upon a similar notion of quantifier ordering. He notes that, since (93b)

- (93)
 a. He walked halfway to New York.
 b. He walked all the way to New York.

entails (93a) but not vice versa, (93b) represents a stronger claim than (93a). *S* may therefore assert (93a) to implicate \neg BEL(*S*, (93b)) if s/he believes that s/he and *H* mutually believe s/he is observing the Maxim of Quantity-Quality. Although 'canonical' orderings like the quantifiers might seem difficult to mention **without** licensing the implicatures described above, in fact these meanings need not be conveyed. For example, the meaning Harnish finds licensed by (93a) need not be licensed by 94.

- (94) A: I hear George bet Tom he couldn't walk at least half the distance from Minneapolis to New York.
 B: Yeah, he bet him \$500.
 A: So, what did Tom do?
 B: He walked halfway to New York.

The felicity of such dialogues, in addition to the cancelability of implicatures licensed via mention of quantifiers, confirm that these meanings do represent conversational implicatures.

5.1.2. Modals

Mention of EPISTEMIC and DEONTIC modals may also license scalar implicatures. Lexicalizations of values in an epistemic ordering *possible/certain* support scalar predication in exchanges like 95. Here B licenses the belief that $\neg\text{BEL}(B, i \text{ was (certainly) in the neighbourhood})$.

- (95) A: You were in the neighbourhood of the pantry at one time, were you not?
 B: I may have been.
Cyril Hare, An English Murder

The DEONTIC modals (*permissible/obligatory*) permit the implicature B conveys in 96,

- (96) A: See my comment was, if we should throw even the \$2000 into an IRA or something for her...
 B: You could do that too.

$\neg\text{BEL}(B, \text{we should throw even the } \$2000 \text{ into an IRA...})$. Similarly, in 97, B licenses the implicature that $\neg\text{BEL}(B, \text{we think Gemayal will survive})$.

- (97) A: Do you still think that President Gemayal will survive?
 B: Well, we think that he should.
This Week with David Brinkley, 19 February 84

And the denial of a higher value in this ordering, *should have to*, in 98, conveys that the lesser deontic *can* is true or unknown. That is, $\neg\text{BEL}(B, \neg(A \text{ can take off the back plate}))$.

- (98) A: I would like to know if I can take off the back plate. (= [Hobbs 79a]'s 4)
 B: You shouldn't have to.

Unlike the quantifier orderings, the modal orderings can be justified in model-theoretic terms: In all modal systems, if NECESSARY(p_i) in some model, then POSSIBLE(p_i) in that model; similarly, if OBLIGATORY(p_i), then PERMISSIBLE(p_i). So, in each case, a modal ordering can be defined for the natural-language counterparts of these operators via the logical entailment of their logical counterparts.

5.1.3. Logical Connectives

The natural-language counterparts of the LOGICAL CONNECTIVES, *or* and *and* have figured in several discussions of quantity implicature. Horn bases his *or/and* ordering on the fact that natural-language *or* is often interpreted by hearers as exclusive: So, ' $p_i \text{ or } p_j$ ' may be

uttered to implicate $\neg\text{BEL}(S, (p_i \text{ and } p_j))$.⁹² Since, in standard first-order logic, ' $p_i \wedge p_j$ ' entails ' $p_i \vee p_j$ ' Horn claims that *and* semantically entails *or*. The lower value in his *or/ and* scale, ' p_i or p_j ', may be uttered to implicate that the higher value, ' p_i and p_j ', is false or unknown, as in 99.

- (99) A: Are you serving tea and coffee?
B: Tea or coffee.

And, in 100, A rejects the implicature she infers from B's '*hell or high water*'.

- (100) A: Have a happy Memorial Day Weekend.
B: Thanks. Same to you. I just wish I could enjoy it. It'll be just like the last weekend, and the last, and the last...
A: But I hope not the next, and the next, and the next...
B: No way. Come hell or high water, it will be done.
A: Come hell and/or high water...

Prince [Prince 82a] suggests that this \vee/ \wedge ordering may also account for scalar implicatures arising from the utterance of conditional sentences. She notes that, when a conditional sentence, ' $p_i \Rightarrow p_j$ ', is transformed into its disjunctive equivalent ' $\neg p_i \vee p_j$ ',⁹³ this \vee/ \wedge ordering can account for the implicature $\text{KNOW}(S, p_j \Rightarrow p_i)$ ⁹⁴ as follows: By asserting the equivalent of $\neg p_i \vee p_j$, S may implicate $\text{KNOW}(S, \neg(\neg p_i \wedge p_j))$ -- the falsity of conjunction. This implicature is itself equivalent to $\text{KNOW}(S, p_j \Rightarrow p_i)$: $\neg(\neg p_i \wedge p_j) \Leftrightarrow p_i \vee \neg p_j \Leftrightarrow p_j \Rightarrow p_i$. Thus the fallacy of 'affirming the consequent' can be explained in terms of quantity implicature via the *or/ and* scale. Note also that, since ' $p_j \Rightarrow p_i$ ' is equivalent to ' $\neg p_i \Rightarrow \neg p_j$ ', this manipulation also accommodates the fallacy of 'denying the antecedent'. (And see Section 3.1.1.2.). So, Prince's analysis would predict that the utterance of (101a) may convey (101b).

- (101)
a. If Bill goes then Mary will go.
b. If Mary goes then Bill will go.

However, this suggestion raises an interesting problem. With equal plausibility, one might define an ordering from the logical connectives, $\Rightarrow/ \Leftrightarrow$, since $p_i \Leftrightarrow p_j$ entails $p_i \Rightarrow p_j$. Then the utterance of a conditional $p_i \Rightarrow p_j$ may also convey $\text{KNOW}(S, \neg(p_i \Leftrightarrow p_j))$ via the $\Rightarrow/ \Leftrightarrow$ ordering. However, by this analysis, the utterance of (101a) could convey $\text{KNOW}(S, \neg 102)$ -- and also the meaning

⁹²Note that, while the use of disjunction may be taken as exclusive it need not be; so ' p_i or p_j ' cannot always be translated $(p_i \vee p_j) \wedge \neg(p_i \wedge p_j)$.

⁹³This translation itself is controversial, of course [McCawley 81:49ff].

⁹⁴Prince follows Gazdar's [Gazdar 79b] view of the epistemic force of quantity implicature and represents these implicatures as $K_S \neg(p_i)$.

(102) If Bill goes then Mary will go and if Mary goes then Bill will go.

KNOW(S , $\neg(101b)$). -- where Prince would predict it could convey KNOW(S , (101b)). (That is, if S asserts ' $p_i \Rightarrow p_j$ ' and implicates ' $\neg(p_i \leftrightarrow p_j)$ ', ' $\neg(p_j \Rightarrow p_i)$ ' follows.) Thus, implicature licensed via one ordering contradicts implicature predicted by an alternate analysis.

Cases such as this recall Gazdar's injunction that the semantic representation from which conversational implicatures are derived must be rather more 'surfacey' than a simple first order representation; the substitution of equivalent schemata during the computation of implicatures is risky business indeed and will not be permitted for the calculation of scalar implicatures.

5.1.3.1. Disjunctive Assertions

According to [Kempson 75, McCawley 78], assertion of a disjunction may also implicate that S is not able to assert either disjunct alone, as in 103.

(103) A: Do you think she's ataxic or she's weak?

B: Pause...sigh...I think she's ataxic [...].

That is, if S says ' p_i or p_j ', s/he may implicate $\neg\text{BEL}(S, p_i)$ as well as $\neg\text{BEL}(S, p_j)$. If S can affirm p_i but instead asserts ' p_i or p_j ', s/he will violate the Maxim of Quantity. Note the oddness of 104, in which the humor of B's

(104) A: What do you think, Grandpa. Will it be a boy or a girl?

B: Probably.

response rests upon just such an implicature -- $\neg\text{BEL}(B, \textit{it will be a boy}) \wedge \neg\text{BEL}(B, \textit{it will be a girl})$. For such scalar implicatures, the ordering $p_i \vee p_j / p_i$ must be salient. However, a different ordering must be seen as salient to explain the scalar implicature licensed in 105 -- $\neg\text{BEL}(B, \textit{I want to go out to dinner})$.

(105) A: Do you want to go out to dinner or find a movie to see?

B: A movie sounds good.

Since $\neg p_j$ is cancelable, it does not follow from logical implication, i.e., ' $p_i \vee \neg p_j$ ' cannot be seen as an exclusive disjunction here. Such implicatures are better explained by postulating the evocation of a set/ member relationship among various alternatives (See Section 5.1.10). By affirming a member of that set of disjuncts, as in 105, S may implicate that other members (alternate values) are false or unknown.

5.1.3.2. Conjunctive Assertions

Just as disjunction figures in several types of scalar implicature, so does conjunction. Since the truth of $p_i \wedge p_j$ entails the truth of p_i and of p_j , one might define an ordering p_i / p_i and p_j by analogy from $p_i / p_i \wedge p_j$. Such an ordering accounts for the implicature in 106 that the dog is not vicious.

- (106) A: Uh, a very large and vicious dog is about to attack me.
B: He's large.

Similarly, in Gazdar's example (reproduced in 107), B conveys

- (107) A: Is your mother well and back?
B: Well she's back, yes.
A: She's not well then.

that her mother is not well, as A recognizes explicitly.

5.1.3.2.1. All-Implication

Implicatures licensed via one of Harnish's measures of 'strength of claim' -- ALL-IMPLICATION appear also to rely upon a $p_i / p_i \wedge p_j$ ordering. Implicitly, Harnish relies upon this relationship in claiming that the Maxim of Quantity-Quality solves the classic problem of how the assertion of (108b) conveys (108c). [Harnish 79:126-9,320ff]

- (108) What color is the flag?(= [Harnish 79]'s (22))
a. (The flag is) red and white.(= [Harnish 79]'s (23))
b. (The flag) is red.(= [Harnish 79]'s (24))
c. The flag is all (only?) red.(= [Harnish 79]'s (25))

He explains that, if H believes S is being cooperative, then, in particular H will believe S is obeying Quantity-Quality. (108a) is a stronger claim than (108b), since *red and white* entails *red*, but not conversely. Recall that, for Harnish, the entailment of u_j by u_i is characterized by the fact that ' u_j and $\neg u_i$ ' represents a contradiction. So, in this case 'the flag is red and white and the flag is not red' would, in his view, represent a contradiction. Of course, one might easily argue that 'the flag is red and white and the flag is not all red' does not. Furthermore, Harnish states that u_j will be stronger than u_i if u_j entails u_i and u_i does not entail u_j . Clearly, 'the flag is red' does not entail that 'the flag is red and white' -- by any definition of entailment. By saying the 'weaker' (108b), then, S implicates that $\neg(108a)$ -- and, therefore, that (108c).

In fact, some more general notion of informativeness here -- e.g., why would S mention only one color in the flag if s/he was able to mention others -- appears preferable to entailment. However, Harnish also notes that adjectives such as *spotted*, *dirty*, *stained*, *torn*, *patched*, *dented*, *wet*, *on fire*, for which we might make a similar relative informativeness argument, do

not support similar implicatures.⁹⁵ And others, such as *twisted*, *curved*, and *steep*, in Harnish's terms "convey only weak ALL-IMPLICATION". Harnish seems to suggest that the possibility of implicature appears dependent upon the degree to which the fact that the property mentioned holds for some part *x* of a whole *y* entails that it holds for all of *y*. So, while part of the phenomenon Harnish describes may be subsumed by identifying a relationship between conjunctions and their conjuncts, other aspects remain to be explained.

5.1.3.2.2. Specificity of Assertion

Harnish's notion of SPECIFICITY OF ASSERTION also relies implicitly upon an ordering derived from natural-language conjunction: When *S* asserts *p_i and p_j*, Harnish claims, *s/he* implicates *not (only) p_i and not (only) p_j*. The former will be seen as more specific than the assertion of *p_i* or *p_j* alone -- and, thus, as a weaker utterance. For example, by asserting (109a) *S* may implicate $\neg(\text{only})(109b)$.

(109)

- a. Jones wants ham and eggs for breakfast.
- b. Jones wants ham for breakfast.

But by Harnish's own entailment measure, (109a) represents a stronger statement than (109b) since $p_i \wedge p_j$ entails both *p_i* and *p_j*. The truth of a conjunction $p_i \wedge p_j$ entails the truth of its conjuncts, and conversational implicatures do not contradict entailments. In fact, *S* may implicate $\neg(109a)$ by asserting (109b). It is not clear why 109 should differ from 108, for instance; so, two of Harnish's own metrics would predict different implicatures for the same utterance. So, I will prefer the '*p_i / p_i ∧ p_j*' ordering over Harnish's 'specificity of assertion' notion.

A number of orderings formed from the logical connectives support scalar implicature, including:

p_i or p_j / p_i and p_j
p_i or p_j / p_i and
p_j / p_i and p_j

Since Harnish's all-implication appears consistent with the latter, this notion too will be included in my understanding of scalar implicature. However, his 'specificity of assertion', which seems to contradict this *p_i / p_i and p_j* ordering, will not be included.

⁹⁵Although, 'the flag is wet' does seem to convey that 'the flag is (all) wet' even if it is difficult to imagine what 'the flag is (all) patched' might mean.

5.1.4. Numerical Orderings

5.1.4.1. Cardinal Predicates

It is generally accepted that mention of a cardinal n may be ambiguous between *exactly n*, *at most n*, and *at least n*. Kempson [Kempson 75:154-155] proposes that the *at least n* reading for cardinals can license quantity implicatures. She notes that, when *at least* explicitly modifies a cardinal n in an utterance, S may implicate *more than n*. Consider 110, in which Kempson

(110) I have at least five dollars.

would predict that S implicates that s/he has more than five dollars. However, native speakers informally polled do not seem to share Kempson's intuition. Furthermore, it seems that whatever sense of *more than n* that may be conveyed by such utterances comes instead from *at least*'s conventional force -- i.e., '*exactly or more than n*' -- rather than from conversational implicature. The problem of distinguishing this meaning from conversational implicature by tests such as cancelability results from the disjunctive nature of the meaning; so, (111a) and (111b)

(111)

- a. I have at least five dollars and no more.
- b. I have at least five dollars and in fact more than five.

both are felicitous, since they cancel only part of the conventional force of what *at least* conveys.

Horn, on the other hand, finds that only the '*no more than n*' reading will license conversational implicatures. I claimed in Section 3.2.1 that intonation will not disambiguate among the various readings. However, context can: In response to A's query in (112a), it seems likely that B intends to convey *at least five dollars* by (112e).⁹⁶

(112)

- a. A: Can you afford the movies?
- b. A: Do you have ten dollars?
- c. A: You know, you must declare your foreign currency upon entering Czechoslovakia.
- d. A: Do you have a dollar?
- e. B: I have five dollars.

However, in response to (112b), (112e) appears to license *at most five dollars*. After (112c), (112e) seems more like to be interpreted as *exactly five dollars*. In the simple cases, (112b) and (112d), prior mention of a larger or smaller cardinal appears to favor the *at most n* or *at least n*

⁹⁶It has been pointed out to me that this interpretation may already be out-dated; assume that it costs less than five dollars to see a movie.

reading,⁹⁷ but, in (112c), it seems likely that world knowledge tells us B is more likely to have exactly as much money as s/he declares. Note also that, (112e) might be viewed as either a positive or a negative response to (112d) -- i.e., *Yes, I have (at least) five dollars* or *No, I don't have (at most, or, exactly) one dollar*.

5.1.4.1.1. Approximating Contexts

Horn examines other clues that appear to disambiguate speakers' use of cardinals. When a cardinal is 'rounded' (as in approximating contexts), it is less likely to be taken as indicative of an upper bound (i.e., as *at most n*). So, without a disambiguating context, (112e) is more likely to be interpreted as *at least five dollars* or even as *approximately five dollars* (asserting no bound) than, say, (113a)

- (113)
- a. I have \$5.50.
 - b. I have \$5.57.

would be. Arguably, (113b) would be even more likely to be taken as an exact specification: Since rounded amounts can be substituted for a number of more specific amounts, the assertion of an ostensibly rounded amount like \$5.00 may stand for many amounts, including \$5.57. However, the assertion of an apparently non-rounded amount like \$5.57 cannot be used when, say, \$5.00 is known to be true. So, the assertion of a more specific amount like \$5.57 can be seen as providing more information than the assertion of a rounded amount like \$5.00. Note the correspondence between these observations and Harnish's notion of 'specificity of assertion' above (Section 5.1.3.2.2). However, the latter would predict that, if *S* makes a more general assertion when a more specific assertion would be relevant, s/he may implicate s/he cannot assert a more specific value. So, the assertion of (112e) may convey *S* cannot be more precise about the state of his/her finances -- cannot say (113b). Horn is more concerned with what the assertion of (113b) will convey. Since cardinals taken as asserting an upper bound are those which can license quantity implicatures, it will be easier for *S* to convey and *H* to interpret implicatures arising via the assertion of these clearly 'non-rounded' cardinals.

5.1.4.1.2. Lexicalized Cardinals

Additionally, Horn claims that lexicalized cardinals (such as *annual*, *bicycle*, *double*, and *monologue* which incorporate a cardinal) 'force' an upper-bound interpretation -- and, thus, must license *not more than n* (See Section 4.2.1). So, he predicts that the utterance of (114a) may implicate $\neg(114b)$, while the utterance of (114c) must implicate $\neg(114b)$.

⁹⁷(112e) might also be interpreted as the proffering of an alternate denomination of bill in response to (112d); 'I have four dollars' however avoids this possibility. And see below (Section 5.1.4.1.1) on approximating contexts.

- (114)
- a. This figure has three sides.
 - b. This figure has four sides.
 - c. This figure is a triangle.
 - d. #This figure is a triangle, and, in fact, it is a square.

However, for (114c), this meaning does not seem to be cancelable; i.e., (114d) is infelicitous. So, it seems more likely that (114c) conveys \neg (114b) conventionally (i.e., that \neg (114b) is part of its conventional force) rather than via conversational implicature. That is, *triangle* is defined as a geometric figure with three and only three sides. Thus, contra Horn, lexicalized cardinals do not appear to license the conversational implicatures that the mention of cardinals license.

5.1.4.1.3. 'Reversed' orderings

Horn acknowledges that his entailment-defined cardinal ordering will not accommodate all quantity implicatures that may be licensed by the mention of cardinals. For example, he finds that the implicature licensed by the utterance of 115 relies upon a 'reversed' cardinal ordering.

(115) Arnie is capable of breaking 70 on this {golf} course.

S's assertion of 70 may implicate \neg BEL(*S*, *Arnie is capable of breaking n on this course*) for $n < 70$. So, in such a case, Horn believes the cardinals should be represented as:

80-----70-----60-----50

While such cases might seem to pose yet another problem for Horn's entailment condition, there is a simple solution. In golf, *breaking 70* does entail *breaking n* for all $n > 70$; so semantic entailment may in fact suffice to define the ordering evoked by *S*. However, this ordering is not defined over the cardinals, but rather over events which involve cardinals. 115 shows how a cardinal may figure in orderings other than a simple ordering of cardinals by entailment; again, in such cases, knowledge of the relative salience of metrics will be required to disambiguate possible scalar implicatures.

In sum, the assertion of a cardinal n may license the scalar implicature \neg BEL(*S*, $n+j$) for $j > 0$ when context indicates that a cardinal ordering is salient. In this regard, cardinals that appear to be 'rounded' are more difficult to use to convey scalar implicatures. However, Kempson's notion '(at least) n ' may be used to implicate '*more than n*', and Horn's contentions that mention of lexicalized cardinals license quantity implicatures and that cardinal ordering may be reversed in some cases will not be adopted in this theory.

5.1.4.2. Ordinal Predicates

Horn claims that the canonical ordering of the ordinal scalar predicates resembles his reversed cardinal ordering -- i.e.,

nth----- ...-----*fourth*----- *third*-----*second*----- *first*

So long as an ordinal refers to a ranking other than *number of instances*, he claims, its mention asserts upper-boundedness, and thus conveys that smaller, or, higher, ordinals are known to be false. So, the assertion of *third* conveys \neg *second* and \neg *first* in (116a).

(116)

- a. Little Herbie came in third out of 200 entries.
- b. #Little Herbie came in third and in fact second out of 200 entries.

However, it seems doubtful that (116a) conveys either of these meanings via conversational implicature, since \neg *second*, for example, is not cancelable ((116b)). Too, (116a) conveys \neg *fourth* as well, but this meaning would not be predicted by Horn's analysis.

Horn further contends that *only* reverses an ordinal ordering, as in (117a).

(117)

- a. The Socialist Worker candidate is expected to finish *only* sixth.
- b. #The Socialist Worker candidate is expected to finish *only* sixth and in fact seventh.

But, again, the meaning in question is not cancelable ((117b)) -- and thus cannot be termed a conversational implicature. It appears rather to represent a conventional implicature from *only*.

Horn's exclusion of ordinals referring to 'number of instances' from consideration is ironic: Although such ordinal references do not rely upon a 'reversed ordering', they do appear to support quantity implicatures -- although these implicatures are quite different from those Horn proposes. These ordinal-based implicatures appear to rely critically upon the type of entity modified by the ordinal. When ordinals modify stages of some (linear) process, for example, a corresponding ordering of the ordinals

first-----*second*----- *third*----- ...-----*nth*

appears appropriate, as in (118a).

(118) A: Have you finished the third grade?

- a. B: I've finished the second.
- b. B: I've finished the fourth.

In (118a), B implicates that s/he has **not** finished the third grade; while, in (118b), B conveys that she **has** completed it. However, when ordinals modify entities that are **not** linearly ordered, as in (119a) and (119b),

it seems likely that both responses implicate that B has **not** read the third chapter. Here, a set/

- (119) A: Have you read the third chapter?
 a. B: I read the second.
 b. B: I read the fourth.

member ordering appears salient. Nothing in Horn's theory can explain either of these cases of quantity implicature.

So, I will claim that, although mention of ordinal values may indeed support scalar implicature, Horn's account of implicatures so licensed is incorrect. Implicatures licensed via an ordinal will also depend upon the entity modified by that ordinal: if a linear ordering of entities modified is salient, *S* may license one sort of meaning by choice of ordinal; if a non-linear ordering is salient, then *s/he* may only license another sort of meaning.

5.1.5. Indefiniteness and Definiteness

Grice [Grice 75] and Prince [Prince 81a] note that the use of an indefinite NP often conveys that *S* cannot or sees no need to be more specific, as in (120a). If *H* owns the

- (120)
 a. A bulldozer just flattened a car in the parking lot.
 b. I've been sitting in a car all morning.

flattened car, that fact will be relevant. By failing to identify the vehicle more specifically, *S* implicates that, so far as *s/he* knows, further identification is irrelevant; thus, the car is not *H*'s (and possibly does not belong to any known acquaintance of *H*'s). But obviously this implicature will not be carried by every use of the indefinite. Consider Grice's example (reproduced in (120b)), in which it seems less likely that *S* is conveying this inability to identify *a car* more specifically.

Gazdar has proposed that such implicatures are best seen as scalar quantity implicatures: that is, indefinites and definites can be ordered *a/ the*. Again, Russell's argument that use of a definite description *the X* entails (for *S*) the existence of some such *X* provides a logical justification for this ordering. Since indefinites other than the indefinite article can be uttered to license similar implicatures, as in 121 (where A and B discuss what kind of connections a potential date must have had to be chosen for a television dating program), a more general version of this ordering might be *indefinite/ definite*.⁹⁸

- (121) A: Works on the show.
 B: Some show.

⁹⁸Such responses are most felicitous with stress or with FALL-RISE intonation over the indefinite, i.e., \Some/ show. *B* successfully conveyed that the unappealing candidate must at least . See Note 61.

Obviously, indefinites need not always license scalar implicatures: Generic indefinites for example, like *A whale has lungs*, will not license the implicature $\neg\text{BEL}(S, \text{the whale has lungs})$.

The uniqueness (within some inferable set or context) that definite descriptions convey can also be explained in terms of scalar implicature: When *S* refers to *one of his books* in 122, he

- (122) Well, fortunately, his book is doing so well -- that, well, one of his books is doing so well -- another one is not doing badly either -- and he will be in a position where he can do that. That's really great.

blocks the possible implicature $\neg\text{BEL}(S, X \text{ has written more than one book})$ that *H* might have inferred from the initial definite reference -- *his book*.

While it is interesting to consider this account of the additional meanings which indefinites and definites may convey, it is unclear how such an analysis fits into more sophisticated accounts of this much-studied phenomenon. So, while a theory of scalar implicature must accommodate the observations discussed above, I will not pursue this particular line of inquiry specifically.

5.1.6. Ranked Entities, States, Actions, and Attributes

A number of authors have noted that quantity implicature can also be supported by phenomena which cannot be described more generally than by the observation that certain nouns, verbs, and modifiers -- or, denotations of certain entities, states, actions, and attributes, as I will term them -- appear to be intuitively rankable. The lexical items which figure in this phenomenon have given support to the notion that generalized conversational implicature are defined by the presence of certain lexical items. Many of these rankable items may be used to evoke orderings that support scalar implicature.

5.1.6.1. Rankable Entities

Walker [Walker 75] has noted that the relationship between nouns such as *friend* and *lover* permits generalized conversational implicatures, as when the utterer of (123a) implicates $\neg(123b)$.

- (123)
 a. This is my friend John.
 b. This is my lover John.
 c. This is my friend John, who is also my lover.

Again, not every use of *friend* ((123c), for example) will convey that $\neg\text{lover}$.

Horn also notes that military ranks (*private/ corporal/ sergeant ...*), life stages (*child/ toddler/ infant/ newborn*), and, of course, his *tort/ misdemeanor/ felony/ capital crime* ordering permit the licensing of similar implicatures. Note that, while *lovers* may also be *friends*, they need not be -- and *sergeants* cannot be *privates*; so, although some of these rankings -- which do appear to support scalar implicature -- may be described by an intuitive notion of entailment, even a very intuitive definition cannot distinguish all and only those orderings supporting quantity implicature. These observations provide further evidence that an entailment definition of utterance ranking is inadequate.

Too, many of these orderings may be quite domain-dependent. For example, 124 is impossible to interpret unless we know whether Dr. X is more or less powerful or respected than the chief surgeon -- or is, in fact, s/he.

- (124) A: Does the chief surgeon concur in your diagnosis?
B: Dr. X thinks I'm right.

Except in the final case, concurrence of the chief surgeon will in no way entail the concurrence of Dr. X. If Dr. X is, say, less powerful/ responsible than the chief surgeon, then B may implicate \neg BEL(B, *the chief surgeon concurs in the diagnosis*) in the above exchange. If Dr. X instead ranks higher the chief surgeon, it does not appear that any inference about the chief surgeon can be drawn here.

5.1.6.2. Rankable Activities, States, or Attitudes

Certain verbs may denote activities, states, or attitudes which are themselves rankable. For example, it is clear from 125, 126, and 128 that verbs indicating 'degree of emotional attachment' may also be ordered. In 125, B implicates her lack of commitment to *like* by affirming *don't mind*.⁹⁹

- (125) A: She likes it.
B: I don't mind it.

While no scalar implicatures are licensed during the exchanges presented in 126-128,¹⁰⁰ they

- (126) A: That cat doesn't like anything.
B: She loves being brushed.

- (127) A: Well, are you in love with this guy?
B: Well, I love him.

⁹⁹We might of course analyze B's response as a denial of *mind*, but it seems clear in this exchange from that the more idiomatic notion of *don't mind* as a lesser degree of attachment than *like* is intended.

¹⁰⁰In which higher values are affirmed.

- (128) A: You don't like snow?
B: I LOATHE it.

motivate the claim that speakers consider *loathe*, *like*, *in love*, *love* to be intuitively rankable. Again, if B merely acquiesces in *don't like*, or *in love* s/he might risk A concluding that B cannot affirm yet higher values. I shall return to this notion in Section 7.1.3. Such dialogues provide the best criteria for the identification of these miscellaneous rankings, in particular, those which are domain dependent.

Orderings of other verbs, such as *want* and *need* will also support scalar implicature. From 129, for example, we see that 130 will be felicitous.

- (129) A: O.K. Oh, you want this.
B: I need this.

- (130) A: Do you need this?
B: I want it.

Mention of epistemic verbs, *think*, *believe*, and *know*, can also support scalar implicature, [O'Hair 69, Gazdar 80, Pinkal 83] as shown above in 64 (repeated here), in which B rejects A's implicature $\neg\text{BEL}(B, X \text{ knows } I'm \text{ stupid})$.

- A: Dan thinks I'm stupid.
B: No, he knows you're stupid.

Similarly, in 131 and 132, the ordering *think/ believe/ know* supports the scalar implicature $\neg\text{BEL}(B, \text{believe } \dots)$ and $\neg\text{BEL}(B, \text{know } \dots)$.

- (131) A: O.K. Do you have my records of like what I took?
B: Yes, I think so.

- (132) A: And you have Data Structures?
B: No, I think I'm going to take that next term.

Alternate lexicalizations of items in this ordering are illustrated in the exchanges in 133 and 134. In 133, B implicates

- (133) A: Do you know how many we have?
B: We can check.

$\neg\text{BEL}(B, I \text{ know how many we have})$ by affirming *can check*, while in 134 B conveys the same implicature by affirming *find out*.

- (134) A: Do you know what time it is?
B: I can find out.

Such use of the epistemics has been discussed in the literature as HEDGING behavior. [Lakoff, G 72, Prince 82b] Hedges other than the epistemics may also license scalar implicatures: In 135 and 136, from a fictional interrogation of witnesses by police, speakers

attribute propositions to others to implicate their own inability to commit themselves to those propositions.

(135) A: Nobody saw anything or anyone. Mrs Praed discovered the body at about six o'clock in the morning--

B: She says.

(136) A: Mr Porle, I believe you were employed by a Miss Marietta Rainer. Is that correct?

B: I used to clean her windows--yes.

A: You know she was recently found dead?

B: So someone told me--yes.

[Prince 82b] terms these uses of performatives and epistemics *shields*, and explains that they introduce fuzziness in the relationship between propositional content and speaker.

[Prince 82b]'s *approximators* also represent hedges whose use may support scalar implicature: Items so identified will introduce "fuzziness within the propositional context" as in [Prince 82b]'s 137:

(137) A: Are his feet blue?

B: They're kind of blue.

B implicates here that $\neg\text{BEL}(B, \textit{his feet are blue})$ by use of what [Prince 82b] terms an ADAPTOR -- a term which adapts an old term to a new instance -- here, *kind of*. Other approximators called *rounders* provide a more general term when the precise term is not available or is unnecessary. So, in [Prince 82b]'s 138,

(138) Her temperature was about 100.

S implicates that $\neg\text{BEL}(S, \textit{her temperature was 100})$.

It has not been noticed in the literature that similar orderings may be evoked by the mention of certain verb/preposition pairs. In 139, for example,

(139) A: Should we have the ham for Christmas?

B: We could have it during Christmas.

B implicates $\neg\text{BEL}(B, \textit{we should have the ham for Christmas})$ by affirming the lesser value *have it during*; that is, the ham should not be the main course for Christmas dinner, but it could be eaten during the Christmas season. Note that this distinction relies also upon the recognition that the referent of *Christmas* in A's question, *Christmas dinner* is distinct from the *Christmas season* to which B refers.

And in the exchange presented in 140, in which a (somewhat unusual) caller describes her

(140) B: I as a result of my extensive investigations and especially on an international basis -- I was involved in industrial and governmental espionage so I had to be very careful about --

A: CIA stuff?
 B: Pardon me?
 A: CIA stuff?
 B: We HAVE worked WITH them.
 A: I see.
 B: I don't work FOR them.
 A: I see.
 B: Uh I work FOR my corporation.
 A: Uh huh.
 B: I have also worked with security organizations, the FBI --
 A: Uh huh.
 B: And the intelligence groups. But what happened was I had to remain silent about my inheritance because uh we were threatened.

past work to a financial advisor, the caller (B) first implicates \neg BEL(B, *we worked for the CIA*) and then reinforces this implicature explicitly. In this exchange, the scalar implicature is licensed by a ranking of *work with* as a lesser item than *work for*; this ordering is plausible if one considers that *working for* someone involves *working with* them, but not vice versa.

5.1.6.3. Rankable Attributes

Horn in particular has noted a number of modifiers which support quantity implicature; some are presented below:

<i>pretty/ beautiful</i>	<i>happy/ ecstatic</i>
<i>warm/ hot</i>	<i>cool/ cold</i>
<i>intelligent/ brilliant</i>	<i>fair/ good/ excellent</i>
<i>middle-aged/ old</i>	<i>adolescent/ adult</i>
<i>critically/ mortally/ fatally (wounded)</i>	

These and numerous other modifier orderings also support scalar implicature.

For example, Horn finds that the affirmation of *pretty* implicates the inappropriateness of every stronger term, which, for him, is every term that entails *pretty* in an ordering defined by *degrees of attractiveness*.¹⁰¹ Thus, the utterance of (141a) conveys that \neg BEL(S, (141b)).

- (141)
- a. That picture is pretty.
 - b. That picture is beautiful.

Scalar implicatures are also licensed in 142, in which a speaker points to a standard deviation that looks rather large and explains:

The audience understood that S was conceding \neg BEL(S, *that's an X amount*), for $X < \textit{tiny}$ in

¹⁰¹Horn does not claim that all these attributes can be ordered by entailment, and even *pretty* is suspect in this regard.

(142) That's not a tiny amount, but...

some ..., *large, medium, small, tiny* ordering. Note here, as in other modifier orderings such as the temperature scale or *degrees of attractiveness*, that the orientation of such rankings varies with context (See Section 6.3.2.3).

Orderings formed from intuitively rankable entities, states and activities, and attributes -- licensed via mention of the NP's, VP's, adjectivals and adverbials that refer to them -- thus support scalar implicature. Of all the orderings discussed previously in the literature, this class has been the most difficult to accommodate in a single measurement of informativeness. Such rankings are also inherently difficult to identify, since they may be quite domain-dependent.

5.1.7. Temporal Implicatures

Harnish and Horn have both identified numerous temporal indicators that may license quantity implicatures, including tense and temporal adverbials such as Horn's *sometimes/always* ordering. Harnish notes that use of the past tense in (143a) seems to convey that "the activity or state indicated by the verb no longer is present". [Harnish 79:388] So (143a) can convey \neg (143b).

- (143)
 a. I x'd (to) p_i .
 b. I now p_i .

In particular, he points out that verbs such as *used to*, *wanted to*, and *believed* may license such implicatures. For example, in 144, *S* implicates that \neg BEL(*S, chocolate cigarettes now come in a box ...*).

- (144) Chocolate cigarettes remind me of my childhood. They used to come in a box that looked like a fake cigarette box for kids. I like the idea of using them, but I would feel silly using this kind of chocolate cigarette.

Harnish also notes that use of the simple past tense for any verb will license the quantity implicature that the queried state does not now hold; so, *I p_i'd* may convey \neg BEL(*S, I now p_i*). For example, in 145 B

- (145) A: Are you on your honeymoon?
 B: Well, I was.

conveyed an indirect negation to the question asked by this means -- implicating \neg BEL(*B, I am now on my honeymoon*).

Similarly, Horn proposes that what he terms 'quasi modals' (and what others have called

'verbs of incompleteness') also license quantity implicatures, via orderings such as *want to/ try to/ succeed in*. For example, in 146

- (146) A: Is she going to get that card renewed?
B: She's trying.

B implicates that her friend has not yet -- and is not sure of eventually succeeding in -- getting a free membership card to a night club renewed. Similarly, in 147, A implicates \neg BEL(A, *I am having V-8 juice*).

- (147) A: Do you want V-8 juice?
B: Are you having V-8 juice?
A: I'm considering it.

And in 148, B implicates that he has *not* actually bought the television set.

- (148) A: Did you buy the TV?
B: I'm going to.

Horn omits verbs such as *able*, *persuade*, *forget*, and *intend* from his scalar predicates, since he claims that these predicates entail either their complements or the negation of their complements and thus cannot be used to implicate that negation. However, *able* and *intend* do indeed appear to support quantity implicatures similar to those supported by his 'quasi-modals'. *Be able to p_i* is not synonymous with *manage to p_i* as Horn claims; possessing the ability to achieve p_i need not entail *achieving p_i* . So, the utterance of 149 need not convey that B has indeed made dinner; in fact, B conveys just the opposite in this exchange.

- (149) A: Have you made dinner yet?
B: I can.

Neither does *intending to do p_i* entail either *doing p_i* or \neg *doing p_i* . In 150, B conveys that the investment in question had been meant as deferred income, even though he questions whether it will count as such.

- (150) A: I would suspect that this was part of a pension plan, was it?
B: Not as a pension plan per se. It was intended to be deferred income.

That is, the assertion of *intended to p_i* may convey \neg BEL(*S*, p_i), and this meaning does represent a conversational implicature in the terms presented in Chapter 2.¹⁰² Both *able* and *intend* do in fact support scalar implicature.

Temporal implicatures may also be licensed about present states by the affirmation of

¹⁰²Note in particular that the meaning *intend to p_i* may license is cancelable, as in vi:

(vi) I intended to p_i and in fact I did.

some future state. For example, B's response in 151 licenses the implicature $\neg\text{BEL}(\text{B}, I \text{ have a master's})$.

- (151) A: You have a master's don't you?
B: I will.

In 152, B licenses $\neg\text{BEL}(\text{B}, \textit{that is the best way to go (in the future)})$.

- (152) A: Now is that the best way to go for me?
B: Presently yes.

As I noted in Section 3.2.2, Harnish claims that denying p_i for t_j may license the belief that p_i held for some period t_i , $i < j$. But similar denials may also license implicatures about future time periods: In 153, for example, B's denial of *today* conveys that B will leave at another -- presumably future -- time.

- (153) A: Are you leaving today?
B: I'm not leaving today.

In short, affirming or denying a temporal may license implicatures about past or future states.

If, like Horn, one sees temporal orderings as both linear and defined by entailment, then one cannot explain such two-sided implicatures: Both affirmations and denials of some temporal value v_j may license scalar implicatures involving both higher v_i and lower v_k : The affirmation of v_i may license $\neg\text{BEL}(S, \neg v_i)$ as well as $\neg\text{BEL}(S, \neg v_k)$; the denial of v_j may license $\neg\text{BEL}(S, v_j)$ as well as $\neg\text{BEL}(S, v_k)$. So, these orderings are clearly not definable by entailment. The truth of future p_i entails nothing about the truth of present or past p_i .

However, from the implicatures licensed in naturally occurring discourse, it appears that temporal orderings may best be seen as sets of temporal alternatives, any number of which may, in general, hold. In 154, for example, B in fact responded with (154a). However, she might felicitously have given any of the other responses:

- (154) A: Do you have a badminton team?
a. B: I had.
b. B: I do now.
c. B: I will have.
d. B: I don't have one now.
e. B: I won't have.
f. B: I didn't have one.

Where (154a) indicated to A $\neg\text{BEL}(\text{B}, S \text{ does (now) have a badminton team})$ and $\neg\text{BEL}(\text{B}, S \text{ will have a badminton team})$, (154b) might convey S's inability to commit herself to past and future, and (154c) a similar lack of commitment to present and past. Alternatively, the utterance of (154d) could license $\neg\text{BEL}(\text{B}, \neg I \text{ had one})$ and $\neg\text{BEL}(\text{B}, \neg I \text{ will have one})$; of (154e), lack

of commitment to the falsity of past and present; and (154f), lack of commitment to the falsity of present and future -- e.g., *I didn't have one but I do now* or *I didn't have one but I will* might each be implicated by (154f).

Implicatures licensed by mention of such alternatives are identical to those observable when a set/ member relationship is evoked. That is, the affirmation of any member may convey *S*'s lack of commitment to other unmentioned members, while the denial of a member may convey that *S* believes other members true or does not know whether they are true or false (See 5.1.10 below.). While our intuitive notion of time may be linear, then, the implicatures that may be licensed by temporal references suggest a different representation here.

5.1.8. Spatial Orderings

Spatial orderings may also support scalar implicature. The directionality of spatial orderings appears to follow the egocentricity observable in most deictic behavior, in which the unmarked point is the deictic center: Just as the central person is the speaker and the central time the time of utterance, so the central place is the speaker's location at the time of utterance. [Levinson 83] So, implicatures licensed via spatial orderings will be calculated with respect to *S*'s location unless otherwise indicated. Below I will assume that deictic center and speaker location are identical for simplicity's sake.

In 155, for example, B appears to affirm a nearer location to deny a queried further location.

- (155) A: Did you get downtown?
B: I went to Bonwit's.

Similarly, in (156a), a bus-driver affirms one street to deny that

- (156) A: Do you go straight up Walnut?
a. B: To Thirty-fourth.
b. B: I don't go to Thirty-fifth.

his bus will go to the end of Walnut, Sixty-Third Street. But B might have conveyed the same information by (156b), denying a further location to implicate the affirmation of a nearer location.¹⁰³ Similarly, in 157, B affirms a nearer location to deny the farther *way to* location.

- (157) A: Is this the way to the Towne Building?
B: This IS the Towne Building.

¹⁰³It seems less likely that B would choose a denial in this situation, although the denial is perfectly interpretable. However, 'I turn before Thirty-fifth' seems somewhat more natural an alternative.

So, S may affirm a nearer location l_i to convey lack of commitment to a farther one l_j (i.e., $\neg\text{BEL}(S, l_j)$) or deny a farther l_j to convey $\neg\text{BEL}(S, \neg(l_i))$.

It also would seem that a speaker may declare ignorance of some location l_j to convey $\neg\text{BEL}(S, \neg(l_i))$ for l_i closer to S than l_j and $\neg\text{BEL}(S, l_k)$ for l_k further from S . So, B might convey that $\neg\text{BEL}(B, \textit{it gets to Thirty-sixth})$ and $\neg\text{BEL}(B, \neg(\textit{it gets to Thirty-fourth}))$ by the response in 158.

- (158) A: Does this bus go up Walnut?
B: I don't know if it gets to Thirty-fifth Street.

5.1.9. Process Stages and Prerequisites

Harnish implicitly recognizes the notion that process or prerequisite orderings may permit scalar implicature in his discussion of how the assertion x finished y may be viewed as a stronger remark than the assertion x started y . Since *finishing* 'entails' *starting*,¹⁰⁴ -- but not vice versa -- the assertion of x started y implicates the falsity of x finished y , as when S implicates $\neg(159b)$ by saying (159a).

- (159)
a. Minnie started mowing the lawn.
b. Minnie finished mowing the lawn.

This intuition seems correct, even though Harnish's explanation is unconvincing.¹⁰⁵ And the denial of *finish* can be employed to implicate $\neg\text{BEL}(S, \neg\textit{start})$ -- that *finish* is the earliest stage in some process S can truthfully deny. As far as S knows, earlier stages like *starting* are true.

¹⁰⁴In the sense that *having finished* 'entails' *previously having started*. This is one example of the disparity between Harnish's abstract characterization of entailment and the intuitive -- and, here, temporally-dependent -- notion he is trying to capture.

¹⁰⁵According to Harnish, since *finishing* entails *starting*, x finishing y is equivalent to (viiia). The denial of (viiia) is (viiib).

- (vii)
(a) $(x \text{ started } y) \wedge (x \text{ finished } y)$
(b) $\neg(x \text{ started } y) \vee \neg(x \text{ finished } y)$
(c) $\neg(x \text{ started } y)$
(d) $\neg(x \text{ finished } y)$

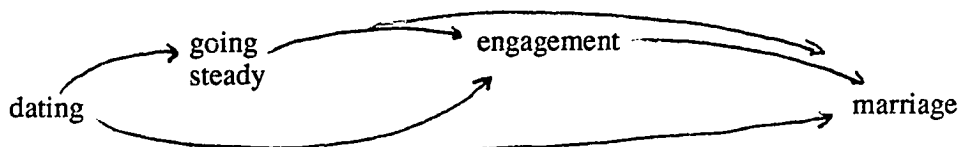
Clearly the truth of the first disjunct of (viiib) ((viiic)) is sufficient for the truth of the disjunction. So S might deny (viiia) simply by affirming (viiic). By affirming (what is, in effect) the disjunction (viiib) instead, S thus makes a weaker statement than would be relevant and suitable if s/he could truthfully affirm (viiic). So it must be that (viiic) is false, i.e., that x started y is true. Of course, the truth of the second disjunct of (viiib) ((viiid)) is also sufficient for the truth of (viiib). So, by the same reasoning we might conclude that S is unable to affirm (viiid) and that x finished y is true! The problem, of course, is that Harnish defines *finish* in terms of itself (i.e., x finishing y is equivalent to (viiia)) and assumes an implicit ordering of conjuncts which his notation does not support.

Note that, in exchanges such as 160, B provides an indirect response to A's query, which we might

(160) A: Did you finish this?
B: I didn't start it.

interpret as an attempt to block the implicature that could be licensed by a simple denial of *finish* -- i.e., that lower values such as *start* are true or unknown to B.

Orderings such as these may be seen as *stages of a process* or *prerequisite* orderings and support scalar implicature. For example, assume the following ordering:



Then we can explain the following implicatures as affirmations of stages in this process: In 161, B implicates that the woman in question is

(161) A: So, is she married?
B: She's engaged.

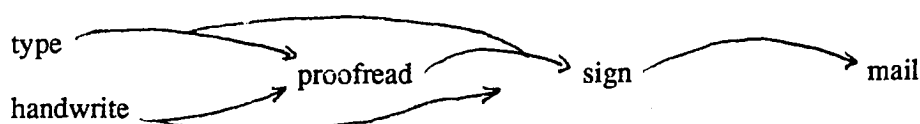
not married by affirming that she is engaged. Note that this response will not commit B to the truth of *going steady*, for example, although this state may sometimes precede engagement: So, process orderings need not be linear.

But note that expressions which may be seen as denoting process stages need not actually serve this function. In some contexts, for example, *taking the GRE's*, *writing a thesis*, *doing a project*, *taking a comprehensive exam* *taking prerequisites* and *taking electives* might be modeled as stages in a process of completing a Computer Science major. But it seems clear that, in an exchange like 162, these expressions are better seen

(162) A: O.K. And for Barnard students, they had to take either GRE or write a thesis, right? But for Computer Science I don't know what to do. Is there any project or...?
B: No, no, not. Our Department doesn't require any project neither a comprehensive exam, so all you need to do is fulfill the requirements which are a couple of prerequisites and four electives.

as an unordered set of prerequisites, rather than as stages in a temporally ordered process.

When orderings like these do include alternative or optional paths, such branching nodes may be seen, like hierarchical siblings, as alternate values in the ordering. For example, signing a letter may be preceded optionally by proofreading it, and also by the alternate stages of typing the letter or writing it out by hand, as represented below:



Then, in an exchange like (163a), the affirmation of an alternative stage *typing a letter* licenses the

- (163) A: Have you mailed that letter yet?
 a. B: I've typed it.
 b. B: I haven't proofread it.

implicature $\neg\text{BEL}(\text{B}, I \text{ mailed the letter})$ about the queried higher stage and $\neg\text{BEL}(\text{B}, \neg(I \text{ wrote the letter by hand}))$ about the unqueried alternate value. Note, however, that B's utterance of (163b) can implicate only $\neg\text{BEL}(\text{B}, \neg(\text{type} \vee \text{handwrite}))$ -- not one particular disjunct.

So, not only do linear orderings, such as those discussed by Horn and Harnish, support scalar implicature, but non-linear orderings do as well. Process stages may be ordered by either type of metric.

5.1.10. Sets and Whole/ Part Relationships

Other non-linear orderings that support scalar implicature, which have not been discussed previously in the literature, include set/ proper subset relationships¹⁰⁶ and part/ whole relationships. I have already noted some naturally occurring data in support of this claim (See examples 59 and 71.). Further examples of scalar implicature licensed via evocation of sets are presented below.

- (164) A: Can you sing a Motels song right now?
 B: Now?
 A: Yeah.
 B: My cousin can.
- (165) A: Do you have apple juice?
 B: I have grape or tomato or bloody mary mix.
- (166) A: Have you ever knitted before?
 B: I've done a lot of crocheting.
- (167) A: I'll have a small lentil soup, and whole wheat bread.
 B: We have rye, pumpernickel, and rolls.

B's implicature in 164 that she is **not** able to sing a Motels song is also based upon affirmation of a subset of a perceived set of relatives. Similarly, in 165 B evokes the set of juices to

¹⁰⁶Defined over the non-null subsets of some salient set.

implicate $\neg\text{BEL}(B, I \text{ have apple juice})$; by committing herself to the set $\{\text{grape}, \text{tomato}, \text{bloody mary mix}\}$, B does not commit herself to the set $\{\text{grape}, \text{tomato}, \text{bloody mary mix}, \text{apple}\}$, and thus, does not commit herself to *apple juice*. The exchanges in 166 and 167 involve similar implicatures. The set/ subset orderings that may permit scalar implicature are limited only by the imagination of conversational partners: For example, in 168, B's reply indicates that she has

- (168) A: You should be able to log on to the VAX -- so long as you've had your granola.
B: I had grapefruit.

decided that the set of breakfast foods must be salient -- although this speaker later confessed to having been quite puzzled by A's remark.

The claim that speakers may implicate lack of commitment even to unmentioned members of some salient set is supported by the following exchanges: In 169, A clearly understands that B has implicated that no one other than her son and daughter live with her.

- (169) A: Isn't there anybody waiting for you at home?
B: Yes there is...my son, my daughter.
A: No one else?

And, in 170, *S* blocks the potential inference that his

- (170) Fine? Alright, get hold of a copy of the July 7 81 WSJ, next to the last page. You can get it down at the Logan Square Library or at any one of the major colleges in the area and, uh, Drexel, Temple, Villanova, LaSalle, St. Joe's, uh, Penn State [sic], Lidener [sic] -- I hope I didn't miss anybody. If I did it was unintentional...

list has exhausted the major colleges in the area. In sum, such exchanges provide evidence that, when *S* affirms a proper subset of some mentioned or inferable set *V*, s/he may implicate that s/he is unable to commit him/herself to unaffirmed members of *V*; in effect, by affirming a proper subset v_i , *S* may license $\neg\text{BEL}(S, v_j)$ where $v_i \neq v_j$ and v_i and v_j are non-null and are in the power set of *V*.

A more complex example of how speakers may license scalar implicature by evoking set-inclusion relationships is presented in 171:

- (171) 1) A: Let me just check whether you have all the prerequisites.
You have Calc 1 and 2?
2) B: Uh-hmm.
3) A: You have Introduction to Programming?
4) B: Oh, yeah.
5) A: You have Data Structures, Fundamental Algorithms.
6) B: No, I don't have Fundamentals.
7) A: O.K.
8) B: Or Comput[ability].

- 9) A: O.K., so you'll need Fundamentals, Comput[ability]. You have...
 10) B: I'll have to take Finite Math, too.
 11) A: O.K. You have Discrete Math?
 12) B: Yeah, I have it and I have Assembly Language.
 13) A: O.K., and do you have any electives?
 14) B: Yeah, I have one -- Intermediate.
 15) A: O.K., so you need three more electives, and you need Comput[ability], Discrete...Finite Math?
 16) B: Finite Math and Fundamentals.

In lines (3) and (5) of this example, A, a faculty adviser lists prerequisite courses for a computer science major for B, a student. In (6), by denying a singleton subset of the set {Data Structures, Fundamental Algorithms} listed in (5), B implicates that s/he has taken the undenied member, *Data Structures* -- that is $\neg\text{BEL}(B, \neg\text{Data Structures})$. In (10), B corrects A's implicature from (9) that the set {Fundamentals, Computability} will exhaust the set of required courses B has yet to take; implicitly, A has conveyed $\neg\text{BEL}(B, \text{Finite Math})$ -- that, as far as A knows, B will not have to take Finite Math -- so B corrects that implicature. In (12), B adds that s/he has taken *Assembly Language* when questioned about *Discrete Math*. A simple response of 'yes' to (11), s/he may believe, could encourage A to conclude that $\neg\text{BEL}(B, \text{I have Assembly Language})$. So, B blocks that potential inference in (12). Finally, B corrects the apparent implicature licensed by A in (15), that A has mentioned all the courses B has left to take -- hence $\neg\text{BEL}(B, \text{courses other than those mentioned in (15)})$ -- by adding that s/he must also take *Fundamentals*.

The implicatures speakers license by the enumeration of part(s) of a whole are similar to those licensed by the mention of a proper subset -- conveniently enough, since particular examples may often be interpreted in either framework. For example, in 172, B implicates he has not read the other half of the book;

- (172) A: Did you finish the book?
 B: Half of it I read.

in 173, B corrects A's implicature that $\neg\text{BEL}(A, \text{there is no air-conditioning in other rooms of the house})$.

- (173) A: I've been told there's no air-conditioning in my bedroom.
 B: There's no air-conditioning anywhere in the house.

In each case, these exchanges might just as easily be analyzed in terms of set/subset relations. However, since the implicatures licensed by asserting a proper subset are identical to those licensed by asserting part(s) of a whole -- i.e., both represent lower values in a salient ordering -- the choice is unimportant for this analysis.

5.1.11. Type/ Subtype, Instance-of, and Generalization/ Specialization Relationships

Scalar implicatures may also be licensed by reference to items related by a type/ subtype, instance-of, or generalization/ specialization relationship. So, in 174, B implicates $\neg\text{BEL}(B, I \text{ have Italian})$ by affirming the alternate value *vinaigrette* in an ordering of *types of salad dressings*.

- (174) A: What kind of salad dressing do you want?
 B: Do you have Italian?
 A: Vinaigrette.

In 175, B conveys that she does have a form of *paste* by affirming the subtype, *rubber cement*.

- (175) A: You don't have paste, do you?
 B: Rubber cement.

Note here that, since the truth of a subtype v_j entails the truth of its type v_i , v_j will represent a higher value than v_i and sibling subtypes v_j and v_k will represent alternate values for the purpose of calculating scalar implicatures. So, *S* may affirm a type to convey lack of commitment to a subtype, may deny a subtype to affirm a type, or may affirm or deny one or more subtypes to convey the opposite commitment to alternate subtypes.

Even MARGINAL or ATYPICAL subtypes¹⁰⁷ may license scalar implicatures. In 176, for example, B's affirmation of *vitamins* conveys that his child is not taking

- (176) A: Is she taking any medication?
 B: Vitamins.

any other medications, although *vitamins* itself is not clearly a type of medicine. And B's response in 177 conveys that she does not have

- (177) A: Do you have a pet?
 B: We have a turtle.

other pets. Because the affirmation of a marginal subtype can be employed to implicate $\neg\text{BEL}(S, \text{other subtypes})$, responses such as those in 176 and 177, as in 178 below, may be interpreted as either *yes* or *no*, depending upon whether *H* believes vitamins constitute medication, turtle constitutes a type of pet, or allergies constitute sickness.

- (178) A: Are you sick?
 B: I've got allergies.

S's implicature will, however, be independent of *H*'s understanding that *S* has given a positive response or a negative one.

¹⁰⁷Items whose membership in the type v_i is debatable, as, penguins are generally classified by subjects as marginal or atypical members of the category of birds. See [Rosch 75, Jolicoeur 84].

Scalar implicatures licensed via type/ subtype relationships are similar to those that may be licensed by evocation of instance-of or generalization/ specialization relationships. For example, in 179, B may implicate \neg BEL(B, *that will be Visa*)

- (179) A: Will that be Visa?
B: It'll be a charge.

by affirming *charge*. And, in 180, where the relationship between *outside* and *car* clearly cannot

- (180) A: Is he outside?
B: He's in the car.

be seen as one of type/ subtype, the fact that B may convey commitment to *outside* by affirming *in the car*, suggests that B may implicate \neg BEL(B, *he's in the car*) by affirming *outside*, as in 181, or convey

- (181) A: Is he in the car?
B: He's outside.

\neg BEL(B, \neg *he's outside*) by denying *in the car* in 182 based upon a generalization/ specialization ordering.

- (182) A: Is he outside?
B: He's not in the car.

Like implicatures licensed via evocation of set/ subset and part/ whole relationships, scalar implicatures licensed by mention of members of type/ subtype, instance-of, and generalization/ specialization relationships are difficult -- but, for our purposes, unnecessary -- to tell apart.

5.1.12. Entity/ Attribute Orderings

Finally, scalar implicatures may also be licensed by evocation of an entity/ attribute relationship, as in 183, 184, 185, and 178. In 183, for example, B implicates that he cannot commit himself to the queried proposition *there is an opera on* by affirming

- (183) A: Is there an opera on?
B: There's a recorded one.

an attribute of an opera presentation (*recorded*). B is not sure a recorded opera constitutes having an opera on, so says 'as much as he truthfully can' by affirming an attribute. Likewise, in 184, B

- (184) A: Are you a doctor?
B: I have a PhD.

affirms an attribute of some *doctors* to implicate her inability to commit herself to *being a doctor*.

Speakers may also license implicatures about their commitment to other attributes by affirming or denying some attribute of an entity. In 185, B implicates $\neg\text{BEL}(B, I \text{ am Greek})$ by affirming an attribute of Greek people (ability to speak the language). By affirming an attribute, in effect,

- (185) A: Are you Greek?
B: I speak some Greek.

a speaker may convey that s/he is unable to affirm the entity itself, i.e., in 185, Greek nationality.

If a mentioned attribute is a definitional characteristic of a salient entity, it seems that the affirmation of this attribute should also affirm the entity -- by logical implication. Thus, implicatures based upon entity/ attribute relations seem likely to involve only non-definitional attributes. However, it is difficult to discover attributes for any class which are truly definitional -- e.g., dogs have four legs but a three-legged dog is still a dog. And the affirmation of even truly definitional characteristics when *S* might, with equal effort, have affirmed the entity itself, seems to convey that *S* is unwilling to commit him/herself to the affirmation of the entity, as in 186.

- (186) A: Are whales mammals?
B: They have lungs and are warm-blooded.

S may also affirm an attribute to convey his/her inability to commit him/herself to other attributes which may be salient in the discourse, as in 187.

- (187) A: Do you speak Greek?
B: I've spent some time there.

Of course, if the salient relationship here is implication (so that *living in Greece* implies *knowing Greek*), a different meaning will be licensed. Similarly, if *S* denies an attribute, as in 188, s/he may confirm other attributes, unless a prerequisite

- (188) A: Do you speak Greek?
B: I've never lived there.

relationship rather than an entity/ attribute one is salient.

In sum, where v_i and v_j are attributes of some entity v_k , v_k can be seen as a higher value than v_i and v_j , and v_i and v_j can be seen as alternate values in this ordering of an entity with its attributes.

5.1.13. Summary

The orderings that permit speakers to license scalar implicatures thus range from those relatively domain-independent 'canonical' orderings inspired by the logical quantifiers and connectives to domain-dependent entity rankings -- and from linear orderings to hierarchical orderings. While most orderings which support Horn's scalar predication, Harnish's quantity-quality implicature, and Gazdar's scalar quantity implicature also support scalar implicature, there are a number of exceptions: Scalar implicature is supported by Horn's quantifier, modal, logical connective, cardinal and ordinal orderings, although I have formulated different accounts of implicatures licensed by the numerical orderings and additional orderings that may be derived from the logical connectives. While Harnish's notion of all-implication can be subsumed under the latter, I have argued that his concept of implicatures licensed via 'specification of assertion' contradicts not only intuitive interpretations of his examples but also his own accounts of quantity-quality implicature. Grice's notion that indefinites and definites can be ranked is clearly subsumable by scalar implicature, although just as clearly definiteness and indefiniteness can be given a more sophisticated analysis than scalar implicature provides. Many authors have identified ranked entities, states, actions, and attributes, including epistemic verbs and other hedges; I have identified still more, including orderings of verb/ preposition pairs; all support scalar implicature. From the temporal orderings noted in the literature as supporting quantity implicature, I have formulated a new analysis of the scalar implicatures that temporal references may license. I have also proposed that spatial and process/ prerequisite orderings support scalar implicature; although Harnish implicitly recognizes the latter, his justification is flawed. In addition to the linear orderings to which previous research has confined quantity implicature, scalar implicature may also be licensed by mention of values ranked in orderings such as set/ proper subset, whole/ part, type/ subtype, generalization/ specialization, instance-of, and entity/ attribute relationships.

While it might be possible to extend this listing process to produce an ostensive definition of the orderings that support scalar implicature, a more general definition is clearly to be preferred. Even if an exhaustive list of such orderings were possible, mere listing is theoretically unsatisfying. In practical terms, it would require that each such ordering be predefined, with no principled basis for including or excluding additional, perhaps domain-dependent orderings. In any event, orderings such as 168 indicate that exhaustive listing is impossible: People can and do form scales from items that may bear little intrinsic relationship to one another; these scales are hard to predict outside the context in which they are evoked.

5.2. Defining Scale

Horn, Harnish, and Gazdar have all attempted to characterize the orderings that support quantity implicature by some form of entailment. Although all have recognized flaws in this approach, none has devised a satisfactory substitute. In this section, I will examine previous attempts at characterizing these 'scales', look at proposed definitions of these scales, and point out their deficiencies.

5.2.1. Limits of an Entailment Condition

As noted in Section 3.2.1, Horn notes serious weaknesses in his own entailment definition of scale -- even for his canonical scales. For example, the largely intuitive definition Horn has been forced to propose for his quantifier orderings (See Section 5.1.1.) poses a serious problem for his entailment condition: The idea that *all x* entails *some x* may be satisfying from an intuitive if not a logical point of view. However, as Horn notes, *some/ all* represents only a portion of people's notion of a quantifier ordering; a more complete ordering would include the negative pole as well -- *no{ne}/ some/ all*. But clearly we would not want to say that *all x* entails *no x* -- nor that (189b) semantically entails (189a).

- (189)
- a. None of the people left early.
 - b. All of the people left early.
 - c. #No people left early and in fact everybody did.

Although (189a) may indeed convey that \neg (189b), it does not convey this meaning via conversational implicature; the infelicity of (189c) shows that (189a)'s conveyance of \neg (189b) is not cancelable. Horn simply omits these poles from his quantifier orderings.

For modal orderings too Horn omits negative poles, lest *possible* entail *impossible* and *permissible*, *impermissible*. Again, in exchanges such as 190, the affirmation of the negative pole *impossible*, while conveying that higher values

- (190) A: I'm beginning to think finding a lover is impossible. B: It's not impossible....

possible and *necessary* are false, does not convey this meaning via conversational implicature. So, again, negative poles must be omitted from such orderings -- not only to preserve an entailment condition, but also to predict possible conversational implicatures correctly.

Although this truncation strategy seems reasonable -- if counter-intuitive -- for quantifier and modal orderings, it is less easily applied to orderings like the miscellaneous modifiers discussed in Section 5.1.6.3. Again, a full specification of an intuitive *degrees of attractiveness*

ordering -- say, *hideous/ ugly/ plain/ pretty/ beautiful* -- cannot be defined by semantic entailment -- lest *beautiful* entail *hideous*, for example. But while it may be sufficient to omit negative poles from quantifier and modal orderings, it is less clear how best to mutilate modifier orderings. It is easy to say that *beautiful* does NOT entail *hideous* or *ugly* -- but does it entail *plain* -- or even *pretty* for that matter?

For the temperature scale *cold/ cool/ tepid/ warm/ hot*, Horn tries to resolve a similar dilemma. He claims that while the assertion of *warm* implicates *not hot*, the assertion of *cool* cannot implicate *not warm*; instead, *cool* asserts the negation of *warm*. The difference between the two cases for Horn is the difference in felicity between (191a) and (191b).¹⁰⁸

(191)

- a. #The soup is cool if not warm.
- b. The soup is warm if not hot.

On this basis, Horn divides the temperature scale into halves, *cool/ cold* and *warm/ hot*, with *cold* the positive pole of its half. So, *hot* entails *warm* but not *cold*, and *cold* entails *cool*. Middle values like *lukewarm* remain a problem, since they do not seem to be entailed by either *cold* or *hot*. But Horn's solution for this ordering -- as for *degree of attractiveness* above -- seems a misguided attempt to save his entailment condition. If we accept that a hot soup is NOT therefore also a cold soup, can we really accept that it IS a warm soup? Whatever *hot* entails about *warm*, e.g., *A hot soup is not only warm, it's hot.*, it should also entail about *cold*. Or are all these values just mutually exclusive states. At any rate, note that, in naturally occurring discourse (192), *cold* and *warm* seem to be perceived as values on a common scale -- and as licensing particular implicatures via their relative positions.

(192) A: It's cold out here.

B: It's not warm.

.....

A: Why didn't you just say 'yes'?

B: Because then you'd think I thought it was colder than it really is.

For B, *not warm* appears to convey $\neg\text{BEL}(\text{B}, \neg\textit{it's cool})$ as well as $\neg\text{BEL}(\text{B}, \neg\textit{it's cold})$; that is, B denies *warm* to implicate his lack of commitment to the falsity of all lesser values on a temperature ordering. B can deny only warm and nothing less than that. So, in this case, Horn's splitting strategy would not explain B's implicature.

Implicatures licensed in 193 and 194 provide further evidence that modifier orderings should not be divided: In the former, A implicates $\neg\text{BEL}(\text{A}, \neg\textit{they're all right})$ and $\neg\text{BEL}(\text{A}, \neg\textit{they're bad})$ on a modifier ordering *bad/ all right/ good*, while B

¹⁰⁸Horn distinguishes between the two in terms of SUSPENDIBILITY (See Section 5.2.2.); a *cool/ warm* relationship is not suspendible, while a *warm/ hot* relationship is.

(193) A: They're not very good though. The camera's lousy.
 B: But they're not bad.

(194) A: Did you do all right?
 B: I didn't do that good, you know.

implicates $\neg\text{BEL}(\text{B}, \neg i \text{ did all all right})$ and $\neg\text{BEL}(\text{B}, \neg i \text{ did (good) well})$ from the reversed ordering *good/ all right/ bad* -- that is, A and B have evoked differently oriented value orderings.¹⁰⁹ The conclusion from this exchange was that 'the pictures are all right' -- the intersection of the speakers' implicatures $\neg\text{BEL}(\text{S}, \neg\text{they are all right})$. In 194, B implicated $\neg\text{BEL}(\text{B}, I \text{ did all right})$.

While quantifier, modal, and some modifier orderings may be truncated or split, orderings such as the temporals and process/ prerequisites discussed above are more difficult to accommodate within an entailment definition. As noted in Section 5.1.7, truth of v_i in the future clearly does not entail the truth of v_i in the present or past. And while process orderings might be considered entailments in the sense that accomplishing stage v_j in a process 'entails' having accomplished stage v_i , this notion of entailment is far from the one or two-sided logical implication usually employed to define semantic entailment. And, of course, orderings such as *tort/ misdemeanor/ felony/ capital crime* and other ranked entity orderings (See Sections 3.2.1 and 5.1.6.1.) simply cannot be defined in terms of even the most general notion of entailment. In consequence, Horn proposes a more general definition of scale -- in terms of SUSPENSION.

5.2.2. Defining Scales by Suspension

Throughout his discussion of miscellaneous scalar predicates, Horn is concerned with identifying linguistic constructions that favor 'upper-boundedness' as a means of identifying additional scalar predicates. For example, he identifies 'scalar clues' in items like *downright* and *absolutely, let alone, even, only, and just*. So, when it becomes clear that his entailment definition of scale must be abandoned, Horn proposes an alternate definition via certain of these constructions.

Constructions like those represented in 195 suggest that *unattractive/ hideous, colonel/ general, good/ great, and unpleasant/ obnoxious* may be seen as (parts of) scales.

¹⁰⁹See Section 6.3.2.3 for a discussion of the hypothesis that mentioned poles can indicate the orientation of evoked orderings.

(195)

- a. George is unattractive if not {downright/absolutely} hideous.
- b. Frank isn't a colonel let alone a general.
- c. This picture is good even if it isn't great.
- d. Leo isn't obnoxious, {just/only} unpleasant.

In each case the explicit relation of v_i to v_j appears to assert the upper-boundedness which S 's assertion of v_i (alone) might implicate. Similarly, Horn notes that modification of a (possible) scalar predicate by *too*, as in 192, can also provide a test of scale.

(196) It's cold out but it's too warm for skiing.

'*Too v_j*', he explains, conveys that a weaker element on a scale than that which actually holds should hold. So, if it's *too warm*, then some v_i less than *warm* (e.g., *tepid*, *cool*) should hold.

Horn proposes to employ one class of such constructions, SUSPENDERS, which he exemplifies by *X if not Y* -- or, v_i if not v_j -- as a substitute for his entailment definition of scale. Suspenders are so termed because they function to suspend implicatures that might be licensed by the utterance of v_i alone. Some examples are shown in Table 5-1:

childish if not infantile	middle-aged if not old
acceptable if not attractive	satisfied, if not pleased
unusual if not unprecedented	excellent if not perfect
glossed over if not entirely overlooked	possible if not probable
lukewarm if not downright	some if not many
unsympathetic	

Table 5-1: Horn's Suspenders

While the first member of each pair may 'suspend' the second, the converse is not true; so, for example, '*many if not some*' is infelicitous if not contradictory.

From these suspensions, a scale Sc can be defined as follows: If ' p_i if not p_j ' is a felicitous utterance but ' p_j if not p_i ' is not, then p_j is a higher value than p_i on Sc . These scales need not be linear but may be hierarchical; that is, scalar predicates need only be partially ordered.¹¹⁰ Horn further notes that these scales may be infinite, as for the cardinals. Horn explains that all his entailment-defined scales can also be defined by felicity of SUSPENSION (See Section 2.4.2.1.). For example, a quantifier scale *some/ many/ most/ all* might be defined as in 197. Note particularly that (197d) is infelicitous; so, this suspension definition produces the 'right' results for negative poles.

¹¹⁰Horn clearly does not use partial order in the technical sense here, since he goes on to allow cycles, such as *days of the week* or *months of the year* as scales. (In fact, these ostensible cycles must be interpreted as linear to support quantity implicature, as I discuss in Section 5.3.1.)

(197)

- a. Some if not many of the people left early.
- b. Many if not most of the people left early.
- c. Most if not all of the people left early.
- d. #None if not some of the people left early.

Moreover, suspension can also define scales for which entailment definitions are suspect if not downright impossible, as in 198 (*condone/ promote*), 199 (*sick/ dying*), and 200 (*misdemeanor/ felony*).

(198) [George Jackson's] jailers condone racial prejudice, if they don't promote it.

(199) Camille is sick if not dying.

(200) Smoking marijuana is a misdemeanor if not a felony in every state of the union.

However, while a suspension definition of scale does indeed accommodate relationships which support scalar predication but which an entailment definition excludes, it is unsatisfactory in other ways: First, definition of scale by felicity of suspension is clearly an unsatisfactory solution from a computational point of view, unless all scales are to be predefined as such or felicity judgments are to be solicited interactively. Second, while values 'close together' on a scale may felicitously be suspended (as in (201a)) to define a suspension scale, attempted suspension of others is less clearly felicitous (as in (201b)).

(201)

- a. one if not two
- b. one if not forty
- c. thirty-nine if not forty

Given only the suspension test and (201b), *one* and *forty* will not appear on single scale. However, if (201a) and (201c) were instead the first and final pairs in a list of cardinal suspensions, the cardinal scale would be definable by suspension. Since suspension appears thus not to be transitive, a definition of scale based on suspension will succeed for Horn's scales only if individual suspension pairs are carefully chosen. And note also that it is not clear that 'forty if not thirty-nine' is infelicitous, as, for example, in 202.

(202) George isn't as old as he looks. I'm not sure how old he really is but I think he's only about forty if not thirty-nine.

So, Horn's suspension definition of scale will not in fact define the cardinals as scales, since it is possible to assert 'thirty-nine if not forty' and 'forty if not thirty-nine'. Third, as noted in Section 2.4.2.1, phenomena other than conversational implicatures are suspendable, including conventional implicature, presupposition, and even entailment. So, Horn's proposed definition

will identify scales such as *only vote/ did vote* from (203a) and *bald/ exists* from (203b).¹¹¹ So, while

(203)

- a. Only Muriel voted for Humphrey if even she did.
- b. The king of France is bald if there is a king of France.

a suspension definition of scale will accommodate those scales that entailment does not, it will also identify scales that do not support quantity implicature.

In sum, Horn's entailment definition of scale is too restrictive; yet his alternative suspension definition is both too weak, including as scalars items which clearly do not support the class of quantity implicatures he wants to define, and too constrained, excluding scales Horn would wish to include.

5.2.3. Gazdar's Definition of Scale

In adapting Horn's notion of scalar predication for computational use, Gazdar attempts a more formal definition of scale. For him, the chief difficulty in defining a quantitative scale is enforcing that items ordered on a scale be qualitatively similar as well as quantitatively ordered. For example, how can *believe/ know* be defined as a scale but not *regret/ know*?

Initially, Gazdar proposes that "identity of selectional restrictions" or "identity of item-induced presuppositions" might serve as similarity criteria. In [Gazdar 79a] he suggests that Thomason's [Thomason 72] notion of SORTAL CORRECTNESS (See Section 4.1.1.) might serve to constrain values on a single scale. He finds Thomason's definition too broad, allowing quantitative orderings such as *John/ person*, which do not support scalar implicature.¹¹²

In the end, however, Gazdar must abandon this attempt at a formal definition of scale, adopting the view that 'semantic informativeness' is a pragmatic notion that cannot be captured by the concept of semantic entailment. As noted in Chapter 4, Gazdar effectively accepts an entailment definition of scale for the purposes of demonstrating his formalism, assuming that such scales will be 'given'.

¹¹¹Note that Horn does not limit his 'suspenders' to 'if not', but includes those in this example.

¹¹²In fact, a *person/ John* ordering does support scalar implicature, as a generalization/ specification ordering:

(viii) A: Did John stop by today?
B: A person did.

5.2.4. Other Definitions of Scale

While analyses of other linguistic phenomena, such as *let alone* [Fillmore 83], *all but, at that, and much less* constructions, polarity [Fauconnier 75] -- and other forms of conversational implicature [Fauconnier 79] have centered upon a notion of scale, these analyses too have foundered on the inadequacy of an entailment definition of scale and the difficulty of finding an alternative. Other discussions of polarity [Ladusaw 80],¹¹³ pragmatic phenomena such as TOPICALIZATION [Prince 81b], EPITOMIZATION [Ward 83], and FOCUS/ CENTERING [Joshi 81, Grosz 83], and prosodic phenomena such as CONTRASTIVE STRESS [Culicover 83] and FALL-RISE intonation [Ladd 80] have proposed simple set membership or functional dependency to account for the relationship between an utterance and its context and the additional meanings that their phenomena may license. However, as noted above, simple set membership represents only one of the metrics which support scalar implicature, and, clearly, many scalar implicatures cannot be explained in this framework.¹¹⁴ And [Joshi 81, Grosz 83] acknowledge that their 'functional dependence' is only an approximation of the relationship they intend.¹¹⁵ So, from previous studies of quantity implicature -- and from work on other linguistic phenomena based upon 'strength of claim', the general consensus is that, while some metric for ranking utterances must be found, the most acceptable choice, entailment, is simply inadequate.

5.3. Ranking Utterances via Posets

From the discussion in Section 5.1 it should be clear that a satisfactory characterization of the orderings that support scalar implicature must accommodate not only orderings defined by entailment, but also orderings which Harnish, Horn, and Gazdar have identified as supporting quantity implicature that are not defined by entailment, as well as the additional orderings I have identified above which are neither linear or not defined by entailment. In this section, I propose a new definition of this group of relations and argue that only these relations support scalar implicature. I then demonstrate how, given this new characterization, it is possible to specify definitions of higher, lower, and alternate values among ordered items and thus to complete the formalization of the scalar implicature conventions I proposed in Section 3.3.

¹¹³Ladusaw in fact proposes that Fauconnier's entailment scales be defined as sets: So, for example, he claims that *father* entails *man* because the set of fathers is a subset of the set of men; and, if p_i entails p_j then the set of circumstances under which p_i is true must be a subset of the set of circumstances under which p_j is true; so 'be a father' \subseteq 'be a man'.

¹¹⁴In fact, as I will propose in Chapter 8, set membership appears to be insufficient to explain some of the phenomena for which it has been employed.

¹¹⁵[Ward 85b] discusses how the notions of utterance ranking proposed here might provide a reasonable substitute.

Above I have used the terms *relation*, *ordering*, *scale*, and *ranking* loosely to refer to those relationships among expressions or among the things denoted by expressions that serve to support scalar implicature. Before proceeding to a general account of these metrics, I will specify a more precise terminology: Henceforth, I will define orderings to be (mathematical) RELATIONS as follows: Let V be a set. Then a RELATION O on V is a subset of $V \times V$. The ordered pair $\langle v_i, v_j \rangle \in O$ is commonly denoted by $v_i O v_j$. The metric by which such a relation is constructed I will term the ORDERING METRIC of the relation, or, simply, its METRIC.

5.3.1. Relations Not Supporting Scalar Implicature

Above I have characterized the relations that support scalar implicature by listing those that do indeed support the phenomenon, much as previous authors have done. However, we can also specify which relations do **not** support scalar implicature. I have already established that a relation supporting scalar implicature does so by providing a way to determine, for any two items ordered in that relation, whether one item is higher or lower than another, or whether the two are alternates with respect to some common higher or lower value. So, relations that do **not** support scalar implicature are just those which fail to support such distinctions. These orderings include CYCLIC relations¹¹⁶, temporal parallelism, and, in general, relations that are not both REFLEXIVE¹¹⁷ and ANTISYMMETRIC¹¹⁸ or relations that are not both IRREFLEXIVE¹¹⁹ and ASYMMETRIC¹²⁰ -- as well as relations that are not transitive¹²¹

First, consider the cyclic relation evoked in the constructed 204. We can reasonably assume that the squares on a monopoly board constitute a cycle from our knowledge of board games.

¹¹⁶A CYCLE is a finite chain of edges connecting vertices (v_1, \dots, v_n) for which v_1 coincides with v_n . That is, there is a (unique) path from v_1 to itself.

¹¹⁷A relation O on a set V is reflexive iff for all $v_i \in V$ $v_i O v_i$. *Is_a_legal_representative_of* is a reflexive relation on the set of independent adults, for example, since each member of this set is his/her own legal representative.

¹¹⁸A relation O is antisymmetric iff for all $v_i, v_j \in V$ $v_i O v_j$ and $v_j O v_i$ implies $v_i = v_j$. The relation *is_as_tall_or_taller_than* is an antisymmetric relation.

¹¹⁹A relation O is irreflexive iff for all $v_i \in V$ $\neg v_i O v_i$. So, *brother_of* is irreflexive, since no man is his own brother.

¹²⁰A relation O on V is asymmetric iff for all $v_i, v_j \in V$, if $v_i O v_j$ then $\neg v_j O v_i$. The relation *is_taller_than* is asymmetric.

¹²¹A relation O is transitive iff for all $v_i, v_j, v_k \in V$, $v_i O v_j$ and $v_j O v_k$ implies $v_i O v_k$. So, a non-transitive relation is simply a relation for which this transitivity condition fails to hold. *Sibling_of* represents a transitive relation, while *friend_of* represents a non-transitive relation. If I am Mary Jo's sibling and Mary Jo is Henry's sibling, then I am also Henry's sibling. However, if Neal is Martha's friend and Martha is Derek's, it need not follow that Neal is also Derek's friend.

- (204) A: Did you pass Go on that turn?
 B: I got to Free Parking.

But, if B's response evokes this relation in the discourse, his/her response will be ambiguous. For, *getting to Free Parking* may or may not have involved *passing Go*. So, *Free Parking* will be both a HIGHER and a LOWER value than Go, since one is spatially located both 'before' and 'after' the other. Since the ranking of these two items with respect to one another is thus ambiguous, A has no basis for interpreting B's response -- nor B for calculating how his/her response will be interpreted, if a cyclic relation is perceived as salient. However, if B perceives the relation of *Free Parking* to *Go* to be linear rather than cyclic, and also believes *Free Parking* is ordered before *Go*, then B may implicate $\neg\text{BEL}(B, I \text{ got to } Go)$ by the response in 204. Alternatively, if B linearizes this cycle such that *Free Parking* is ordered after *Go*, his/her response may be interpreted as "I not only passed Go but I got as far as Free Parking."

Similarly, in 205, if the process of playing cards is seen as

- (205) A: Have you shuffled the cards?
 B: We've finished the game.

circular, B will be unable to convey any scalar implicature by this response. *Finishing a game* may be seen as both HIGHER and LOWER than *shuffling the cards* in this exchange. So, B might convey either that s/he hasn't shuffled because s/he doesn't intend to play another game, or s/he may be conveying that not only has s/he shuffled the cards -- but in fact s/he's finished a game.

Convention or context may suggest the linearization of these cycles to permit scalar implicature. For example, cycles such as days of the week or month, months of the year, and global spatial orderings appear to have conventional linearizations. With no further specification of context, B's response in (206a) is likely to be

- (206) A: Will the report be finished by the 30th?
 a. B: The first.
 b. B: The fifteenth.

interpreted as *no*. *The first* will be seen as that subsequent to the queried *thirtieth* rather than the *first* preceding it, perhaps because of some convention that conversational participants assume the closest value to the queried value in these linearizations when ambiguity exists. In (206b), with *thirtieth* and *fifteenth* diametrically opposed in the cycle, it is less clear that hearers will prefer one interpretation over another, unless perhaps some principle of ellision might be proposed such that, all other things being equal, reference to a month other than that of the deictic temporal center requires specification of that month.

So, cycles will not permit scalar implicature -- unless, by some such linearization, they become non-cyclic. For, any values v_i, v_j in a cyclic relation O must both precede and succeed

itself; that is, both $v_i O v_j$ and $v_j O v_i$. So, (the transitive closure of such) cycles are symmetric.¹²²

Other symmetric relationships, like *sibling_of* in 207, also fail to support scalar implicature:

- (207) A: Is John Bill's brother?
B: Mark is John's.

If A knows that Mark is also Bill's brother, B's response will allow him to deduce the direct response *no*. However, note that this conclusion would come not from implicature but from deduction. No scalar implicature can be licensed by this response.

A relationship such as temporal parallelism also fails to support scalar implicature. If *watching television* and *sitting down* are seen as related only in that they may be performed at the same time¹²³ in 208,

- (208) A: Are you watching television?
B: I'm sitting down.

then, not only will no scalar implicature be licensed, but B's response will probably not be deemed cooperative -- or even coherent -- by A.

Also, relations that are not transitive fail to support scalar implicature. For example, a response of (209a) cannot be interpreted as conveying a scalar implicature via the linear *isa_friend_of* relation apparently evoked, because such a relation is not transitive.

- (209) A: Is Jane Mary's friend?
a. B: She's George's friend and George is Mary's.
b. B: George is.

While A might infer that B does not believe Jane is Mary's friend, simply from B's failure to provide a direct response, this inference will not be conveyed via scalar implicature. But A cannot conclude that the proposition '*Jane is Mary's friend*' is either true or false unless she infers that B believes the *isa_friend_of* relation is transitive. Otherwise, B's additional information provides no basis for either implicature or deduction of a direct response. A response of (209b) can, of course, license the scalar implicature $\neg \text{BEL}(\text{B}, \textit{Jane is Mary's friend})$ via a set/ subset relationship.

Consider finally the relation of 'liking', which is neither reflexive and antisymmetric nor

¹²²A relation O on a set V is symmetric iff, for all $v_i, v_j \in V$, $v_i O v_j \Rightarrow v_j O v_i$.

¹²³They may be seen in other relationships that do support scalar implicature, such as members of the set of possibly activities or in a prerequisite relationship, of course.

irreflexive and asymmetric, and which is not transitive. Then B's response of (210a) will not permit the licensing of scalar implicatures:

- (210) A: Does Jane like Mary?
 a. B: Mary likes Jane.
 b. B: Harry likes Arthur.

In fact, it is difficult to know what conclusions to draw from (210a). It might be that B means to convey that, indeed, Jane must like Mary since Mary likes Jane. But it might also be that B conveys ignorance of whether or not Jane likes Mary. While the latter might seem a more likely interpretation, consider that (210b) will license the same implicature. So, it is not an ordering by 'liking' that supports the inference, but rather the set of liking relationships that B evokes given this interpretation of (210a) and (210b).

In sum, the relations that support scalar implicature -- including those relations identified previously in the literature as well as the orderings I have proposed above -- turn out to be just the class of PARTIALLY ORDERED SETS, or POSETS.¹²⁴

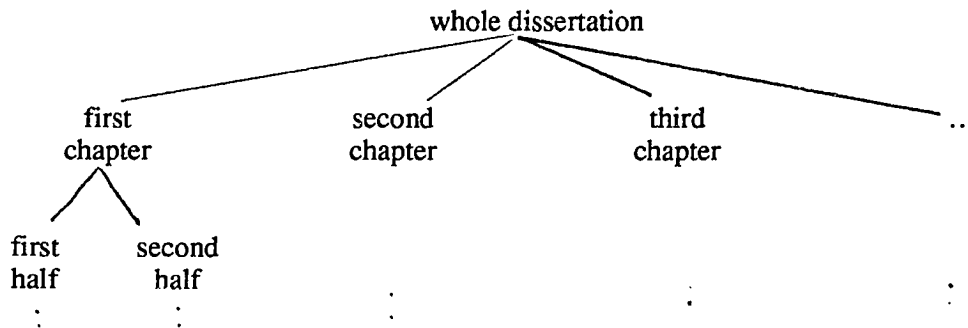
5.3.2. Poset Rankings

A partially ordered set or poset is defined in terms of a set V and a metric σ that partially orders it and denoted by $[V, \sigma]$. Posets supporting scalar implicature may be defined over (expressions that denote) entities, actions, attributes, times, places, or concepts, including those concepts ordered in Horn's canonical quantifier, modal, and number scales -- or any other items. So, a property may be ranked with respect to some entity which exhibits it via an *attribute_of* relation; an event may be ranked with other events according to temporal precedence; types and their subtypes may be ordered in an *isa* hierarchy; actions related by a *prerequisite_of* metric; subsets of a set ranked with respect to one another by an inclusion relation; and so on. Thus, a poset condition on scalar implicature allows us to accommodate not only linear but hierarchical orderings (ordered from root to leaf) in a single concept, and consequently provide uniform definitions of the HIGHER, LOWER, and ALTERNATE predicates introduced in Chapter 3.3.4.

¹²⁴While posets are commonly defined as orderings that are reflexive, antisymmetric, and transitive, Knuth notes that posets may be defined equivalently as irreflexive, asymmetric, and transitive [Knuth 73:258-59]. A relation satisfying the first definition of poset is *is as tall or taller than*, and one satisfying the second is *is taller than*. Note that one can always start with a relation R_1 satisfying the latter definition and produce one satisfying the former R_2 , simply by adding the ordered pairs $v_i \sigma v_i$ for all v_i in V . In effect, R_2 will represent the transitive and reflexive closure of R_1 , as with *is taller than* and *is as tall or taller than*. Not only do all the orderings discussed above satisfy one of these definitions of poset, but any poset -- in an appropriate context (See Section 6.3.2.3.) -- can support scalar implicature.

5.3.2.1. Defining Posets over Expressions

Posets are defined not over entities, actions, and so on, but over expressions (as defined in Section 4.1.2.2), or semantic representations, which correspond or refer to them. For example, model we might conceive of an ordering ‘parts of a dissertation’ defined over the set of *dissertation parts* by the metric *isa_part_of*. For such a conceptual ordering, we can define a corresponding partial ordering on expressions, such as *first chapter*, *first half*, *whole dissertation*, that refer to these parts. Such a poset might be represented by the following directed graph, where each node is labeled by such an expression:



So, the partial orders which we may use to calculate scalar implicatures will consist of sets of expressions, ordered via their correspondence to real-world entities and relationships, much as Gazdar defines his scales in terms of expressions (See Section 4.1.2.2.). (In a database environment, this level of representation is equivalent to the modeling of entities and relationships ‘in the world’ by database entities and relationships.)

Now we can require that, for S to license a proposition represented by some sentence p_j by an utterance represented by a sentence p_i , p_i and p_j must be related by the presence in p_i and p_j of expressions e_i and e_j , respectively, which can be partially ordered by some salient σ . Alternately, we might say that p_i and p_j must be related by some e_i and e_j which appear in some salient poset O . In either account, S 's belief that some partial ordering relation is mutually believed salient by S and H is a necessary condition of scalar implicature.

5.3.2.2. Higher, Lower, and Alternate Values

With poset to substitute for Horn's and Gazdar's ‘scale’ and Harnish's ‘measures of informativeness’, it is now possible to provide a semantics for the notions of HIGHER, LOWER, and ALTERNATE values used throughout the previous discussion of scalar implicature. In particular, we can provide definitions of the HIGHER, LOWER, and ALTERNATE predicates which were stipulated in Chapters 3.3 and 4.

For any partial ordering O on a set of expressions E , for all $e_i, e_j \in E$, $e_i \neq e_j$ ¹²⁵:

1. e_j is HIGHER than e_i wrt O iff $e_i O e_j$;
2. e_i is LOWER than e_j in O iff $e_i O e_j$;
3. e_i and e_j are ALTERNATE values in O iff e_i and e_j are INCOMPARABLE ELEMENTS¹²⁶ of O and there exists some $e_k \in E$ such that e_k is higher than both e_i and e_j or e_k is lower than both e_i and e_j in O .

These possible orderings of e_i and e_j in O are illustrated in the constructed example 211, where O is the ordering depicted above, 'parts of a dissertation'.

(211) A: Did you read the first chapter?

a. B: I read the first half of it.

b. B: I read the whole dissertation.¹²⁷

c. B: I read the third.

Let e_i be the expression queried by A (*first chapter*) and e_j be the expression mentioned in B's responses. Then, in (211a), e_i is LOWER than e_j in O ; e_i is higher than e_j in (211b); and e_i and e_j are ALTERNATES in O in (211c).

Now recall the scalar implicature conventions Imp_{1-3} presented in Section 3.3.4 as an axiomatization of Condition 3 on conversational implicature. By mentioning the lower value *first half* in (211a), B may implicate $\neg \text{BEL}(\text{B}, I \text{ read the (whole) first chapter})$. This implicature is predicted by the convention Imp_1 , since *first chapter* is a lower member of the poset 'parts of a dissertation' than *first half*. That is,

$$\begin{aligned} & (\text{BMB}(\text{B}, \text{A}, \text{SALIENT}(\text{'parts_of_a_dissertation'}, C_i)) \wedge \text{REALIZE}(I \text{ read the first} \\ & \quad \text{half of it, AFFIRM}(\text{B}, \text{first_half_of_chapter_one}, \text{BEL}(\text{B}, \\ & \quad \text{read}(\text{B}, \text{first_half_of_chapter_one})))) \wedge \text{IS_COOP}(\text{B}, C_i, \\ & \quad \{\text{QUANTITY, QUALITY}\}) \wedge \text{HIGHER_SENT}(\text{read}(\text{B}, \text{chapter_one}), \\ & \quad \text{read}(\text{B}, \text{first_half_of_chapter_one}), \text{'parts_of_a_dissertation'}) \\ & \Rightarrow \text{SCALAR_IMP}(\text{B}, \text{A}, I \text{ read the first half of it, } \neg \text{BEL}(\text{B}, \\ & \quad \text{read}(\text{B}, \text{chapter_one})), C_h) \end{aligned}$$

Similarly, B's implicature in (211c) may also be accounted for by Imp_1 :

$$(\text{BMB}(\text{B}, \text{A}, \text{SALIENT}(\text{'parts_of_a_dissertation'}, C_i)) \wedge \text{REALIZE}(I \text{ read the third,$$

¹²⁵This condition is necessary for posets defined as reflexive, antisymmetric, and transitive lest some e_i be defined as higher or lower than itself.

¹²⁶Elements are incomparable if they are not ordered with respect to one another in O , that is, if neither $e_i O e_j$ nor $e_j O e_i$.

¹²⁷While some have found this response odd, consider it in the following context: B wonders whether to interpret A's query as 'Did you read *only* the first chapter?' or as 'Did you read *at least* the first chapter?'

$$\begin{aligned} & \text{AFFIRM}(B, \text{third_chapter}, \text{BEL}(B, \text{read}(B, \text{third_chapter}))) \wedge \\ & \text{ALT_SENT}(\text{read}(B, \text{chapter_one}), \text{read}(B, \text{third_chapter}), \\ & \text{'parts_of_a_dissertation'}) \\ & \Rightarrow \text{SCALAR_IMP}(B, A, I \text{ read the third, } \neg \text{BEL}(B, \text{read}(B, \text{chapter_one})), \\ & C_i) \end{aligned}$$

There are no restrictions on those posets which support scalar implicature. However, (at least) one restriction does exist on which posets may be viewed as salient in a given exchange: Above (Section 5.1.6.3) I noted that, for most metrics that rank utterances, both a given metric and its dual (converse) may be candidates for salience in an exchange. However, no metric σ_i which orders values v_i and v_j such that a) v_i is higher than v_j and b) the truth of v_j entails the truth of v_i can support scalar implicature -- for the simple reason that, in such a case, a sentence p_i ranked higher than a sentence p_j by σ_i since then the implicature licensed would be inconsistent with the utterance licensing it. In terms of the formalism presented in Chapter 2, such a meaning would not be reinforceable. Consider, for example, (212a):

- (212) A: Are you planning to buy a dog?
 a. B: A German Shepherd.
 b. B: I'm buying a German Shepherd and I'm not buying a dog.

While one might identify either an ordering defined by 'isa' (i.e., a German Shepherd *isa* dog) or by 'subsumes' (i.e., a dog subsumes the subtype German Shepherd) as salient in this exchange, only the latter permits scalar implicature here. B cannot implicate that she is not buying a dog via this response, since buying a German Shepherd entails buying a dog. The attempted reinforcement of (212b) fails. However, we cannot rule out 'isa' relations as potential supporters of scalar implicature: In 213, for example, B's response might evoke either an 'isa'

- (213) A: Would you like a dog?
 B: I'd like a German Shepherd.

hierarchy -- or its dual. Apparently, any poset can support scalar implicature, although other tests for conversational implicature may rule out some particular posets in particular exchanges.

5.3.2.3. Representing Scalar Implicature Orderings as Posets

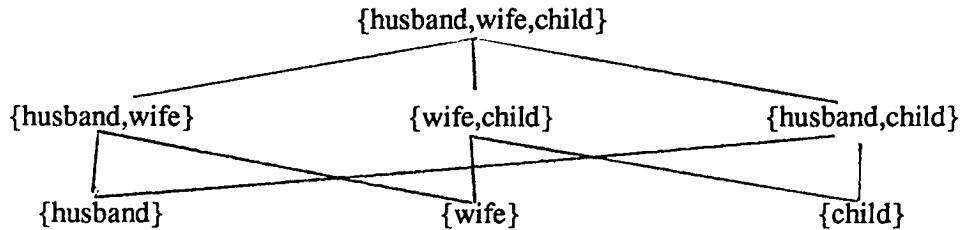
I have demonstrated above how part/ whole relations can be represented. To demonstrate that the other orderings discussed in Section 5.1 are accounted for by a poset condition, I will describe how representative orderings can be accommodated by this condition so that potential scalar implicatures are correctly predicted by Imp_{1-3} .

Relations defined by ordering the non-null members of the power set of some set x by set-inclusion allow a poset representation of x and its non-null proper subsets as follows: Any non-null proper subset of a set may be ranked as LOWER than the set which includes it, and that

set, in consequence, will represent a HIGHER value in the ordering. Subsets which are neither included in, nor include, one another, will be ALTERNATE values in this poset. Consider how the salient ordering in the following exchange might be represented:

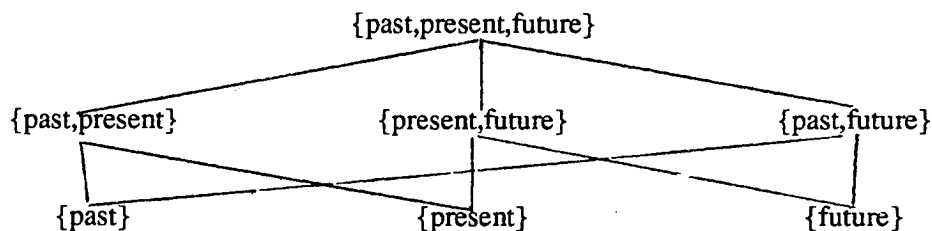
(214) A: Do you speak Portuguese?
 B: My husband does.

The inclusion ordering which supports the implicature in 214 might be represented as follows:



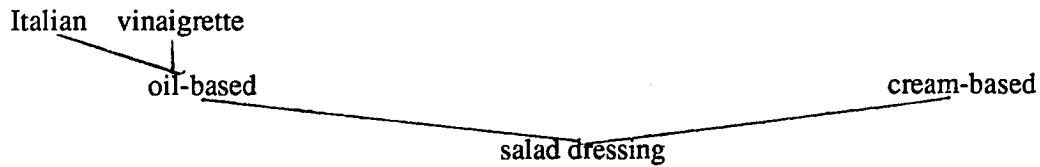
So, *{husband,wife,child}* will be the highest value in this ordering, with the alternate doubletons *{husband,wife}*, *{wife,child}*, and *{husband,child}* lower values and the alternate values, *{husband}*, *{wife}*, and *{child}* lower values still in this poset. By the scalar implicature conventions, then, *S* may affirm, say, *{husband,wife}* to convey $\neg\text{BEL}(S, \{husband,wife,child\})$ as well as $\neg\text{BEL}(S, \{husband,child\})$ and $\neg\text{BEL}(S, \{wife,child\})$. Note, particularly, that there may be some redundancy in scalar implicatures predicted from this representation. Also, any subsets so represented may be lexicalized in various ways -- as, the expression *{husband,wife}* might be lexicalized as 'couple' or as 'husband and wife'. The theory presented in this thesis will not distinguish between these.¹²⁸

As noted in Sections 5.1.7, temporal orderings may also be represented as sets of temporal for the analysis of licensed scalar implicatures. So, these orderings too will be defined by set inclusion, as:



Posets defined by a type/ subtype metric, such as that which supports 174, may be illustrated by the (partial) classification hierarchy:

¹²⁸But see [Corella 84, Kalita 84] for some approaches to this problem.



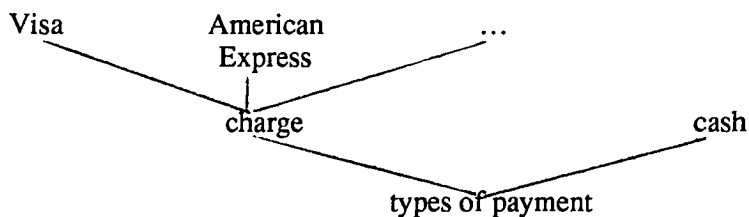
In this poset, *salad dressing* will be a lower value than both *Italian* and *vinaigrette*; thus, the latter types of dressing will be alternate values in the poset. The affirmation of a subtype e_i will thus implicate $\neg\text{BEL}(S, e_j)$ for alternate subtypes e_j in the poset, by Imp_1 . Given the same salient poset, S might convey $\neg\text{BEL}(S, \neg I \text{ want salad dressing})$ by denying *Italian* or *vinaigrette*, as in 215.

- (215) A: Do you want salad dressing?
B: I don't want Italian.

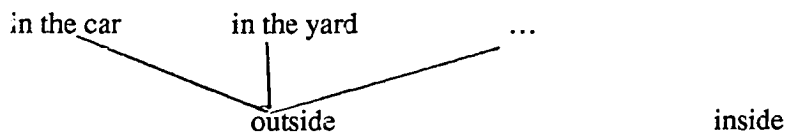
And S may convey $\neg\text{BEL}(S, I'd \text{ like vinaigrette})$ or $\neg\text{BEL}(S, I'd \text{ like Italian})$ by affirming *salad dressing*, as in 216.

- (216) A: Would you like vinaigrette?
B: I'd like some salad dressing.

Instance-of and generalization/ specialization relations will be represented similarly, with instances represented higher values than the classes they are instances of, as in the following representation of a relation which might account for the exchange in 179:

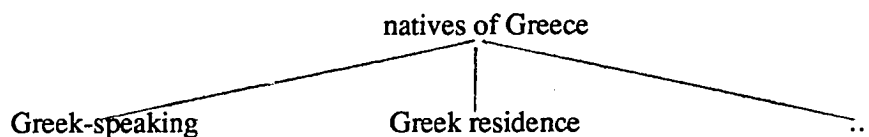


Generalization/ specialization hierarchies may also be represented in this form, as in the following representation of a relation supporting the implicature in 180, 181, and 182:



Specializations are higher values in these orderings than their generalizations.

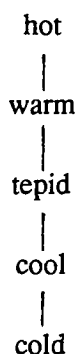
Entity/ attribute orderings may be represented in an 'is_an_attribute_of' relation. So, the following digraph represents an orderings which supports the implicature in 185:



Since linear orderings¹²⁹ are themselves partial orderings, orderings such as those defined on the quantifiers, modals, logical connectives, number scales, and on indefinites and definites and locations, as well as those ranked entities, states, actions, and attributes and those process and prerequisite orderings that are linearly ordered may all be represented by posets. For example, quantifiers might be ranked by the informal notion of entailment discussed above (See Section 5.1.1.), and ordered as follows:



Instead of splitting temperature scales as Horn does, we can now define an ordering on temperature values by a metric like *is_cooler_than* to form a poset supporting implicatures like that licensed in 192:



Alternately, the dual to this poset, formed from a metric such as *is_warmer_than* on the same set of temperature values will support the implicature licensed in 217 -- i.e., $\neg\text{BEL}(B, \textit{the lemonade is cold})$.

¹²⁹Orderings for which any two elements are comparable.

- (217) A: Is the lemonade cold?
 B: It's cool.

For rankable entities or concepts, such as Horn's *bête-noir*, *tort/ misdemeanor/ felony/ capital crime*, values now may simply be ordered by a metric such as *is_a_lesser_offense_than* to account for the implicatures discussed in Section 3.2.1.

5.4. Summary

All of the relations described in Section 5.1 as supporting scalar implicature share the characteristics of the class of partial ordering relations. Relations that do not support scalar implicature do not exhibit these characteristics. So, a poset condition on scalar implicature appears to capture just those intuitions I have described in previous chapters in a simple and principled way. Furthermore, it seems that any poset may support scalar implicature, in an appropriate context. In addition, this approach allows for the use of novel relations to license scalar implicature. In Chapters 6 and 7, I will describe further advantages of this method of utterance comparison, including ways to compute these relations from an existing knowledge base.

I have mentioned throughout previous chapters some of the problems presented by the role that world or domain knowledge as well as the current discourse context plays in the licensing and interpreting of quantity implicature in general. Definitions of ordering relations may vary from person to person and context to context and thus influence what implicatures may be conveyed. A poset condition permits the specification of the extent to which these definitions may vary without communication breakdown occurring. Even if *S* and *H* define some ordering relation similarly, a *S* can only license a quantity implicature if s/he perceives that that ordering is salient in the discourse and believes that *H* too recognizes this salience. In Chapter 6, I will propose how certain aspects of the discourse context may be accessed to determine which ordering relations may be salient for a given exchange.

CHAPTER VI

Computing Scalar Implicatures

Almost zero is better than less than nothing.
Brian Henry, *The Winds of War*, 2/11/83.

In my discussion of scalar implicature and its predecessors, scalar predication, quantity-quality implicature, and scalar quantity implicature, I have noted the critical role that context plays in the licensing of scalar implicatures. I have proposed that this role may be interpreted as the specification of salient orderings in the discourse which permit inferences about values in such orderings which *S* has not explicitly committed him/ herself to. To represent this specification so far, I have proposed a predicate $SALIENT(O, C_h)$ to denote the salience of *O* in a discourse context C_h and have claimed that *S* must believe this salience to be a mutual belief of him/ herself and of *H* in order to make some implicature involving recognition of *O*.

In the present chapter I will examine how context in general and the salience of expressions and posets in particular affects *S*'s licensing of scalar implicatures and *H*'s interpretation of them in greater detail. I will first examine possible disparities in speaker/ hearer definitions of particular posets and how a notion of poset COMPATIBILITY allows us to relax the condition that *S* must believe the salience of some particular poset is a mutual belief of *S* and *H*. I will then discuss strategies for determining the salience of particular orderings in given discourse contexts. Finally, I will present a revised version of the theory of scalar implicature, incorporating salience information and allowing for the presence of multiple salient expressions in an utterance.

It should be noted throughout the following discussion that no claims are made for the cognitive reality of the algorithms or definitions presented below. I will only claim plausibility through computational feasibility. Neither do I claim any contribution to the literature on salience. I will only suggest what salience information is required and why, and how trends in current research, when successful, might accommodate these needs. So, I will be proposing tentative and partial solutions to how our current understanding of salience and apparently related concepts may be used to incorporate contextual information into the calculation of scalar implicature.

6.1. Context and Scalar Implicature

Recall that Gazdar's approach to the calculation of utterance meanings is to define functions which, for any utterance, will produce sets of entailments, potential presuppositions, and potential implicatures which that utterance might license. He then determines the actual meanings licensed by adding first entailments, then potential clausal quantity implicatures, potential value quantity implicatures, and potential presuppositions, to the discourse context. So long as each addition produces a consistent context, the context will be incremented and potential meanings will become actual meanings of the utterance. Meanings whose addition would produce an inconsistent (in the sense defined in Section 2.4.2.1) context Gazdar deems to be canceled by that context.

But this notion of context involves only prior assertions which are logically inconsistent with a potential implicature. In effect, if u_i can ever, under any circumstances, license p_j , Gazdar's system, as he has presented it, will predict that it will -- unless $\neg p_j$ or some sentences that logically imply it have previously been realized in the discourse.

Of course, it is not a simple matter to determine whether or not some expression will evoke an ordering or which ordering some expression will evoke. For example, in 218, B apparently perceives an ordering

- (218) A: Have you, me, and Ellen ever had dinner together?
B: We've had lunch.

defined over the *set of meals* as salient.¹³⁰ B implicates $\neg\text{BEL}(\text{B}, \textit{we've had dinner})$ -- as well as $\neg\text{BEL}(\text{B}, \textit{we've had breakfast})$ -- since *dinner* and *breakfast* are each alternate values with *lunch*. But B might perceive other orderings as salient from the same query, as in 219.

- (219) A: Have you, me, and Ellen ever had dinner together?
B: We've been to the movies.

Here the salient ordering appears to be defined on some *set of social events*. While one of the implicatures licensed by B's response will be the same as one licensed by 218 -- $\neg\text{BEL}(\text{B}, \textit{we've had dinner})$ -- another might be $\neg\text{BEL}(\text{B}, \textit{we've gone for drinks})$. But clearly, even if A and B both recognize some *set of social events* as salient, how will they know that it is the same

¹³⁰This ordering may be represented as follows:

	{breakfast,lunch,dinner}		
{breakfast,lunch}	{breakfast,dinner}		{lunch,dinner}
{breakfast}	{lunch}		{dinner}

such set? Can B really implicate $\neg\text{BEL}(\text{B}, \textit{we've had drinks})$ unless s/he believes s/he and A both recognize not only the salient ordering metric but also the set of events over which it is defined? And, in 220, the salient ordering supporting B's implicature that $\neg\text{BEL}(\text{B}, \textit{you, me, and Ellen have had dinner together})$ is not only different from the salient orderings evoked in 218 and 219, but is evoked by B's perception of a different expression -- $\{you, me, Ellen\}$ -- as salient.

(220) A: Have you, me, and Ellen ever had dinner together?
B: Ellen and I have.

Finally, 221 illustrates how *S* may evoke multiple salient orderings from multiple expressions in a single utterance:

(221) A: Have you, me, and Ellen ever had dinner together?
B: Ellen and I have had lunch.

In this exchange, orderings defined over $\{you, me, Ellen\}$ and $\{breakfast, lunch, dinner\}$ are both evoked by B's response -- to implicate $\neg\text{BEL}(\text{B}, \textit{you, me, and Ellen have had dinner together})$. However, how we might model other apparent implicatures, such as $\neg\text{BEL}(\text{B}, \textit{you, me and Ellen have had lunch})$ and $\neg\text{BEL}(\text{B}, \textit{Ellen and I have had dinner})$, is not so clear.

So, to calculate a scalar implicature from an utterance with semantic representation p_i , we need some way of determining which of all the expressions in p_i *S* is employing in licensing some scalar implicature (and, thus, must anticipate *H* will recognize) and which of potentially infinitely many possible ordering relations will be salient for those expressions. We must also determine how *S* and *H* can assume mutual recognition of the actual definition of such orderings.

6.2. Defining Posets in Discourse

Harnish, Horn, and Gazdar all implicitly assume that *S* and *H* share knowledge of a common set of ordering relations, each of which is automatically evoked when one of its values is referenced. But, of course, even when *S* and *H* perceive orderings that seem intuitively quite similar as being evoked by a given value, they may differ critically for the purpose of successful communication of scalar implicature. If *S* believes some ordering O_i is salient and relies upon *H*'s perception of O_i as such for the interpretation of his/her implicature, but if *H* believes some distinct ordering O_j is salient instead, then *S* may not accurately anticipate the inference(s) that *H* will derive from *S*'s utterance -- and thus *H* may infer meanings *S* has not intended. Or, *H* may fail to infer meanings that *S* has intended to convey. So, in anticipating *H*'s general ability to 'work out' a scalar implicature, *S* must, in particular, anticipate whether *H* will be able to identify the ordering that relates *S*'s utterance to other utterances s/he might have made and whether *H* will define this ordering as *S* does.

6.2.1. Sources of Divergent Definition

It seems reasonable to assume that *S* and *H* are more likely to disagree on the definition of domain-dependent and non-canonical orderings than on canonical orderings. Speakers' knowledge of canonical orderings such as the cardinals, quantifiers, and modals and many modifiers appears to stem from their linguistic competence or very general world knowledge and thus seems more likely to be shared among speakers. That is, knowledge of the relative positions of *one* and *four* or *some* and *all* or *pretty* and *beautiful* in their respective orderings should vary little from speaker to speaker. But even for such canonical orderings as the quantifiers, the relative positions of *many*, *most*, and *quite a few* may not be agreed upon by *S* and *H*. And for orderings such as those which permit scalar implicature in 222, it would

- (222) A: Do you want a Rainbow?
 a. B: I want a 1200 baud modem.
 b. B: I want an IBM PC.
 c. B: I want a piece of gum.

be most surprising if *S* and *H* could define the same set of salient entities evoked by each response.

While *S* and *H* may identify divergent orderings as salient under any circumstances, such divergence seems most likely when orderings are domain-dependent and when *S* and *H* possess unequal expertise in that domain. In 223, for example, the distinction between *pneumonia* and

- (223) A: You mean, I have pneumonia!?
 B: A mild case.

mild pneumonia may not be apparent to a patient. However, such exchanges may convey to the less-expert partner the information that some term *H* had not viewed as a potential value might in fact be so viewed, i.e., that an ordering exists on which *pneumonia* and *a mild case of pneumonia* are values. Thus speakers need not always avoid the use of orderings their hearers may not recognize, but simply recognize their possible consequences -- at best a request for clarification, at worst, miscommunication. And, by their use, *S* may actually convey to *H* that some previously unknown ordering in fact exists. So, in 223, while A may not actually understand the ordering B employs, s/he will be able to infer that there are 'degrees' of pneumonia.

Alternatively, if *H* has more expertise than *S*, s/he may derive implicatures *S* does not intend by interpreting some reference in a too technical light. In 224, if A interprets B's response in a more technical sense than B, s/he may conclude $\neg\text{BEL}(\text{B}, \textit{the child is an idiot})$, if she knows that idiots are technically classed as

(224) A: Is that child handicapped?
 B: He's an imbecile.

having greater intelligence than imbeciles. In both cases of unequal expertise, some modeling of S or H is required to assess licensed implicatures properly. Such user modeling will not be pursued here, but its utility should be noted for these cases.

6.2.2. Defining Poset Compatibility

Although above I have noted the potential for communication breakdown when S and H do not share a common definition for some ordering, and I have in previous chapters demanded that some particular ordering's salience be a mutual belief of S and H for scalar implicature, this constraint is in fact too strong. I will now propose a relaxation of this condition, expressed in terms of the metric σ and the set E that, together, define all orderings O .

From the examples given above it can be noted that two orderings O_i and O_j may differ from one another either in the set of expressions they are defined over or in the metrics which define those expressions -- or, of course, in both. By our definition of orderings as posets, for any two posets O_i and O_j to be termed identical, they must be defined over identical sets of referents E_1 and E_2 by identical metrics σ_i and σ_j . So, for each ordered pair $e_i\sigma_i e_j \in O_i$, $e_i\sigma_j e_j \in O_j$, and for each $e_i\sigma_j e_j \in O_j$, $e_i\sigma_i e_j \in O_i$. That is, orderings are identical iff they are defined over the same set of referents and if the metrics ordering these sets order their members in exactly the same way. And orderings may differ if either E_1 and E_2 are different or if σ_i and σ_j are different. But note that the definition of a poset is independent of the label S and H place on either the set or the metric or the partial ordering itself: That is, σ_i and σ_j or E_1 and E_2 or O_i and O_j may be LABELED differently by S and H -- yet still be identical according to this definition.

But even if S does not believe that the ordering O_i s/he believes to be salient will be recognized by H , s/he may still use O_i to support a scalar implicature given that s/he believes certain less restrictive conditions hold. Suppose S believes O_i to be defined by σ_i over E_1 but s/he also believes that H may only recognize a subset E_2 of E_1 as ordered by σ_i . For example, say in 225, B recognizes a *set of desserts*

(225) A: Do you have jell-o?
 B: We have pie.

ordering defined over the set $\{pie, jell-o, cake, ice-cream, fruit\}$, but A's *set of desserts* does not include *fruit*. Although the ordering by which B anticipates and A interprets scalar implicatures in this exchange will thus differ, the inferences A will draw given an ordering defined over $\{pie, jell-o, cake, ice-cream\}$ will be at least consistent with those A implicates via the larger set.

That is, while B may intend also to convey $\neg\text{BEL}(\text{B}, \textit{we have fruit})$ by 225, all those scalar implicatures that are understood by A will be scalar implicatures -- i.e., will be intended by B.

When O_i and O_j are identical except in that the set of referents over which O_j is defined, E_2 , is a proper subset of that over which O_i is defined, E_1 , then O_j will also be a proper subset of O_i . When S believes that O_i is salient but has reason to believe that H may recognize some O_j , then S will nonetheless believe that scalar implicatures inferred will be consistent with scalar implicatures intended, since the former will be a subset of the latter.

If instead S has reason to believe that H may recognize a larger set of salient referents ordered in O_j than S recognizes as salient in O_i -- if E_1 is a proper subset of E_2 and, thus S 's O_i is a proper subset of O_j -- then S may suppose that his/her knowledge will be accurately interpreted by H . Say, in 225, it is B, not A who fails to include fruit among his/her set of desserts. Then, B may utter his/her response while believing that s/he does have fruit -- but A may understand that $\neg\text{BEL}(\text{B}, \textit{we have fruit})$. In such cases, when S has reason to believe that O_i is a proper subset of O_j , S may indeed anticipate that H may draw false conclusions from his/her implicature.

So, I will term S 's salient ordering O_i COMPATIBLE with any H 's ordering O_j if O_j is a subset of O_i ; that is, O_j is COMPATIBLE with O_i if O_j is either identical to O_i or if O_j is a proper subset of O_i . In other words, O_i will only be incompatible with O_j if O_i is itself a proper subset of O_j . S may then employ any O_i to license scalar implicature if O_i is compatible with the O_j (s) H may recognize as salient.

Two orderings O_i and O_j may also differ in that, while they are defined over the same set $E_1=E_2$ of referents, two different metrics, σ_i and σ_j , order E_1 . That is, for some pair of expressions e_i and e_j , while $e_i, e_j \in E_1$, and $e_i, e_j \in E_2$, $e_i, \sigma_i, e_j \in O_i$ but $e_i, \sigma_j, e_j \notin O_j$ or $e_i, \sigma_j, e_j \in O_j$ but $e_i, \sigma_i, e_j \notin O_i$.¹³¹ So, for example, in 226, if B cannot anticipate that

(226) A: Is the Pacific Fleet on red alert?
B: The First Battle Division is.

A will recognize that *Pacific Fleet* is a higher value than *First Battle Division*, s/he will not be able to implicate $\neg\text{BEL}(\text{B}, \textit{the Pacific Fleet is on red alert})$ -- even if s/he does believe A will recognize that *Pacific Fleet* and *First Battle Division* appear in a whole/part ordering. For, if A instead believes that *Pacific Fleet* is a part of *First Battle Division*, then A will conclude that $\text{BEL}(\text{B}, \textit{the Pacific Fleet is on red alert})$ instead.

¹³¹Again, remember that metrics σ_i and σ_j are identical just in case, for all $e_i, \sigma_i, e_j \in O_i$, $e_i, \sigma_j, e_j \in O_j$, no matter what the metrics are labeled.

However, like the sets of expressions over which orderings are defined, the metrics which define them need not be identical for S to believe that his/her implicature will be 'worked out' by H . If, for all $e_i \sigma_j e_j \in O_j$, $e_i \sigma_i e_j \in O_i$, then, again, we can say that O_j is compatible with O_i . For any relations O_i and O_j on a set E_1 , O_i on E_1 COVERS O_j on E_1 if $O_j \subseteq O_i$. So, if O_i covers O_j , then, again, O_j will be compatible with O_i , and S may anticipate that any implicatures based upon H 's recognition of O_i will be inferred -- though not all those implicatures may be inferred. So, O_j will be compatible with O_i when $O_j \subseteq O_i$. That is.¹³²

Ordering Compatibility:

$$O_j \subseteq O_i \Rightarrow \text{COMPATIBLE}(O_j, O_i)$$

If S believes that H will recognize an O_j that is compatible with the O_i S recognizes as salient, S will at least believe that the inferences H draws will be consistent with those S has implicated. In such case, we will say that S believes some ordering to be AVAILABLE for some H in some context C_h :

$$(\text{BEL}(S, \text{SALIENT}(O_i, C_h)) \wedge \text{BEL}(S, \text{BEL}(H, \text{SALIENT}(O_j, C_h)))) \wedge \text{BEL}(S, \text{COMPATIBLE}(O_i, O_j)) \Rightarrow \text{BEL}(S, \text{AVAIL}(H, O_i, C_h))$$

However, while we will want to demand that S believe O_i to be compatible with all O_j that H might infer as salient in its place, we will not want to demand that S believe that all orderings s/he believes salient in some context C_h are compatible with all orderings H believes salient -- e.g., that, in 221, the ordering of $\{you, me, Ellen\}$ is compatible with the ordering of $\{breakfast, lunch, dinner\}$. For this and other reasons which will become clearer in Section 6.3, I will redefine $\text{SALIENT}(O_i, C_h)$ as $\text{SALIENT}(e_i, O_i, C_h)$ -- that is, some ordering is salient in a context for some particular expression which appears in it. Then we can redefine the notion that S believes O to be available to H via mention of e_i in C_h as follows:

Belief in Availability of Ordering:

$$(\text{BEL}(S, \text{SALIENT}(e_i, O_i, C_h)) \wedge \text{BEL}(S, \text{BEL}(H, \text{SALIENT}(e_j, O_j, C_h)))) \wedge \text{BEL}(S, \text{COMPATIBLE}(O_i, O_j)) \Rightarrow \text{BEL}(S, \text{AVAIL}(H, e_i, O_i, C_h))$$

With ordering compatibility and availability to H defined, and with ordering salience redefined, we can now revise the salience condition presented in the definitions of the scalar implicature conventions -- $\text{BMB}(S, H, \text{SALIENT}(O, C_h))$ -- in terms of inference consistency as $\text{BEL}(S, \text{AVAIL}(H, e_i, O, C_h))$. So, one mutual belief requirement in our definition of conversational implicature can be relaxed.

¹³²As Bonnie Webber (Personal Communication) has suggested, in 226, if B believes that A considers *Pacific Fleet* and *First Battle Division* to be alternate values in some ordering, then s/he can also license the desired implicature. So, there seem to be other ways in which orderings are compatible in the sense described here of licensing only implicatures consistent with S 's beliefs. However, in such a case, B should surely 'square away' the disparity between his/her beliefs and A's [Joshi 82].

6.3. Salience and Scalar Implicature

While studies of conversational implicature have readily acknowledged the overriding importance of 'context' in the generation and interpretation of implicature, little attempt has been made to define how the particulars of an exchange interact with the conventions that underly implicature. I noted in Section 3.2.2 that, of the authors who have studied quantity implicature, only Harnish explicitly demands that the metric underlying a particular implicature must be 'relevant' in the discourse. Although Horn mentions that context, as well as explicit cancelation or suspension, can block implicatures, he goes no further than this. And, given his discussion of how implicatures involving positive poles **must** be inferred, it is difficult to see how the mention of some expression like *some* will **not** license \neg *all* in his theory -- unless it is explicitly canceled. Although Gazdar includes a more computational notion of cancelation in his theory, his too in effect predicts implicature from the semantic representation and explicit cancelation alone. His value quantity implicature is context-dependent only insofar as it can be canceled by prior utterances in the context. Otherwise, all possible implicatures which might be generated by an utterance are generated, regardless of context or speaker intention.

If *S* does **not** believe it relevant to 'affirm the highest value he truthfully can', surely s/he may -- cooperatively -- choose not to do so. For example, it seems equally likely that plausible that *A* **not** have implicated \neg BEL(*A*, *you took a lot of computer courses*) in 89,¹³³ even though in this exchange *B* has clearly inferred it. As O'Hair [O'Hair 69:41] notes (summarizing Grice) "Unless there are outweighing good reasons to the contrary, one should not make a weaker statement rather than a stronger one *if the audience is interested in the extra information that could be conveyed by the latter.*" (My italics) Such factors justify the inclusion of a salience condition as described above in the definition of AVAIL. I will incorporate these notions of audience interest and relevance by claiming that the licensing of a scalar implicature is dependent upon whether the ordering that might support such an implicature is perceived as SALIENT in the discourse.

The notion of SALIENCE in AI and natural-language literature has so far been confined to discussion of the salience of objects and their attributes [Carbonnell 70, Conklin and McDonald 82, McCoy 85]. The salience of orderings or relations -- that is, of entities together with orderings on them -- examined in [Herskovits 82] and mentioned in recent work by Grosz and Sidner [Grosz and Sidner 85], is much less well understood. Psychological literature on human

¹³³That is, in

A: And you took some courses...

B: I've taken a lot of computer courses in high school.

propensities for viewing objects in terms of one relationship or another has been largely confined to children's ability to construct taxonomic hierarchies.

There are two possible approaches to the identification of the salient expressions and orderings necessary for the calculation of scalar implicatures in discourse: First, we might identify one or more expressions which denote items marked as salient in discourse and search for any posets they may appear in. Second, we might identify relationships marked as salient in the discourse, see if they represent partial orderings, and then identify any expression in the sentence under consideration which may appear in this poset.¹³⁴ I will examine each of these possible strategies below.

6.3.1. Identifying Salient Expressions

As Gazdar notes, identifying orderings evoked by an utterance is not a simple matter of identifying a lexical item in the surface representation of an utterance as a value in some ordering. However, neither is it a matter of identifying predefined values in some semantic representation -- as Gazdar and Horn in effect do. An expression may evoke an ordering in some contexts which it will not invoke in others -- even when it is part of an identical sentence. So, mention of *some* or *could* or *three* need not always convey \neg *all*, \neg *should*, and \neg *four*. However, it does seem plausible that domain independent orderings may more confidently be assumed recognizable by both *S* and *H*. So, all other things being equal, a domain independent ordering will be more likely to be assumed salient for the purpose of supporting scalar implicature than will a domain dependent orderings.

Identifying a broader range of relations that support scalar implicature, as I have done in Chapter 5, makes it even more difficult to determine whether or not some expression should be seen as evoking a salient relation -- i.e., whether the identification of some relation is necessary for the interpretation of the utterance -- by expanding the class of expressions which may evoke such orderings. In effect, any expression may license a scalar implicature, as suggested by 227.¹³⁵

¹³⁴While it seems likely that, if relations are salient in the discourse, the items they order should also be salient. However, this is an open question.

¹³⁵Lexical items corresponding to salient expressions are italicized.

- (227) A: Did the girl in the red dress spill a diet coke?
 a. B: She spilled a diet *pepsi*.
 b. B: She spilled a *regular* coke.
 c. B: She spilled *a glass of tomato juice*.
 d. B: *Jane* spilled a diet coke.
 e. B: The girl in the red *slacks* spilled a diet coke.
 f. B: The girl in the *green* dress spilled a diet coke.
 g. B: The girl in *the green slacks* spilled a diet coke.
 h. B: The *boy* in the red dress spilled a diet coke.
 i. B: The girl in the red dress *will spill* a diet coke.
 j. B: The girl in the red dress *drank* a diet coke.
 k. B: The girl in the red dress *spilled the* diet coke.

How salient expressions may be identified in some utterance is of course a hard problem, since salience in general is far from understood. Potential solutions to the problem at hand may lie in ongoing work in linguistic pragmatics and in natural-language processing on sentence TOPIC (or LOCAL FOCUS or CENTERING in the AI literature),¹³⁶ which may be signalled by syntactic, prosodic, or certain semantico-pragmatic cues. Among these, the best understood are syntactic and prosodic markings.

6.3.1.1. Syntactic Cues

Syntactic marking is commonly seen as way of focussing an item. Topicalizations and other preposings ((228a)), it-clefts ((228b)), pseudo-clefts ((228c)), and there-insertion ((228d)) may serve to mark items syntactically.¹³⁷

- (228)
 a. A cat she's going to buy.
 b. It's a cat that she's going to buy.
 c. What she's going to buy is a cat.
 d. There's a cat she's going to buy.

If focus is an acceptable indicator of salience, then such clues may be included among the methods of identifying salient expressions for the purposes of interpreting scalar implicatures. However, syntactic marking too can accomplish only part of the task: In general, only one item can be marked in an utterance by syntactic means and not all salient expressions need be syntactically marked.

¹³⁶Below, I will use 'focus' to denote this general notion of *S*'s centering attention on some particular aspect of the discourse. The item focussed on serves as an index into the discourse structure, as well as an indication of which items inferences -- such as scalar implicatures -- are drawn about.

¹³⁷Among other things they do.

6.3.1.2. Intonational Cues

It has long been noted that intonation can serve to mark sentence FOCUS or TOPIC. How prosodic marking is actually interpreted has, however, been little studied. Recent work on so-called CONTRASTIVE STRESS [Culicover 83] and on FALL-RISE intonation [Ward 85a] suggests how these two prosodic features mark expressions as salient. For each, an expression including the semantic representation of the lexical item which contains the syllable receiving SENTENCE STRESS¹³⁸ would be a salient expression in our terms.

6.3.1.2.1. FALL-RISE Intonation

FALL-RISE intonation (See Note 61.) also helps to limit the range of salient expressions in a sentence uttered with this contour. Elsewhere [Ward 85a], Gregory Ward and I have proposed a procedure for identifying this range for a particular utterance made with FALL-RISE, which I will summarize below. Once this range has been identified, we can limit the set of possible posets from which *S* has chosen a salient ordering to those posets in which one of these potentially salient expressions appears.

In sentences uttered with the FALL-RISE contour, the syllable SYL receiving sentence prominence may mark a number of expressions as potentially salient. Consider the following example (recall that \SYL/ indicates SYL):

(229)

- a. A: Did Liberals in southern Oregon vote for Hart?
 - b. A: Did Liberals in the West vote for Hart?
 - c. A: Did Liberals anywhere vote for Hart?
 - d. A: Did anybody vote for Hart?
- B: Liberals in southern Cali\for/nia voted for Hart.

In B's response in 229, SYL is contained in *California*, whose representation must therefore be included in any salient expression e_i . But only in B's response to (229a) does it seem that *California* constitutes e_i , where an ordering defined over the set of U.S. states appears to be evoked. If B's response is to (229b), however, then it appears that we should look at the expression denoting *southern California* to identify e_i , and that *parts-of-the-western-U.S.* is an appropriate ordering here. In response to (229c), *in southern California* appears to represent e_j in an ordering of locations of U.S. voters. Finally, in response to (229d), e_j represents a value in the ordering of U.S. voters, *liberals in southern California*. So, the range of of potentially salient expressions includes at least these four. Note, however, that 230, where SYL is contained in *liberals*,

(230) B: \Lib/erals in southern California voted for Hart.

¹³⁸The most prominent syllable in a sentence.

cannot be a felicitous response to any of the queries in 229. Similarly, in 231, where SYL falls in *southern*, the only possible salient expression will in fact be *southern* itself, whatever

(231) B: Liberals in \south/ern California voted for Hart.

the context.

These observations can be accounted for by drawing upon Liberman and Prince's [Liberman 77] relational theory of stress, in which stress is analyzed as a binary feature with the values WEAK and STRONG. For example, Liberman and Prince's rules would assign the values STRONG (S) and WEAK (W) to the sentence *Liberals in southern California voted for Hart* (uttered with neutral intonation) as in Figure 6-1, where R labels the root of the tree.

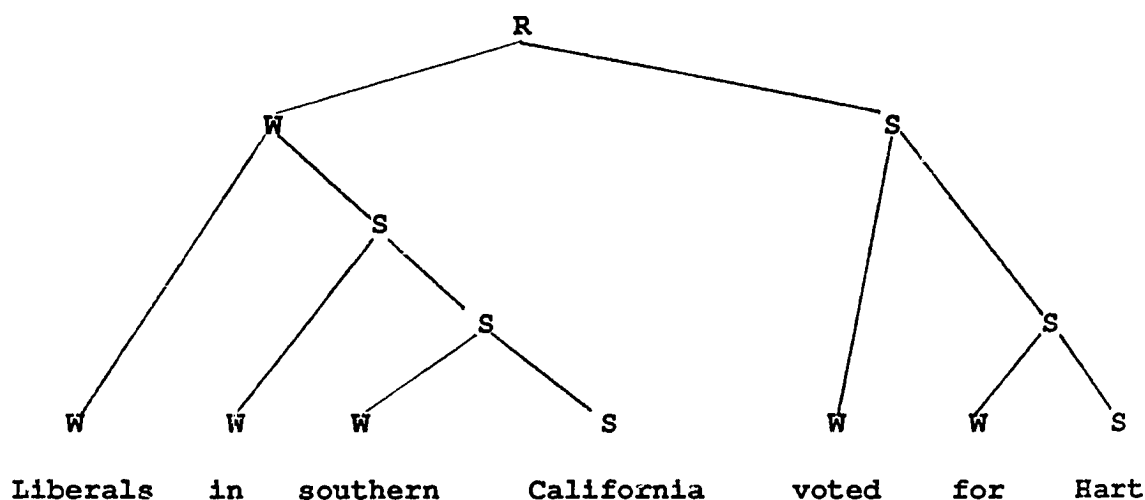


Figure 6-1: Metrical Stress Tree

Intuitively, when SYL occurs in a lexical item LEX which is dominated by a weak node in the prosodic tree for the corresponding utterance with unmarked intonation, only the expression including the representation of LEX can be marked as salient by FALL-RISE. Otherwise, larger expressions representing strings dominated by nodes in the prosodic tree dominating more than just LEX will also be candidates for expression salience.

Assuming a mapping between the semantic representation of an utterance u_i and its prosodic tree via the lexical items represented in each, this structure can be used to determine the set of potentially salient expressions E from such a prosodic tree as follows:

1. For any utterance u_i uttered with FALL-RISE, identify SYL and the lexical item it is included in, LEX.

2. Find the prosodic tree for u_i with unmarked/ neutral intonation.
3. Traversing the shortest path from LEX to R in this tree, for each Node $_i$ in this path (where i indexes the nodes from LEX to R):
 - A. If Node $_i$ is weak (W), add the expression representing the leaves of the subtree rooted at Node $_i$ to E and end the traversal;
 - B. If Node $_i$ is strong (O), add the expression representing the leaves of the subtree rooted at Node $_i$ to E and test Node $_{i+1}$.
4. Each of the expressions in E yields a possible salient expression for u_i .

Since *liberals* is immediately dominated by a W node in the tree above, in 230, where LEX is *liberals*, the traversal defined above will end with the testing of this first node on the path to R; thus only the subtree rooted at this node will be in E . In 231, only *southern* will be in E . However, in 229, no W node is encountered in the traversal until the trees represented by *California*, *southern California*, *in southern California*, and *liberals in southern California* have been added to E . So, four nodes are added to E for this utterance, as discussed above. By using an independently motivated relational theory of stress, we can provide a straightforward account of the range of potentially salient expressions for an utterance uttered with FALL-RISE -- if prosodic information is available.

6.3.1.2.2. Contrastive Stress

In a study of the indirect speech acts performed by yes-no questions, Kiefer [Kiefer 80] proposes that certain yes-no questions which he terms FOCUSED actually function as *wh*-questions.¹³⁹ Focussed yes-no questions for Kiefer are simply yes-no questions in which one constituent is marked by stress to indicate a BACKGROUND ASSUMPTION which S and H typically share.

(232a)-(232c) are focussed yes-no questions, with a different expression focussed in each, while (232d) contains no focussed element.

(232)

- a. Is John leaving for Stockholm TOMORROW?
- b. Is John leaving for STOCKHOLM tomorrow?
- c. Is JOHN leaving for Stockholm tomorrow?
- d. Is John leaving for Stockholm tomorrow?

While any of the four may be answered with *yes* or *no*, it is also possible that if S asks (232a) she wants H to answer the question 'When is John leaving for Stockholm?'; if she asks (232b) she may want to know 'Where is John going tomorrow?'; and if she asks (232c) she may want to know 'Who is leaving for Stockholm tomorrow?' To obtain these *wh*-questions, Kiefer

¹³⁹Similarly, [Hoepelman 83] claims that stress can indicate which element of a sentence is being queried.

constructs an open proposition from each focussed yes-no question by substituting a pro-word for the focussed element; this represents the query's background assumption. The questioner's request, then, is for an instantiation of the variable in this open proposition.

Although Kiefer does not recognize it, not all instantiations of his open proposition will be equally felicitous. (233a) is an odd

- (233) A: Is John leaving for Stockholm on Thursday?
 a. B: He's leaving for Stockholm on the train.
 b. B: He's leaving for good on Thursday.

response to (232a), and (233b), while a possible response to (232b), is a bit strange. However, since his variables are not typed, his theory will not preclude such infelicitous responses.

Kiefer also errs in claiming that only FOCUSED yes-no questions can elicit such indirect responses; nor does he recognize that more than one item in the query can be treated as queried. For example, all of the responses in 234 are possible responses to (232d).

- (234)
 a. B: Henry is.
 b. B: He's leaving for Stockholm on Wednesday.
 c. B: He's leaving for Paris tomorrow.
 d. B: He's leaving for Paris on Wednesday.
 e. B: Henry's leaving for Paris on Wednesday.

Finally, Kiefer's observations cannot account for the difference between responses that are appropriate to (235a) and those appropriate to (235b).

- (235)
 a. Is John leaving for the North POLE tomorrow?
 b. Is John leaving for the NORTH Pole tomorrow?
 c. He's leaving for Greenland.
 d. He's leaving for the South Pole.

While both (235c) and (235d) might be appropriate responses to (235a), only (235d) is an appropriate response to (235b).¹⁴⁰ As with FALL-RISE, contrastive stress on *pole* in (235a) may in fact focus several expressions, including *the North Pole* and *pole*; but stress on *north* can focus only *north* in (235b).

Culicover and Rochemont's [Culicover 83] account of contrastive stress does accommodate the focussing of more than one expression via stress in one lexical item. They claim that the sentence FOCUS marked by contrastive stress can be determined directly from

¹⁴⁰That is, without some accompanying 'He's not going to either pole but...'

prosodic trees. However, their analysis would predict that any constituent that dominates the item containing the prominent syllable defines a possible focus constituent (1983:148). For contrastive stress, as for FALL-RISE, this account allows too broad a focal range. For example, Culicover and Rochemont's analysis would predict that contrastive stress on *in* in ix could focus the entire expression *liberals in*

- (ix) A: Did Conservatives vote for Hart?
 B: #Liberals IN California voted for Hart.

California, and thus make it an appropriate response to A's query.

It seems likely that an algorithm similar to that presented in Section 6.3.1.2.1 can identify the range of possibly salient expressions marked by contrastive stress as well. Although such information would thus be quite useful for the calculation of scalar implicatures, it will only occasionally be available even in naturally occurring discourse. Also, intonational cues that can identify more than one salient expression in an utterance have not been studied. And, of course, computational recognition or generation of such cues is not yet possible.

6.3.1.3. Semantic and Pragmatic Cues

The determination of sentence focus from non-prosodic semantic and pragmatic cues has been proposed in studies of given/ new information status, discourse coherence, thematic roles, definiteness and indefiniteness, reference and anaphora, and the goals of conversational participants.

6.3.1.3.1. Given and New Information

In distinguishing between GIVEN and NEW information,¹⁴¹ Chafe [Chafe 76] observes that mention of some class member can bring about the 'givenness' of the class. For example, in his 'bulldog' examples (reproduced in 236) the initial utterance might be followed by (236a), (236b), or (236c).

- (236) Yesterday I had my class disrupted by a BULLDOG.
 a. I'm beginning to DISLIKE bulldogs. (= [Chafe 76]'s 16b)
 b. I'm beginning to dislike dogs. (= [Chafe 76]'s 20b)
 c. I'm beginning to dislike animals. (= [Chafe 76]'s 22)

Chafe observes that, the higher the subsequent generic term in its classification hierarchy -- or, actually, the more distant from the initially mentioned item (e.g., *bulldog*, above), the more difficult it is for the generic to be treated as given (and deaccented, for example). That is, *animals* is less likely to be treated as 'in *H*'s consciousness' if *S* utters (236c) than would

¹⁴¹Classically, given information is information which can be assumed to be in *H*'s consciousness, through prior mention or some other means. It is often associated with linear position in an utterance; that is, given information is generally thought to come before new information.

bulldogs in (236a) or *dogs* in (236b); thus, moving up a hierarchy diminishes the implicit givenness that can be transferred. Likewise, moving down the taxonomic hierarchy does not produce a similar sense of givenness for *bulldog* in 237.

(237) I'm beginning to dislike DOGS. (= [Chafe 76]'s 23)
Yesterday I had my class disrupted by a BULLDOG.

Mention of an item higher in a classification hierarchy will not impart 'givenness' to lower items.

Chafe [Chafe 76:130] accounts for this phenomenon in terms of *S*'s knowledge and awareness:

Whenever a speaker's knowledge is such that, for him, consciousness of X necessarily entails consciousness of Y, he will assume that the addressee's consciousness of X entails consciousness of Y also.

He notes that other orderings can impart a similar *S* assumption of mutual consciousness, as the whole/ part ordering evoked in 238.

(238) Yesterday my FANBELT broke. (= [Chafe 76]'s 25)
I couldn't use the CAR all day.

Mention of *fanbelt* allows car to be treated as given.

If givenness can indeed be associated with salience, as has been claimed [Bock 77], then Chafe's observations might help identify salient expressions for the purpose of calculating scalar implicature. If mention of FANBELT can make *car* salient for a subsequent utterance, then presumably *S* and *H* can anticipate and identify this possibility. However, Chafe's notion fails to allow us to predict when items such as *bulldog* or *fanbelt* will **not** represent salient expressions -- i.e., when items which may be seen in some classification hierarchy or some part/ whole relationship will not evoke these orderings. So, although these observations will prove useful below in identifying salient orderings for some salient expression, they do little to help us select among potentially salient expressions.

6.3.1.3.2. Centering and Discourse Coherence

Studies of discourse CENTERING seek to relate the coherence of a discourse to conceptual and positional relationships among items and references to them in a discourse and assumptions that can be made about *H*'s awareness of those relationships. Work by Joshi, Weinstein, and Grosz [Joshi 81, Grosz 83] suggests a simple approach to the identification of salient expressions which does work for a restricted class of utterances.

Joshi et al. claim that a discourse will be more COHERENT when the BACKWARD-LOOKING CENTER of an utterance is either a member of the set of FORWARD-LOOKING CENTERS of the preceding utterance or is functionally dependent upon one such forward-looking center.

Forward-looking centers represent potential foci for subsequent utterances, and are defined tentatively in this work to include the set of noun-phrases in the current utterance.¹⁴² Since 'functional dependence' is both difficult to define and intuitively too narrow, Ward [Ward 85b] extends centering to encompass backward-looking centers which appear in partial orderings with some forward-looking center.

So, if coherent discourse is a goal of *S* and *H*, as is generally assumed, presumably we might restrict the range of possibly salient expressions for u_j to the forward-looking centers of its predecessor utterance u_i or to a set of preceding utterances in C_h . For utterances which license scalar implicature, salient expressions would have to represent a center that is functionally dependent upon a forward-looking center, or, following [Ward 85b], an expression in some partial ordering with an expression in the set of forward-looking centers. So, the location of sentence centers would seem to represent a method of locating our salient expressions.

[Joshi 81, Grosz 83]'s sequential/ conceptual relating of items in discourse does suggest a simple solution to the problem -- at least for certain types of exchanges. In any of the exchanges in 227 (repeated here for convenience),

- (1) A: Did the girl in the red dress spill a diet coke?
 - a. B: She spilled a diet *pepsi*.
 - b. B: She spilled a *regular* coke.
 - c. B: She spilled *a glass of tomato juice*.
 - d. B: *Jane* spilled a diet coke.
 - e. B: The girl in the red *slacks* spilled a diet coke.
 - f. B: The girl in the *green* dress spilled a diet coke.
 - g. B: The girl in *the green slacks* spilled a diet coke.
 - h. B: The *boy* in the red dress spilled a diet coke.
 - i. B: The girl in the red dress *will spill* a diet coke.
 - j. B: The girl in the red dress *drank* a diet coke.
 - k. B: The girl in the red dress spilled *the* diet coke.

for example, we might compare B's response with A's query to locate salient expressions as follows: Let p_i be the semantic representation of the DESIDERATUM¹⁴³ of A's query and p_j be the representation of the respondent's answer. Then, if p_i and p_j represent simple expression alternates -- differing only in expressions e_i and e_j -- and if e_i and e_j appear together in some poset O , we can term $e_j p_j$'s backward-looking center, related to the forward-looking center, e_i . And, e_j will be a salient expression of p_j . For A's query in 227 and B's response of (227a), for

¹⁴²It is unclear whether they include noun-phrases from prior utterances as well.

¹⁴³A specification of the epistemic state a questioner desires. Although centering is currently defined only for declaratives, this strategy will accomplish the desired transformation.

example, *pepsi* would represent a salient expression in this exchange and *set of sodas* might represent a potentially salient ordering, since *pepsi* and *coke* might be characterized as members of such a set. In this way it is possible to calculate expressions which B has taken to be salient -- and, thus, to calculate those scalar implicatures B may be licensing.

When more than one prior reference has been made to distinct values on some ordering, the selection of a salient ordering for a particular value reference should be even simpler. The more points specified on some ordering, the easier it is to identify it. For example, in 239, A's first query helps *H* to select among

- (239) A: How many women in the company are managers?
 B: Thirteen.
 A: Do women employees take more days of sick-leave than male employees?
 B: Non-management women do.

possible orderings on which *women employees* might appear to formulate a cooperative response to A's second query.

In fact, it does appear that a simple expression alternate condition alone is sufficient to identify salient expressions. Although the responses in 240 do appear

- (240) A: Did the girl in the red dress spill a diet coke?
 a. B: She spilled the beans.
 b. B: The girl in the red car did.

intuitively less 'coherent' than the responses in 227, to make any sense of them A must postulate some set of *diet coke* and *beans* or *dress* and *car* -- and $\neg\text{BEL}(B, \text{the girl in the red dress spilled a diet coke})$ does seem to be a reasonable inference from both. However, when these values appear in more than one poset, other means must be used to identify which such poset is salient (See Section 6.3.2.3.).

Of course, the concept of centering which Joshi et al describe is so far defined only for entities, while salient expressions may refer to things other than entities, as noted in Section 5.1. Still, Joshi et al clearly intend some extensions to other items.

More serious for our purposes is that a simple expression alternate p_i immediately preceding p_j in the discourse will not always be easily discernible to permit the location of a salient expression in p_j , since scalar implicatures may be implicated when such a p_i is not the immediate predecessor of p_j or when no such p_i exists. In 241, for example, p_i is separated from p_j by several intervening sentences. While the salience of *apples and pears* might be inferred by simple comparison of '*like(B, fruit)*' with '*like(B, {apples, pears})*', how one locates the former

- (241) A: Do you like fruit?
 B: Why, do you have some around?
 A: I can get some.
 B: Weil, I like apples and pears a lot.

is not clear. So, the licensed implicature $\neg\text{BEL}(B, I \text{ like fruit other than apples and pears})$ cannot be calculated simply by matching semantic representations of contiguous utterances. And for other utterances -- say 242 --

- (242) Some people left early.

no p_i at all may be available for comparison. So, this procedure for locating salient expressions is of limited use.

Finally, although Joshi et al define discourse coherence in terms of forward- and backward-looking centers, they provide few insights on how to identify these centers (beyond the *post facto* 'what is pronominalizable') and no strategies for ordering forward-looking centers once they are found. So, although centering will allow A to identify salient expressions in B's response in 227, for example, it will not help B to choose the most salient expression from the forward-looking centers in A's query, '{*the girl in the red dress, the girl, girl, the red dress, ..., a diet coke ...*}'. And, too, postulation of a single backward-looking center will not help us to identify scalar implicatures arising from the salience of more than one expression in an utterance.

6.3.1.3.3. Local Focus

In her computational account of LOCAL FOCUS, Sidner [Sidner 79, Sidner 83] presents a strategy for tracking the focussing process throughout a discourse and using information thus obtained to resolve anaphoric reference. Using syntactic features and thematic (case) roles of utterances, together with data provided by subsequent anaphoric reference, Sidner provides algorithms to hypothesize discourse and local foci, confirm or reject these hypotheses, and propose substitutions where necessary.

Noun-phrases in an initial utterance are preferred for the DEFAULT EXPECTED FOCUS LIST in the following order:¹⁴⁴

- the subject of a there-insertion clause or an 'isa' sentence;
- the THEME of the sentence, which Sidner defines as the verb relation that indicates the property of being affected by the action of the verb;¹⁴⁵
- the other case roles present in the sentence, with a weak preference for GOAL or for any role filled by an indefinite, with AGENT ordered last;

¹⁴⁴Sidner notes that prosody is not well enough understood to figure in these calculations.

¹⁴⁵Presumably, the *patient* in other thematic typologies.

- the sentence's VP, which may be specified by a nominalization of the verb.

An item selected by this means may be rejected when unacceptable coreference is predicted later in the discourse and another member of the default focus list substituted as discourse focus.

Sidner's theory accommodates FOCUS SHIFT during the discourse to any term in a subsequent utterance. Potential foci for such shifts may be anticipated on the same basis as the choice of the default focus list; actual shift is confirmed or rejected by subsequent anaphoric reference. Provision is also made for a return to previous foci -- FOCUS POPPING -- in the discourse, through the maintenance of a FOCUS STACK. Potential referents are collected in a focus stack, created by ranking NP's from most to least likely to represent the 'focus of attention' in the discourse; This stack will be searched after the discourse focus and potential focus list.

Sidner's restriction of possible foci to noun- and verb-phrases, like [Joshi 81, Grosz 83]'s restriction of centering to noun-phrases, clearly represents a serious problem for our purposes. Since any expression in a sentence can represent a salient expression for the purpose of supporting scalar implicature, any constituent's semantic representation may serve as a salient expression. The commonly accepted notion that contrastive stress may mark focus provides independent support that any item can be focussed, as in 243, where *UNDER* seems obviously to be the item focussed.

(243) I left the book UNDER the table.

Similarly, in 229, identifying the sentence's focus with the NP

(244) I left the RED book under the table.

the red book does not capture the intuitive notion that it is not the entity but one of its attributes that is being attended to. Exclusion of all but NP's and *vp*'s from consideration as foci thus limits the use of existing algorithms for the definition of salient expressions.

6.3.1.4. Communicative Dynamism

Since several expressions may be salient in a given sentence, an ideal focus algorithm for our purposes would also provide us with information about the degree to which items may be 'in focus' in a discourse. Work in COMMUNICATIVE DYNAMISM (CD) by the Prague School has long studied the 'distribution of various degrees of communicative dynamism (CD) over the elements within the sentence' [Firbas 65:170]. The degree of CD for any element is determined by the extent to that element contributes to the development of the communication and is related to the amount of 'givenness' of the element. For example, the lowest degree of CD of a sentence is associated with its 'theme' and the highest with its 'rheme'. Other elements are

associated with communicative ‘transition’, and carry an intermediate degree of CD. Again, this research has far to go before it can predict reliable orderings of focussed elements.

6.3.1.5. Defining Salient Expressions Using Focus

Limited though our current understanding of focus may be, research on this topic holds out hope of an (eventual) independent source of the salience information required to calculate scalar implicatures. Although I will neither propose a new theory of focus assignment nor commit myself to an existing focus algorithm, I will assume in the remainder of this work that some such algorithm is available to rank expected foci. In this section I suggest how focus information might be used to predict salient expressions for the generation and interpretation of scalar implicature.

Assuming the availability of information about can we make of knowledge about the degree of speaker/ hearer attention each item is receiving in the discourse for the purpose of anticipating or interpreting scalar implicatures? An obvious approach is to say that the more in focus some item is the more likely it is figure in a scalar implicature. Say *S* wishes to convey a scalar implicature via evocation of some ordering *O*. Then, if *S* believes that s/he and *H* mutually believe that some e_i in *O* denotes an item that is already salient in the discourse, *S* will anticipate that his/her use of a sentence containing e_i to license a scalar implicature will be likely to be correctly interpreted by *H*. The justification for this claim is as follows: If some item v_i represented by e_i is salient in the discourse, it seems plausible to assume that items it is related to will also be salient, along with the metric by which these items are related to v_i . Kiefer’s account of focussed yes-no questions supports this strategy. In effect, the focussed item is the item most likely to be mentioned in a subsequent utterance or ‘replaced’ with another value in a salient ordering. Of course, when an item is linked to a focussed item by functional dependence, a poset relation [Ward 85b], or some other relationship [Grosz 77], most studies count this a form of focus maintenance.

An alternate strategy might also seem plausible: that *S* seeks to maintain discourse coherence by continuing to mention focussed items and by ‘replacing’ items not in focus by values in some ordering. However, this hypothesis seems intuitively less satisfying than the approach I have taken. Consider the oddness of B’s response in (245a), when A has marked *George* as

- (245) A: Did GEORGE burn the cake?
 a. B: He burned the cookies.
 b. B: Harry did.

in focus by stress.¹⁴⁶ (245b), on the other hand, seems much more appropriate. So, focussed items do seem to be salient items for the purposes of licensing or interpreting scalar implicature: That is, when *S* licenses some scalar implicature, that scalar implicature will either draw upon the salient of some item v_i already focussed in the discourse by mentioning some value in an ordering in which v_i appears, or *S* will focus attention upon some v_j in his/ her utterance which s/he intends to license a scalar implicature.

A look at some naturally occurring discourse illustrates this strategy: In 140 (repeated here for convenience, with stressed items are in upper case), A first queries a lower value in an ordering

- (1) B: I as a result of my extensive investigations and especially on an international basis -- I was involved in industrial and governmental espionage so I had to be very careful about --
 A: CIA stuff?
 B: Pardon me?
 A: CIA stuff?
 B: We (PAST) HAVE worked WITH them.
 A: I see.
 B: I (PRES) don't work FOR them.
 A: I see.
 B: Uh I (PRES) work FOR my corporation.
 A: Uh huh.
 B: I (PAST) have also worked with security organizations, the FBI --
 A: Uh huh.
 B: And the intelligence groups. But what happened was I had to remain silent about my inheritance because uh we were threatened.

of *industrial and governmental espionage employers*. B first evokes tense and VP orderings, *past/ present* and *work with/ work for* by stressing them, to clarify her relationship with the focussed item *CIA*, while maintaining *CIA* as the topic of the discourse. She implicates \neg BEL(B, *We now work for the CIA*), as she states in the next utterance. B then refers to A's ordering of *industrial and governmental espionage employers* by mentioning *my corporation*. In the same utterance, however, she also refers to the previously evoked tense and prepositional orderings. In her next utterance, B refers to other values in the same three orderings, mentioning tense, *work with*, and *security organizations* in general and the *FBI* in particular. In this utterance she implicates \neg BEL(B, *I now work for security organizations*) and \neg BEL(B, *I now work for the FBI*).

Without a more reliable independent test of focus, it is impossible to test this association

¹⁴⁶I use stress here to mark focus since it seems the least ambiguous method.

between focus and salient expressions more objectively. Tentatively, however, I will propose that, given any focus algorithm, the ordering of salient expressions for the purpose of inferring or anticipating the inference of scalar implicatures will be the order of focussed items -- from most to least in focus. However, given an ordering of possible foci, we must then determine the point at which items become 'too little in focus' to be counted as salient. This remains an open question.

Clearly, however, if some expression e_i is marked as salient by some focus algorithm but if S cannot discern or does not believe that H can discern any ordering in which that expression occurs, S will not make reference to e_i to license a scalar implicature. Alternatively, if more than one such ordering in which e_i appears exists for S -- and may be assumed discernible by H -- S will need to choose that ordering which H is most likely to discern to ensure that his/her implicature is properly understood -- i.e., S must select the most salient ordering in which e_i appears.

6.3.2. Identifying Salient Orderings

Although the salience of relations in discourse has been less studied than the salience of entities and attributes, [Grosz and Sidner 85, Conklin and McDonald 82] there has been considerable discussion of those relations which may hold among propositions in the discourse to account for discourse structure. [Hobbs ??, McKeown 82, Mann and Thompson 83, Mann 84] However, these authors have been more concerned with identifying the range of possible relations between discourse units than with how speakers select from this range. Even if salient expressions can be identified in the discourse, knowledge of which relations are also salient will be needed to select among possible orderings for salient expressions. In addition, where independent information about salient expressions is ambiguous or unavailable, salient orderings may themselves point to salient expressions: That is, if an ordering O_i is salient in the discourse and if some expression e_i appears both in the current sentence and in O_i , the salience of O_i might indicate the salience of e_i . Thus, the salience information necessary for the generation and interpretation of scalar implicature might be acquired by looking first for a salient expression or by first locating a salient poset.

It should be evident from examples such as 218 and 219, 222, 225, and other exchanges presented above that a given expression may evoke a number of different orderings. A cooperative S should seek not only an ordering that can be identified by H -- and defined as S defines it -- but the most salient such ordering for the particular exchange. Of course, S may instead attempt to make an ordering salient for H which s/he does not believe is already salient for H -- by evoking it via mention of some value on it -- but only if s/he thinks H will recognize both the ordering and the process by which s/he has evoked it.

While the need for *S* to reason about which ordering *H* may perceive as salient for a given expression is easiest to see where domain-dependent orderings are involved, even expressions used to evoke the canonical, domain-independent orderings like the modals may instead evoke some other ordering, as in the following hypothetical exchanges between a hospital visitor and patient in 246:

- (246) A: Can you get out of bed?
 a. B: I'm supposed to at least once a day.
 b. B: I can sit up.

Since *can* may have either a modal or an 'ability' reading, it may evoke a deontic ordering ((246a)) or an ability ordering ((246b)).

Tentative and partial answers to the question of how one locates the most salient relation for a particular expression in a particular sentence may nonetheless be sought in this domain independent/ dependent distinction as well in information about which orderings have been salient previously in the discourse, in cognitive science studies of human perception of relations, and in clues that particular metrics provide as to the salience of relations they order.

6.3.2.1. Selecting Salient Classification Hierarchies

To date, the major studies of human choice among potentially salient ordering relations have focussed upon children's ability to classify. [Inhelder and Piaget 64, Rosch 76, Markman 83, Gelman and Baillargeon 83, Sugarman 83, Waxman 85] That is, they have been mainly concerned with identifying the development of children's ability to classify taxonomically as opposed to imposing thematic or idiosyncratic organizations upon groups of items. However, one segment of literature on adult classification does suggest how to identify salient classification hierarchies when some such hierarchy is found to be salient.

Scalar implicatures licensed via a classification ordering present peculiar problems for the selection of a salient ordering, since classification hierarchies are potentially infinite. That is, we might represent the following classification of pets finitely, as below:

lilac point blue point

Siamese

Burmese

cat

dog

pet

But, for any classification hierarchy, the most specific levels in that hierarchy may always be

further specified, even if that specification is somewhat idiosyncratic, as say, *round-faced lilac points*, *round-faced lilac points with double paws*, and so on.

While it is clear that B may license the implicature $\neg\text{BEL}(\text{B}, I \text{ have a Siamese})$ by his/her response in 247, it is not clear that we want to say all the (potentially infinite) higher values in this and similar hierarchies will also be the subject of implicatures via this utterance -- e.g., $\neg\text{BEL}(\text{B}, I \text{ have a lilac-point Siamese})$ and $\neg\text{BEL}(\text{B}, I \text{ have a round-faced lilac-point Siamese})$.

(247) A: Do you have a Siamese?
B: I have a cat.

And, in 248, we would not want to say that B's response licenses

(248) A: Do you have a pet?
B: I have a cat.

the implicature $\neg\text{BEL}(\text{B}, I \text{ have a Siamese})$, although *Siamese* is a higher value in the hierarchy than the mentioned *cat* and the ordering is clearly salient in this exchange. Also, in 249, B's response may indeed license $\neg\text{BEL}(\text{B}, I \text{ have a Siamese})$, just as,

(249) A: Do you have a Siamese?
B: I have a pet.

in 250, B's affirmation of *pet* may implicate $\neg\text{BEL}(\text{B}, I \text{ have a cat})$.

(250) A: Do you have a cat?
B: I have a pet.

But it seems odd to say that B may implicate $\neg\text{BEL}(\text{B}, I \text{ have a lilac point Siamese})$ in either 249 or 250. How then do we explain the limits on scalar implicature licensed in 247-250? A classification hierarchy is clearly salient in these, but it does not appear to be the same in each. And for none of these responses would we want to say that the salient ordering is an infinite one -- but how can we bound such orderings in a principled way?

6.3.2.1.1. Basic and Entry Levels Defined

Students of human classification have long debated whether or not certain levels of abstraction in classification hierarchies have particular psychological significance. BASIC categories have been defined by Rosch [Rosch 73, Rosch 76] as those categories in classification hierarchies which carry the most information; are most differentiated from other members of the hierarchy at their level; are the most inclusive categories whose members have a significant number of attributes in common; have similar motor programs and similar shapes; and can be most easily identified from the averaged shapes of members of the class. While Rosch's first examples were of color and form, her findings were soon generalized to biological classifications and to hierarchies such as *food* and *furniture*.

Rosch was particularly interested in identifying prototypical members of these classes.

Empirical studies lead her in [Rosch 75] to adopt Wittgenstein's 'family resemblance' notion, that prototypical members of categories are those which possess the most attributes in common with other members of their category and the fewest in common with members of other categories. So, robins are prototypical birds because they exhibit a high proportion of those attributes common to all birds and fail to exhibit attributes which might link them with competing categories, such as *fish* or *reptile*. Bats are atypical exemplars of the category *bird*, since, while they have wings, they have fur instead of feathers. Thus they share important attributes of the class *mammal*.

In a standard classification hierarchy entities are assigned to one of four levels, ranked on the basis of decreasing generality, from SUPERORDINATE to INTERMEDIATE to BASIC to SUBORDINATE level. An example of such a hierarchy is:



Intermediate levels are often difficult to determine. The basic category *dog* in this hierarchy, like other basic categories, can be identified as 'basic' by a number of generally convergent tests.¹⁴⁷ For example, a basic category often corresponds to the highest-frequency item of the hierarchy in vocabulary counts -- as does *dog*. Also, the word for the basic category tends to be learned earlier than words expressing other levels. [Anglin 79] And this category is often the least abstract member of its hierarchy that is monomorphemically lexicalized, as is *dog* or *chair* in the hierarchy *furniture/ things to sit on/ chair/ armchair*. For some domains, such as color and shape, it has even been argued that this latter feature is universal; for example, few or no languages have a single word for 'a washed-out pinkish-red' or have polymorphemic 'red'. [Berlin 69]

Basic categories are also natural candidates for ostensive introduction: [Fodor 81] that is, one cannot teach *poodle* ostensively to child who has not already been taught *dog*, and similarly with *pale red* and *red*. Basic categories are typically the most abstract members of their hierarchies which differentiate individuals of approximately similar appearance [Rosch 76] While subjects can draw a dog, for example, they cannot draw an animal; similarly for *chair*

¹⁴⁷Where particular basic categories are identified in examples in this section, they have been verified by (others') empirical research along the lines presented below.

and *piece of furniture*. Objects belonging to same basic category have a greater degree of shape overlap (when considering canonical depictions) than do objects belonging to the same superordinate category but not to the same basic category; and objects belonging to the same basic category do not have significantly less overlap than do objects belonging to the same subordinate-level category.¹⁴⁸ There is also considerable evidence that perceptual identifications which involve access of basic categories are faster than those involving non-basic categories. [Intraub 81] For example, when subjects were asked to find a picture of a woman walking a dog they accomplish this task more rapidly than when they are asked to find a picture of a woman walking a poodle. Subjects can also name objects or match names with pictures faster at basic level [Brownell 78, Hutcheon 70, Rosch 76, Segui 68, Smith 78] These results suggest that basic categorizations may be made on the basis of the visual properties of objects.

Basic categories also yield INFORMATION PEAKS when subjects are asked to list properties of categories: They can list more properties of basic categories than of more abstract categories and very few more properties for more specific categories than for the basic level. Also, more subjects tend to list the same attributes of basic categories than they do of other levels. Subjects were also asked to describe motor behaviors that they would emit in the presence of specific objects; more common motor behaviors were listed for basic categories than for other level categories. It has been claimed that basic categories encode the most information per unit of judgment [Fodor 83]; thus, by Grice's maxims of Quantity and Manner, this observation predicts that, *ceteris paribus*, basic categories are the natural ones to use for describing things. [Fodor 83] claims that the basic level thus represents the preferred or default level of description even where more information is known -- unless special circumstances indicate otherwise. That is, basic categories are phenomenologically given, providing the natural level for describing things to oneself as well as to others.¹⁴⁹

While accepting the basic notion of a psychologically privileged level, [Jolicoeur 84] contends that a more sophisticated notion of ENTRY LEVEL is required to accommodate subjects' responses to some of the tests of this level when atypical members of basic categories

¹⁴⁸When subjects were presented with drawings created by averaging the outlines of two shapes and asked to name the category of the object, they could perform the task when shapes were drawn from two members of a basic category or from two members of a subordinate category, but had great difficulty when shapes of different basic categories belonging to a given superordinate were averaged; e.g., it was hard to identify the average of apple and banana as fruit.

¹⁴⁹Of course, the basic level for particular hierarchies may be influenced by culture and individual experience. So, if I have grown up in a fishing village *fish* -- generally considered an intermediate category with *mammal* and *reptile* -- may be a basic category for me.

are involved. Their empirical studies reveal that, for atypical members of a basic category (as for *chicken* an atypical member of the basic category *bird*) the atypical subordinate (e.g., *chicken*) will pass the tests for ‘basic level’ discussed above. That is, the ‘entry’ level to a classification hierarchy will be a function both of an item’s position in the classification hierarchy and of its typicality at that level. Atypical members of a basic category tend to have their privileged or entry level at the subordinate level, as in (251a), while typical members have

- (251) A: Did you see any birds?
 a. I saw a chicken.
 b. I saw a robin.

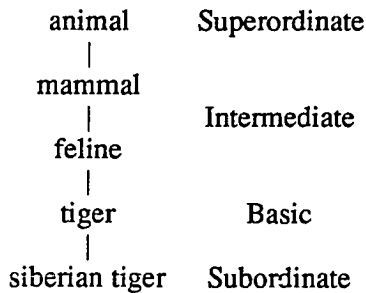
their entry level at the basic level, as in (251b). That is, while accepting the notion of a privileged level in classification hierarchies, [Jolicoeur 84] find that, for atypical members of a basic category, the diagnostics for basic level such as those described above are actually satisfied more often at the subordinate level.

6.3.2.1.2. Bounding Hierarchies by Entry Level

Now, returning to the original problems identified in 247-250, I propose to account for the apparent bounding of the classification hierarchies salient in these exchanges in terms of a hierarchy’s entry level. In the pet hierarchy illustrated above, *pet* will represent the superordinate level, *cat* the basic level, and *Siamese* the subordinate level, with *lilac* point a member of some sub-subordinate level. *Siamese* is not an atypical member of the type *cat*, so the entry level for this hierarchy should be its basic level, *cat*.

Note that, while items more specific than the entry level for the pet hierarchy may indeed be mentioned to license scalar implicature, where all other things are equal the entry level will pose a default bound upon licensed implicatures. So, when a classification metric is salient, we can normally take the resultant salient ordering to be bounded from (above in my representation but below in standard classification hierarchies) by the hierarchy’s entry level. So, in 248, for example, B’s response will not license $\neg\text{BEL}(B, I \text{ have a } Siamese)$ because this level will not be a part of the exchange’s salient classification hierarchy; that is, while B’s affirmation of *cat* will license his/her lack of commitment of higher values, *Siamese* will not be a higher value in a ‘pet’ classification hierarchy bounded (for *Siamese*) by the basic level *cat*. Of course, if some more specific level of the hierarchy is mentioned, a larger ordering must of course be taken as salient. In 247 and 249, thus, the salient ordering will include the subordinate level. However, more specific levels will be added to the salient hierarchy only if they, or still more specific levels, have been explicitly evoked by mention of a member -- as, *Siamese*. So, in such cases, scalar implicatures will be licensed only up to and including the most specific evoked level of a classification hierarchy. B’s response in 247, for example, will license only the implicature $\neg\text{BEL}(B, I \text{ have a } Siamese)$. In effect, then, scalar implicatures will not be licensed ‘above’ (at a more specific level than) the basic level or above a more specific evoked level. Some further examples should illustrate these findings.

Say we propose a path in a classification hierarchy as follows:



where the levels are as indicated, with an intermediate level at either *feline* or *mammal*. Then, if B responds to A's query in 252 with (252a), s/he can implicate $\neg\text{BEL}(\text{B}, (252b))$, since *tiger* represents this hierarchy's basic level and is a higher value in a type/ subtype ordering than *feline*. Since *tiger* represents a

- (252) A: Was that an animal over there?
 a. B: It was something feline.
 b. B: It was a tiger.
 c. B: It was a Siberian tiger.

typical animal, then it seems reasonable that B has considered this and other typical animals but is not able to affirm some particular one of them. But, if B responds with (252b), it seems unlikely that she intends to implicate that the animal is not a *Siberian tiger* or some other more specific variety. That is, it seems that the entry level of a classification hierarchy establishes its 'appropriate level of detail' -- not only for simple description but for implicit meanings such as scalar implicature. Of course, B may indeed choose to respond at a more specific level than the entry level, as in (252c). Note, however, while such a response may alter the 'default' appropriate level of detail -- in this case to the subordinate level -- no scalar implicatures will be drawn beyond this mentioned level. That is, A will not infer that B cannot commit him/herself to some more specific description of the *Siberian tiger*, such as *blue-eyed Siberian tiger*.

I propose that a classification hierarchy's entry level will establish a default boundary beyond which scalar implicatures will not be drawn -- in effect, creating finite type/ subtype orderings from potentially infinite ones. While *S* may evoke a larger ordering by explicit mention of a more specific level, this mention will only serve to raise the default boundary to the mentioned level.

S may alter this default appropriate level of detail in (at least) one other way: Barcelu [Barcelu] notes that, when *S* says 253,

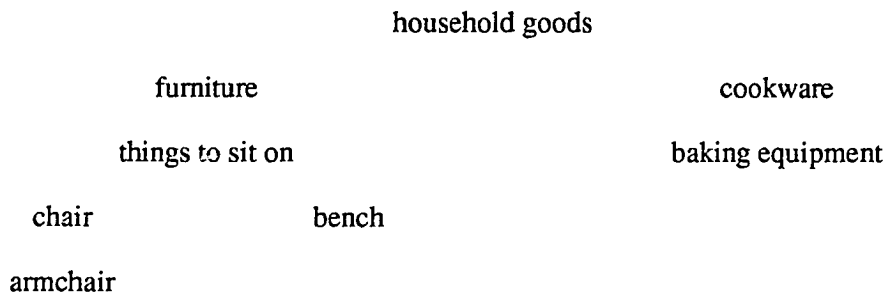
- (253) I'm thinking of something and it's not a dog.

H will infer that *S* is thinking of 'cat' say, over choices like 'parakeet' or 'hamburger'. We might account for this preference by noting that *cat* is intuitively 'at the same level' as *dog* in a classification hierarchy. Furthermore, we could interpret this licensed inference and those licensed by similar utterances as scalar implicatures: In denying a value in a classification hierarchy (or, type/ subtype ordering), *S* may license $\neg\text{BEL}(S, \text{---alternate values})$, by Imp_2 , where *dog* and *cat* are alternate values. This observation suggests another way by which *S* may modify the default salient level of a classification hierarchy. *Cat* and *dog* are both basic categories, and, thus, an exchange like 254 will license no implicatures about values more specific than *dog* in the pet

(254) A: Do you own a dog?
B: I have a cat.

hierarchy. However, when *S* mentions a value in a classification hierarchy which is alternate to a value already salient in the discourse -- i.e., a value at the same level in that hierarchy, even when that level is not the entry level for that path in the hierarchy, it appears that *S* may set the default level of the hierarchy at that level.

Consider, for example, the following hierarchy of household goods, from the superordinates *furniture* and *cookware* down to the subordinate *armchair*:



exchange in 255. While *cookware* is not at the basic level for

(255) A: Does Jane need furniture?
B: Cookware.

this hierarchy, its mention does not seem to license scalar implicatures about more specific (higher) members of the hierarchy, such as *baking equipment*. Yet, clearly, part of this hierarchy is salient, since B can implicate $\neg\text{BEL}(B, \text{Jane needs furniture})$ by the response. So, it seems that *S* may alter the default appropriate level of detail by reinforcing a more general level previously salient in the discourse. In 255, A has made the superordinate level salient by his/her mention of *furniture*, and B responds at the superordinate level with *cookware*. In such a case, the reinforced level will represent the salient level for the discourse, and the hierarchy will be bounded at that level.

In sum, there are at least two ways *S* may alter the default level of a classification hierarchy and thus bound these potentially infinite orderings for the interpretation of scalar implicatures. The default boundary for any classification hierarchy is the entry level for any path in the hierarchy. *S* may reset this default either by mentioning a more specific item than the entry level or by mentioning a less specific item at the same level (alternate to) as an item previously salient in the discourse. Whatever action *S* chooses to take, scalar implicatures will not be licensed beyond (higher or more specific than) this revised level.

These observations help determine how to define the portion of a classification hierarchy that will be salient in an exchange -- given that some classification hierarchy is salient. Chafe's examples (See Section 6.3.1.3.1.) suggest that classification hierarchies may generally be assumed salient, since mention of some value in them imparts 'givenness' to other values. That is, *S* may anticipate that, if *H* does not already view such orderings as salient, s/he will be able to recognize their evocation. So, *S* may anticipate that *H* will be able to work out implicatures which rely for their interpretation upon perception of a classification hierarchy (type/ subtype relationship).¹⁵⁰

6.3.2.1.3. Extending 'Entry Level' Beyond Classification

While no similar investigation of human perception of other ordering relations has been made, it does seem reasonable that other orderings too might have their 'appropriate level of detail': The general problem has been recognized in the literature. In discussing his example of clash between the maxims of Quantity and Quality (illustrated in 256, Grice observes that B's response is clearly not appropriately detailed, although it is as detailed as B can truthfully make it: Recall that A is planning his French holiday. B knows A wants to visit C.

- (256) A: Where does C live?
 a. B: Somewhere in the south of France.
 b. B: In Marseille.
 c. B: On Rue de Janvier in Marseille.
 d. B: In a tiny one-bedroom apartment in a three-story brick building on Rue de Janvier in Marseille.

By saying (256a), Grice claims that B implicates s/he does not know which town C lives in. For, if s/he did, s/he should have mentioned it, since A will need that information to locate C. The question is, if B had responded (256b), could she have implicated $\neg\text{BEL}(\text{B}, (256\text{c}))$? Or, if she had responded with (256c), could she have implicated $\neg\text{BEL}(\text{B}, (256\text{d}))$? So, again, we have potentially infinite hierarchies -- in this case defined in terms of specificity of information. However, it is less clear how to assign *a priori* to any such hierarchy a privileged

¹⁵⁰It is also interesting to note that, if licensable scalar implicatures can be associated with entry levels in the way described above, scalar implicature may provide yet another test for entry level in classification hierarchies.

level of informativeness. Consider how the 'appropriate level of detail' in 256 will change if, instead of planning to visit C, A is merely inquiring as to what has become of his/her old friend C. So, when generalization/ specialization relationships are salient, while it does seem both appropriate and even essential to assign some 'entry level' into the hierarchy, it seems less easy to assume that this level will be independent of context.

6.3.2.2. Salient Orderings from Speakers' Goals

Linguistic and AI research on speech acts, planning, and goal inference has convincingly argued that speakers' goals can determine what is relevant or salient in a discourse. [Morgan 78, Hobbs 79b, Allen 80, Pollack 86] For example, depending upon the goals of the exchanges in 257, B might choose to evoke one of several possible orderings.

- (257) A: Are the Yankees in town this week?
 a. B: Guidry is.
 b. B: The Mets are.
 c. B: They're in town next week.

Each of the responses in 257 will be appropriate in the corresponding contexts in 258.

- (258)
 a. A: I need to interview a pitcher about this new beanball rule. Are the Yankees in town this week?
 b. A: I want to take some clients to a game this afternoon. Are the Yankees in town this week?
 c. A: Maybe we can see Henderson break the record. Are the Yankees in town this week?

If A's goal is *interviewing a pitcher*, then the subset of Yankee pitchers will be salient after his/her query. If A's goal is *finding a ball game to take clients to*, then the set of baseball clubs in New York will be salient. And, if A's goal is to see X break a particular record, then a temporal scale will be salient. Although it is easy to analyze these exchanges 'after the fact', however, it is not clear how S goals can be translated into salience information in general -- and into information about salient items and relationships in particular.

[Grosz 77, Hobbs 79a] propose that, in task-oriented domains, the stages of the process by which that task can be accomplished will be relevant, since S's goal is to accomplish a process in order to accomplish a task. If this assumption is correct, then we might propose that process orderings may be supposed to represent salient relations in such domains. So, in [Hobbs 79a]'s example (reproduced in 259), where the

- (259) A: Have you disconnected the air line?
 B: I loosened it.

task is to assemble an air compressor, *disconnecting the air line* is a relevant process since B

must accomplish it in order to accomplish the task as a whole. So, we can say that stages in this process are salient here and, thus, B can anticipate this recognition on A's part when A interprets the implicature B licenses that $\neg\text{BEL}(B, I \text{ disconnected it})$. Of course, even in non-task-oriented domains, such implicatures may also be generated and understood, so this explanation is partial at best.

Pollack [Pollack 86] proposes a more general model of plan inference which also permits simple location of some types of salient relations for our purposes. Distinguishing GENERATION relations between actions from ENABLEMENT relations, following studies in the philosophy of action, Pollack defines speaker goals in terms of plans which are built up from such relations. In simple terms, when one action α generates another action β , then an agent need only do α and β will automatically be done also. So 'turning on the light' by 'flipping the switch' represents a generation relation between the mentioned actions; 'flipping the switch' generates 'turning on the light'. When α enables β , then an agent needs to do something more than α to guarantee that β will be done. 'Knowing where the switch is' enables 'turning on the light by flipping the switch'. The plans Pollack infers are in effect partial orderings of actions related by generation or enablement. Thus they support scalar implicature, as in 260. Implicatures based on similar orderings

(260) A: Did you turn on the light?
B: I flipped the switch.

have been discussed in Section 5.1.9 as implicatures deriving from the perception of process orderings or prerequisite orderings as salient. Since plans inferred from a discourse can plausibly be assumed salient for S and H in that discourse, these plans will thus represent salient relations for our purposes.

6.3.2.3. Selecting Between Duals

A problem in discerning salient relations by identifying salient expressions which refer to them involves differentiating between duals, or, posets O_i and O_j which differ from one another only in that, where the metric σ_i ordering O_i orders any pair of expressions e_i and e_j as $e_i\sigma_i e_j$, the metric σ_j ordering O_j orders e_i and e_j as $e_j\sigma_j e_i$. This is the notion discussed in Section 5.1.4.1.3 that some orderings appear to 'reverse' other orderings defined over the same set of expressions.

Consider orderings of temperature degrees, ordered by either *is-colder-than* or its dual, *is-warmer-than*. Each mention of a temperature may thus evoke at least two posets. Suppose that, after a cold winter the weatherman has predicted relief from the bitter cold. B returns from getting the paper to be greeted by A, who is just getting up.

- (261) A: Is it warm outside?
 B: It's no 60 degrees out there.

Here B appears to license the implicature $\neg\text{BEL}(\text{B}, \textit{it's warm})$; so, the salient temperature ordering for this exchange is apparently .../ 50/ 60/ 70/ ..., ordered by *is-colder-than*. By denying *60 degrees*, B appears to confirm some lower value on the scale. Alternatively, after a blazing summer the weatherman has promised relief from the heat. B's response in 262 appears to license the implicature $\neg\text{BEL}(\text{B}, \textit{it's cool})$ by implicitly conveying commitment to some lower value

- (262) A: Is it cool outside?
 B: It's no 60 degrees out there.

on a 'reversed' temperature defined by *is-warmer-than*, ...70/ 60/ 50/

In cases such as these we might propose some appropriate 'script' to explain how A and B understand why one scale is appropriate in one context and another in another: Shared knowledge of a prior weather forecast or of the reliability of weather forecasters in general might prove useful in this strategy. However, a strategy less dependent upon such broad contextual cues is also possible: In these and similar discourses, it appears that prior salience (explicit mention in 261 and 262) of a value in the set of referents under consideration may indicate the orientation of the salient ordering. When some such value e_i is salient in prior discourse, e_i may establish the orientation of an ordering in which it appears (or which includes values e_i can characterize) in that the pole closest to e_i (or values associated with it) will be the positive pole of the salient ordering. So, in 261, mention of *warm* favors an ordering which places 'warmer' values at the positive pole, while, in 262, mention of *cool* favors the dual to this ordering.

In some of the naturally occurring data examined for this work an interesting phenomenon emerged: Speakers referred alternately to one poset and then to its dual, apparently to negotiate a range of values which both could agree to. For example, in 263, A first denies a value in

- (263) A: It's not cool outside.
 B: Well, it's not warm.

a temperature ordering *hot/ warm/ tepid/ cool/ cold* to license $\neg\text{BEL}(\text{A}, \neg\textit{it's X outside})$ where $X < \textit{cool}$. Then, B denies a value in *cold/ cool/ tepid/ warm/ hot* to license $\neg\text{BEL}(\text{B}, \neg\textit{it's Y outside})$ for all $Y < \textit{warm}$. In effect, A's and B's implicatures agree only at the value *tepid* -- the intersection of their implicatures. Similar exchanges occur for modifiers like *young/ old, bad/ good, and happy/ sad*.

The problems of selecting salient orderings should by now appear considerable -- and the solutions proposed tentative and partial. While much remains to be explored in the

incorporation of ‘salience’ into the calculation of scalar implicature, some progress has been made. First, as noted in Section 6.2, it is possible to relax the condition that *S* and *H* identify the same ordering as salient and demand only that orderings be **compatible**. So, where several orderings are potentially salient for some expression, it may not always be necessary to select the ‘most salient’, so long as *S* believes all orderings *H* might see as salient to be compatible with *S*’s ordering. Second, it should be possible to use research conclusions from studies of focus to identify salient expressions in an utterance, even if, to date, no satisfactory focus algorithm is available. As we learn more about ‘what is attended to’, we will be better able to test the claim made here that focussed items identify salient expressions for the purpose of generating and interpreting scalar implicature. Finally, while even less is known about the focussing of relationships in discourse than about focussed entities, salient orderings may be identifiable in other ways: For some orderings metrics, it is possible to choose from among several candidate orderings on the basis of properties particular to the defining metric: For classification hierarchies, we can propose a principled mechanism for bounding hierarchies. Where the choice is between dual orderings, we can propose a principled way to select one over another. Knowledge of the domain or of the intentions of conversational participants also may provide clues to the salience of orderings.

While these observations are very far from a comprehensive account of how salient relations may be identified by *S* and *H*, they do suggest fruitful avenues for future research. For the implementation described in Chapter 7, salient expressions and orderings are identified interactively.

6.4. Calculating Scalar Implicatures

In contrast to Gazdar’s two-stage calculation of implicature -- first potential, then actual -- I have proposed that actual scalar implicatures be calculated directly -- but that this calculation be based upon additional information about speaker and hearer beliefs and about the discourse context. Given an utterance, information about speaker intention and mutual speaker-hearer belief about speaker cooperativity, together with contextual information about salient entities and relationships in the discourse, we can use the definition of *S*’s conversational implicature and of *H*’s inference of conversational implicature presented in Chapter 2, the scalar implicature conventions discussed in Chapter 3.3 and refined in Chapters 4, the definitions of utterance ranking presented in Chapter 5, and the conditions on ordering compatibility and ordering salience presented above to compute those scalar implicatures a speaker licenses as well as to calculate inferences which a hearer is entitled to draw.

Recall from Chapter 2 that p_j can be said to represent a conversational implicature of *S* in a context C_h iff:

- S intends to cause H to believe p_j by an utterance u_i (Condition 1);
- S believes that his/her own cooperativity in C_h is a mutual belief of S and H (Condition 2);
- S belief that s/he and H mutually believe that the saying of u_i in C_h must license H 's belief in p_j given belief in S cooperativity (Condition 3);
- p_j is cancelable (Condition 4), non-detachable (Condition 5), and reinforceable (Condition 6).

Recall also that, for scalar implicature, Condition 3 can be satisfied by the satisfaction of the LICENSE predicate, which can be accomplished by S cooperativity plus the truth of any of the scalar implicature conventions introduced in Chapter 3.3 (and modified in Chapters 4 and above). For convenience, I will repeat these here:

$$\begin{aligned}
 Imp_1: & \exists O \text{ BEL}(S, \text{AVAIL}(H, e_i, O, C_h)) \wedge \text{REALIZE}(u_i, \text{AFFIRM}(S, e_i, \text{BEL}(S, p_i))) \wedge (\text{HIGHER_SENT}(p_i, p_j, O) \vee \text{ALT_SENT}(p_i, p_j, O)) \\
 & \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, p_j), C_h) \\
 Imp_2: & \exists O (\text{BEL}(S, \text{AVAIL}(H, e_i, O, C_h)) \wedge \text{REALIZE}(u_i, \text{DENIAL}(S, e_i, \text{BEL}(S, \neg p_i))) \wedge (\text{LOWER_SENT}(p_j, p_i, O) \vee \text{ALT_SENT}(p_i, p_j, O))) \\
 & \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, \neg p_j), C_h) \\
 Imp_3: & \exists O (\text{BEL}(S, \text{AVAIL}(H, e_i, O, C_h)) \wedge \text{REALIZE}(u_i, \text{IGN}(S, e_i, \neg\text{BEL}(S, p_i)))) \Rightarrow \\
 & ((\text{LOWER_SENT}(p_j, p_i, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, \neg p_j), C_i)) \vee \\
 & (\text{HIGHER_SENT}(p_j, p_i, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg\text{BEL}(S, p_j), C_i)) \vee \\
 & (\text{ALT_SENT}(p_i, p_j, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \text{BEL}(S, p_j), C_i)))
 \end{aligned}$$

Then, a scalar implicature p_j can be said to be licensed by S 's utterance of u_i iff S is known or assumed to intend to convey p_j , S and H mutually believe S to be cooperative in uttering u_i , if S is obeying the maxims of QUANTITY and QUALITY, if one of the scalar implicature conventions holds, and if the meaning licensed is cancelable, nondetachable, and reinforceable. One of the scalar implicature conventions will hold when, roughly speaking, p_j represents a higher, lower or alternate sentence with respect to the p_i realized by u_i via the presence in p_i and p_j of expressions e_i and e_j which appear in some ordering O believed by S to be salient in the discourse and compatible with all the possible orderings S believes H might find salient. The following algorithm encodes the process by which the scalar implicature conventions can be tested for a given utterance:

Given S 's u_i with the semantic representation p_i :

1. decide whether any e_i in p_i evokes a salient O ;
2. determine whether p_i represents the affirmation, denial, or declaration of S 's ignorance about some proposition with semantic representation p_k which contains

- e_i , if p_k is simple with respect to e_i .¹⁵¹
3. create an open sentence p_x by replacing e_i in p_k with x ;
 4. locate other values e_j appearing with e_i in O ;
 5. for each e_j create a sentence p_j by replacing x with e_j in p_x -- so that p_j and p_k are simple expression alternatives (See Section 4.1.3.);
 6. determine whether p_j satisfies any of the the scalar implicature conventions.

Returning to examples 218-221 (repeated here for convenience), we can now account for licensed implicatures in terms of this algorithm. Recall that, in 218,

- (1) A: Have you, me, and Ellen ever had dinner together?
 B: We've had lunch.

assuming B's intentions and cooperativity, we can say that $\neg\text{BEL}(\text{B, we have had dinner together})$ is a scalar implicature of B's response, so long as B believes an ordering defined by set inclusion on the set of *meals* ($= \{\text{breakfast, lunch, dinner}\}$) is salient and that A will recognize either this ordering or one compatible with it -- say, perhaps just the explicitly mentioned items $\{\text{lunch, dinner}\}$. Since *lunch* and *dinner* represent proper (singleton) subsets of this set (See Note 130.), they represent alternate values in a salient ordering. By Imp_1 , affirmation of some value v_i (denoted by an expression e_i) may license the belief that alternate values v_j are false or unknown. So, the queried *dinner* -- and other alternates such as *breakfast* -- can be implicitly marked as false or unknown in this way.

In terms of the algorithm presented above, this implicature might be analyzed as follows: Let p_i , the representation of B's response, be 'BEL(Past(eat({B, A, Ellen}, {lunch}), t_0))'.¹⁵² In this sentence, the expression *lunch* evokes a salient ordering of *meals*, so $e_i = \text{lunch}$ and $O = \text{the set of meals}$. p_i represents the affirmation by B of p_k , 'Past(eat({B, A, Ellen}, {lunch}), t_0)'. p_k contains e_i . We can create the open sentence 'Past(eat({B, A, Ellen}, X), t_0)' from p_k by replacing e_i with a variable X. Then for all alternate values in the above representation -- $\{\text{breakfast}\}$ and $\{\text{dinner}\}$ -- we can create simple sentence alternatives 'Past(eat({B, A, Ellen}, {breakfast}), t_0)' and 'Past(eat({B, A, Ellen}, {dinner}), t_0)'. Since these instantiations of X do represent alternate values in *the set of meals*, by the scalar implicature conventions, we can

¹⁵¹i.e., contains no negation with scope over e_i . See Section 4.1.3.

¹⁵²The algorithm is independent of any particular representation. For this example I will employ a higher order modal logic which permits representation of tense (here, 'Past(x, t)' indicates that x held at some time before time t) and quantification over sets. The latter is necessary here to provide a clumsy solution the well-known problem of representing *together*: That is, utterances such as *Jane and Bill are eating lunch together* are not properly represented by 'eat(lunch, jane) \wedge eat(lunch, bill)'.

calculate that $\neg\text{BEL}(\text{B}, \text{'Past}(\text{eat}(\{\text{B}, \text{A}, \text{Ellen}\}, \{\text{breakfast}\}), t_0)) \neg\text{BEL}(\text{B}, \text{'Past}(\text{eat}(\{\text{B}, \text{A}, \text{Ellen}\}, \{\text{dinner}\}), t_0))$ may be licensed in 218.

In 219, however, B perceives an ordering of *social events* to be salient. So, we might calculate this response by representing p_k as

- (1) A: Have you, me, and Ellen ever had dinner together?
B: We've been to the movies.

'Past(go({B, A, Ellen}, {movies}), t_0)', p_x as 'Past(X({B, A, Ellen}, Y), t_0)', and implicatures licensed by B's response, as $\neg\text{BEL}(\text{B}, \text{Past}(\text{X}(\{\text{B}, \text{A}, \text{Ellen}\}, \text{Y}), t_0))$ for all expressions e_i of the form 'X({B, A, Ellen}, Y)' \in *social events* that are alternate values to 'go({B, A, Ellen}, {movies})'.

Finally, in 220, where the ordering supporting B's implicature is defined over the set of {you, me, Ellen}, the implicature $\neg\text{BEL}(\text{B}, \text{you, me, and Ellen have had dinner together})$ may be calculated as follows:

- (1) A: Have you, me, and Ellen ever had dinner together?
B: Ellen and I have.

Let p_k be 'Past(eat({B, Ellen}, {dinner}), t_0)' and p_x be 'Past(eat(X, {dinner}), t_0)'. Then, for all e_i that are alternate expressions to the set {Ellen, B} or that are higher expressions in the salient ordering, B may license the implicature $\neg\text{BEL}(\text{B}, \text{Past}(\text{eat}(e_i, \{\text{dinner}\}), t_0))$. Hence, for the higher expression {B, A, Ellen}, B licenses $\neg\text{BEL}(\text{B}, \text{Past}(\text{eat}(\{\text{B}, \text{A}, \text{Ellen}\}, \{\text{dinner}\}), t_0))$.

So, given an utterance, an expression from its semantic representation, and a salient ordering, it is a fairly straightforward task to calculate the scalar implicatures which may be licensed via this utterance. However, the algorithm presented above will not accommodate the calculation of scalar implicature when more than one expression is salient in this utterance, as in 221 (repeated below).

- (1) A: Have you, me, and Ellen ever had dinner together?
B: Ellen and I have had lunch.

In this exchange, orderings defined over the subsets of {you, me, Ellen} and {breakfast, lunch, dinner} are both evoked by B's response to implicate $\neg\text{BEL}(\text{B}, \text{you, me, and Ellen have had dinner together})$. While p_k is 'Past(eat({B, Ellen}, {lunch}), t_0)' and p_x , 'Past(eat(X, Y), t_0)', it is not clear how licensed implicatures should be represented -- or even just what they are. Intuitively, B seems to be licensing $\neg\text{BEL}(\text{B}, \text{Past}(\text{eat}(\{\text{B}, \text{A}, \text{Ellen}\}, \{\text{dinner}\}), t_0))$, but does she also license $\neg\text{BEL}(\text{B}, \text{Past}(\text{eat}(\{\text{B}, \text{A}, \text{Ellen}\}, \{\text{breakfast}\}), t_0))$? $\neg\text{BEL}(\text{B}, \text{Past}(\text{eat}(\{\text{B}, \text{A}, \text{Ellen}\}, \{\text{dinner}\}), t_0))$? These more complex implicatures will require some revision of the algorithm presented above.

6.4.1. Multiple Salient Expressions

When several expressions in a sentence are salient, the calculation of potential and actual implicatures is much more complex than has previously been recognized. A major weakness of Gazdar's method for calculating potential implicature is that implicatures are calculated separately for each expression recognized in a sentence. That is, for salient expressions e_i , e_j contained in a sentence p_i , with associated orderings O_i and O_j , implicatures licensed via e_i will be calculated by replacing e_i in p_i by higher¹⁵³ values on O_i . Then, implicatures licensed via e_j will be calculated -- also by 'replacing' e_j with higher values in O_j in the original p_i . Implicatures that may involve, say, higher values on O_i in addition to higher values on O_j will not be predicted by Gazdar's method -- or by the theory I have presented so far.

So, for example, applying these methods to B's response 221 would predict that B may license \neg for any p_j from (264a)-(264d) -- but neither (264e), nor (264f), nor (264g). If indeed the orderings stipulated are salient, then

(264)

- a. Ellen, you, and me have had lunch.
- b. Ellen and you have had dinner.
- c. Ellen and I have had dinner.
- d. Ellen and I have had breakfast.
- e. Ellen, you, and me have had dinner.
- f. Ellen, you, and me have had breakfast.
- g. Ellen, you, and me have had breakfast and dinner.

it seems clear that (264e) and (264f) should be considered scalar implicatures as well.

A simple addition to the definitions of higher, lower, and alternate sentences permits calculation of implicatures licensed when one or more expressions are salient in a sentence: Recall that sentences are defined as higher, lower, or alternate to other sentences when they differ in only a single expression as follows:

Higher Sentences:

$$\text{HIGHER_SENT}(p_i, p_j, O) \leftrightarrow \exists O \exists e_i \exists e_j (\text{HIGHER}(e_i, e_j, O) \wedge \text{SIMPLE_EXP_ALT}(p_i, p_j, e_i, e_j))$$

Lower Sentences:

$$\text{LOWER_SENT}(p_i, p_j, O) \leftrightarrow \exists O \exists e_i \exists e_j (\text{LOWER}(e_i, e_j, O) \wedge \text{SIMPLE_EXP_ALT}(p_i, p_j, e_i, e_j))$$

Alternate Sentences:

$$\text{ALT_SENT}(p_i, p_j, O) \leftrightarrow \exists O \exists e_i \exists e_j (\text{ALTERNATE}(e_i, e_j, O) \wedge \text{SIMPLE_EXP_ALT}(p_i, p_j, e_i, e_j))$$

¹⁵³Gazdar does not deal with lower or alternate values.

Intuitively, we want to extend the notions of utterance ranking so that one sentence p_i is higher than another p_j , for at least one pair of expressions in which p_i and p_j differ, p_i ranks higher than p_j , and for no other pair of differing expressions does p_j outrank p_i . So, (264g) will represent a higher sentence than B's response in 221, as well as (264a)-(264f). Similarly, (264e) will represent a higher-ranked sentence than (264c), since the expression {Ellen, A, B} in (264e) includes and thus is higher than the expression {Ellen, B} in (264c) in an ordering defined on the non-null subsets of {Ellen, A, B}. However, (264e) will also represent a higher-ranked utterance than B's response in the original exchange 221 -- '*Ellen and I have had lunch.*' Since {*lunch*} and {*dinner*} represent alternates in the *set-of-meals* ordering, even though (264e) and the response in 221 differ in these expressions, (264e) will still be ranked higher via the {Ellen, A, B} ordering. So, B may implicate $\neg(264e)$ in 221.

Then, p_i should rank lower than p_j whenever the dual holds: for at least one ordering p_i ranks lower than p_j , while nowhere else ranking higher than p_j . B's responses in 221-(264f) are, then, lower sentences than (264g). Sentences should be alternates whenever every expression pair in which they differ represents a pair of alternate expressions. So, (264b) and (264d), for example, will represent alternate sentences.

These extended notions of sentence ranking can be captured as follows:

A sentence p_i is extended_higher than another sentence p_j just in case, for at least one expression pair $\langle e_i, e_j \rangle$ in which p_i and p_j differ, e_i is higher than e_j in the appropriate ordering, and for all remaining expressions e_i and e_j in which p_i and p_j differ, each e_i is higher than or alternate to each e_j in the appropriate ordering.

An extended notion of what it means for one sentence to be ranked lower than another will be the dual of the extended notion of higher sentences:

A sentence p_i will be extended_lower than another sentence p_j just in case, for at least one expression pair $\langle e_i, e_j \rangle$ in which p_i and p_j differ, e_i is lower than e_j in the appropriate ordering, and for all remaining expressions e_i and e_j in which p_i and p_j differ, each e_i is lower than or alternate to each e_j in the appropriate ordering.

Extended sentence alternates are simply:

p_i will be extended_alternate to p_j just in case, for every expression pair $\langle e_i, e_j \rangle$ in which p_i and p_j differ, e_i is alternate than e_j in the appropriate ordering.

These extended notions of sentence ranking can be defined in terms of the simple notions of higher, lower, and alternate sentences as follows:

$$\begin{aligned} \text{EXT_HIGHER_SENT}(p_i, p_j) &\leftrightarrow (\exists O_i \text{HIGHER_SENT}(p_i, p_j, O_i)) \vee \\ &(\exists p_k \exists O_j (\text{HIGHER_SENT}(p_k, p_j, O_j) \vee \text{ALT_SENT}(p_k, p_j, O_j)) \wedge \\ &\text{EXT_HIGHER_SENT}(p_i, p_k)) \end{aligned}$$

$$\text{EXT_LOWER_SENT}(p_i, p_j) \leftrightarrow (\exists O_i \text{LOWER_SENT}(p_i, p_j, O_i)) \vee \\ (\exists p_k \exists O_j (\text{LOWER_SENT}(p_k, p_j, O_j) \vee \text{ALT_SENT}(p_k, p_j, O_j)) \wedge \\ \text{EXT_LOWER_SENT}(p_i, p_k))$$

$$\text{EXT_ALT_SENT}(p_i, p_j) \leftrightarrow (\exists O_i \text{ALT_SENT}(p_i, p_j, O_i)) \vee (\exists p_k \exists O_j \text{ALT_SENT}(p_k, p_j, \\ O_j) \wedge \text{EXT_ALT_SENT}(p_i, p_k))$$

Then the resulting version of the scalar implicature conventions, incorporating these extended notions of sentence ranking, will be:

$$\text{Imp}_1: \quad \exists O (\text{BEL}(S, \text{AVAIL}(H, e_i, O, C_h)) \wedge \text{REALIZE}(u_i, \text{AFFIRM}(S, e_i, \text{BEL}(S, \\ p_i))) \wedge (\text{EXT_HIGHER_SENT}(p_j, p_i, O) \vee \text{EXT_ALT_SENT}(p_i, p_j, O))) \\ \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg \text{BEL}(S, p_j), C_h)$$

$$\text{Imp}_2: \quad \exists O (\text{BEL}(S, \text{AVAIL}(H, e_i, O, C_h)) \wedge \text{REALIZE}(u_i, \text{DENIAL}(S, e_i, \text{BEL}(S, \\ \neg p_i))) \wedge (\text{EXT_LOWER_SENT}(p_j, p_i, O) \vee \text{EXT_ALT_SENT}(p_i, p_j, O))) \\ \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg \text{BEL}(S, \neg p_j), C_h)$$

$$\text{Imp}_3: \quad \exists O (\text{BEL}(S, \text{AVAIL}(H, e_i, O, C_h)) \wedge \text{REALIZE}(u_i, \text{IGN}(S, e_i, \neg \text{BEL}(S, \\ p_i)))) \Rightarrow \\ ((\text{EXT_LOWER_SENT}(p_j, p_i, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg \text{BEL}(S, \\ \neg p_j), C_i)) \vee \\ (\text{EXT_HIGHER_SENT}(p_j, p_i, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \neg \text{BEL}(S, p_j), \\ C_i)) \vee \\ (\text{EXT_ALT_SENT}(p_i, p_j, O) \Rightarrow \text{SCALAR_IMP}(S, H, u_i, \text{BEL}(S, p_j), C_i)))$$

Then the algorithm for calculating scalar implicatures must be altered to accommodate multiple salient expressions as follows: Given S 's u_i with the semantic representation p_i :

1. determine whether any expressions e_i in p_i evoke salient orderings O_i and, if so, associate each such e_i with the appropriate O_i ;
2. determine for each e_i whether p_i represents the affirmation, denial, or declaration of S 's ignorance about some proposition with semantic representation p_k which contains e_i (i.e., if p_k is simple with respect to e_i ;¹⁵⁴
3. create an open sentence p_x by replacing each e_i in p_k with a (distinct) variable x_i ;
4. locate other values e_j appearing with each e_i in the appropriate O_i ;
5. for every p_j resulting from the replacement of one or more x_i in p_x by some e_j such that p_j and p_k are simple expression alternatives,¹⁵⁵ determine whether p_j represents a scalar implicature by the scalar implicature conventions.

These revisions complete the theory of scalar implicature presented in this thesis.

¹⁵⁴See Section 4.1.3.

¹⁵⁵See Section 4.1.3.

6.5. Summary

In this chapter I have presented some limited solutions to the problem of how context might be incorporated into the calculation of scalar implicature. I have proposed a number of ways in which salience information might be derived from focus information and from knowledge of relations that may be salient in the discourse. Finally, I have identified certain revisions which permit the calculation of calculate scalar implicatures licensed when utterances include one or more salient expressions and have presented an algorithm which can be used to calculate these implicatures.

While the calculation of licensed implicatures should be important to any attempt at natural-language generation or interpretation, permitting the anticipating or understanding of pragmatic inferences derivable from any natural-language output or input, it is helpful to show how the computation of scalar implicatures can be turned to particular computational use. In Chapter 7, I will describe how knowledge of scalar implicature permits the generation of one form of cooperative response in question/ answer systems.

CHAPTER VII

Scalar Implicature and Question-Answering

I don't want you, but I need you.
I don't like you, but I love you.
Eddy Rabbit

Studies of question-answering by philosophers, linguists, and computer scientists have proposed numerous accounts of speakers' propensity for providing more or less information than has been requested by a questioner. In this chapter I add to this literature. A large class of indirect and modified or qualified direct responses to yes-no questions can be explained as attempts by speakers to block scalar implicatures which hearers might otherwise be expected to infer. First, I situate this proposal in previous studies of cooperative question-answering and theoretical studies of yes-no questions. Next, I describe how scalar responses can be explained as attempts to block potential inferences of scalar implicatures. Finally, I describe QUASI, a question-answering system that provides cooperative responses to yes-no questions by calculating the scalar implicatures users might be entitled to infer from direct responses to these questions. While it would be foolish to interpret or to generate responses to any class of questions solely on the basis of a single phenomenon such as scalar implicature, QUASI does demonstrate the feasibility of calculating scalar implicatures and one area in which their calculation is useful. In Chapter 8, I will speculate about some further ways in which the calculation of scalar implicatures should be important to more general studies of discourse.

7.1. Appropriate Responses to Yes-No Questions

Cooperativity in question-answering has been defined by a variety of disciplines in a variety of ways: Philosophical studies of question-answering have defined the notion of ANSWERHOOD: the conditions under which a response counts as an answer to a natural-language query. Students of natural-language processing have identified types of cooperative speaker behavior from studies of naturally occurring dialogues and from introspection about the sort of

behavior computer-human interaction should support. In particular, AI's approach to question-answering has focussed on identifying when systems should provide more information than their users have requested and, of course, what that information should be. Anticipation of follow-up questions, explanation of questioners' violated expectations, provision of information relevant to questioners' inferred goals, and correction of misconceptions perceived in a query are only a few types of behavior which AI systems have been or are being designed to support.

7.1.1. Theoretical Accounts of Yes-No Questions

Theorists of question-answering commonly define questions in terms of the set of all their possible (true) answers [Kiefer 80, Hamblin 71]. Yes-no questions are generally seen as a disjunction of all the possible response to them. Traditional accounts portray them as propositional questions (?P) or as a special type of alternative question (?P \vee ?Q) in which the second alternative is simply the negation of the first (?P \vee ? \neg P). So, 'I ask you whether x' or 'Tell me which of the following is true, x or \sim x' [Katz 64] or 'Bring it about that I know that x or \sim x' [Hintikka 78], where x is the DESIDERATUM¹⁵⁶ of the query have all been employed as representations. The 'meaning' of these questions then is the set of all their possible answers, i.e., the set of propositions which might truly be asserted in response.

However, both theoretical work and empirical studies of naturally occurring question-answer exchanges have shown this approach to be inadequate: *Yes* and *no* or even *yes*, *no*, and *unknown* form only a small portion of the set of all appropriate responses to a yes-no question. Furthermore, for some yes-no questions, none of these simple direct responses alone is appropriate.

I have previously noted (See Section 6.3.1.2.2.) Kiefer [Kiefer 80]'s claim that certain yes-no questions actually function as *wh*-questions and indicated its limitations. Aqvist [Aqvist ??] also notes that representing yes-no questions as alternative questions does allow proper treatment of negative queries. In Bolinger's [Bolinger 78] view the term YES-NO QUERY has hypnotized scholars into assuming that, simply because a class of question can be answered by a *yes* or *no*, these alternatives are criterial, and every yes-no question is intended to elicit one or the other. He proposes instead that yes-no questions be viewed as hypotheses put forward for confirmation, amendment, or disconfirmation -- in any degree. Thus, in his example 265, the

(265) Q: Do you like Honolulu?
R: Just a little.

¹⁵⁶A specification of the epistemic state desired by the questioner.

Q's hypothesis 'you like Honoíulu' is amended by the R in a response which is neither *yes* nor *no* but somewhere in between. In his example 266, Q's hypothesis 'it is

- (266) Q: Is it difficult?
R: It's impossible.

difficult' is confirmed by R's assertion of a more positive response than a simple *yes*.

While Bolinger makes a good case for the inadequacy of standard views of yes-no questions, the revision he proposes is itself too limited. 'It's impossible', in 266, does more than simply present a strong affirmation of the hypothesis 'it is difficult' -- it provides new and unrequested though pertinent information. In fact, 'strong affirmation' might better be provided by a response such as 'I am absolutely sure it's difficult' than by the response he suggests. And there are equally appropriate responses to the queries in 265 and 266 that are not easily explained in terms of degree of hypothesis confirmation, as shown in 267 and 268.

- (267) Q: Do you like Honolulu?
a. R: I don't like Hawaii.
b. R: I like Hilo.

- (268) Q: Is it difficult?
a. R: It could be.
b. R: Mike says so.

An alternative account of certain yes-no questions from a COMMON-SENSE REASONING point of view is presented by Sadock [Sadock 77] in his work on MODUS BREVIS. He suggests that exchanges like 269

- (269) A: Is a bat a bird?
B: Well, it's got wings.

in fact represent forms of fallacious reasoning -- in 269, the fallacy of *asserting the consequent*. So, in this exchange, A argues for inclusion of *bats* among *birds* by the implicit 'Well, if something is a bird, it's got wings.' Similarly, according to Sadock, in 270, B invokes a *modus tollens* argument by his/her response, which also is fallacious:

- (270) A: Do you have Pepsi?
B: We have Coke.

That is, B assumes the premise 'We have Pepsi or we have coke' and incorrectly assumes the exclusivity of *or*.

Numerous additional examples are presented in further evidence of such 'informal' reasoning; Sadock concludes that speakers' failure to make their premisses and conclusions explicit makes it possible 'to cover up logical abuses that might be more flagrant if the argument were stretched out [Sadock 77:547]. However, it is not clear that B is necessarily

conveying an affirmative response in 269. While Sadock's explanation is ingenious, the theory of scalar implicature presented in this thesis seems to present a simpler explanation of much of his data without imputing wide-spread fallacious reasoning to speakers.

7.1.2. Computational Approaches to Yes-No Questions

A number of natural-language question-answering systems have recognized the desirability of supporting the sort of indirect responses to yes-no questions examined here. However, to date, few have proposed independent principled bases for the behavior they model. The 'cooperativity' of the additional information provided must largely be assumed and often is driven solely by the fact that the information has already been calculated; I will classify such systems as 'availability-driven' with regard to their provision of additional information to queries. Also, of the systems discussed below, only the HAM-ANS system is particularly concerned with yes-no questions.

7.1.2.1. Availability-Driven Responses

An early example of systems which provide additional information essentially as a side-effect of other system processes is Lehnert [Lehnert 77]'s QUALM, a question-answering system which answered questions about stories conceptually represented in the system KB. QUALM included a verification option which could, under certain conditions, give the effect of providing modified direct response to yes-no questions by verifying affirmative responses. When QUALM was in *talkative* mode and the question category was *verification*, and the direct response to a question was affirmative, dialogues like 271 could result:

(271) Did John go to New York?
 Yes, John went to New York by bus.
 Did John eat?
 Yes, John ate lasagna.
 Did someone pick John's pocket?
 Yes, a thief picked John's pocket.
 Did John pay the check?
 Yes, John paid the bill.

QUALM's elaboration mechanism, in generating verifications from conceptual representations in its KB, output 'Yes, *X*' where *X* was the conceptualization found in the story representation that matches the question concept. Where this conceptualization did not exactly duplicate the question concept matched to it, additional information (in effect, QUALM's own representation of the concept) was presented.

[Kaplan 79]'s COOP system provided early evidence that pragmatic phenomena could

help in the selection of system responses: COOP defined cooperative responses to wh-questions in terms of their PRESUMPTIONS, which Kaplan derived from the pragmatic notion of presupposition.¹⁵⁷ While the correction of false presumptions was the heart of the COOP system, Kaplan also suggested a parallel account of indirect responses to yes-no questions in his FOCUS responses, in which respondents identify information likely to be requested in a subsequent query with the focus of the current query, as in 272, and his SUGGESTIVE indirect responses

(272) Q: Is there a mailbox on this block?(= [Kaplan 79]'s 1)

R: No, but there's one down the street.

as in 273. In the latter, which are clearly included in

(273) Q: Is John a senior?(= [Kaplan 79]'s 8)

R: He's a junior.

my class of scalar responses, Kaplan claims that speakers phrase questions to avoid a negative or trivial response like 'none' or 'zero'. If respondents are asked questions eliciting such responses, they are justified in assuming that some (perhaps rather weak) expectation the speaker had has been violated. So, a suggestive indirect response like that in 273 is appropriate. However, this claim is not fully developed or independently motivated.

The most significant existing system to address the problem of providing modified direct responses to yes-no questions directly is the HAM-ANS system [Hoepfner 84b, Hoepfner 84c, Hoepfner 84a]. Disclaiming any attempt at linguistic or cognitive justification [Hoepfner 83], the designers of this system provide EXTENDED RESPONSES to yes-no questions based upon a variety of heuristics, such as finding explicit partial answers to each conjunct of conjoined NPs to avoid an uncooperative *no* when at least one conjunct is successful, as in 274; 'modifying' universal

(274) Is there an easy-chair or a large chair in the room?

A large chair is not available, but an easy-chair.

quantifiers or cardinals when a queried quantifiers cannot be satisfied to a one which can be, as in 275;¹⁵⁸

and filling optional deep case slots in the case frame of a verb used in the query, as in 276.

¹⁵⁷Kaplan defined LOADED questions as those which indicate a questioner presumes something to be true about the domain which is in fact false. So, in x, if Q knows that CIS 110 has not been offered this year, s/he could be assumed

(x) Q: How many students failed CIS 110 this year?

R: CIS 110 wasn't offered this year.

to know the direct response to his/her query, *none*. By asking the query, Q reveals a false presumption, which R cooperatively corrects.

¹⁵⁸Note that the direct response *yes* is not itself appropriate here.

(275) User: Is there a lamp next to every bed?

HAM-ANS: Yes, next to almost every one.

(276) Did a yellow car and a green stationwagen turn off earlier?

No. A yellow car turned off Hartungstreet onto Schlueterstreet, but the green stationwagen did not turn off first.

The quantity of additional information provided is limited by other heuristics and by the amount of information needed to compute the response.¹⁵⁹ While HAM-ANS results often appear quite 'natural', the lack of independent linguistic or cognitive motivation behind them can produce some inconsistency; the amount and nature of information provided is ultimately dependent upon the case frame of the verb or the presence of a few particular lexical items. So, a questioner's use of one verb may result in the inclusion of an additional modifier in the response, where use of another will not. The critical link between the behavior produced and its appropriateness is not established.

7.1.2.2. Goal Inference and Indirect Responses

A more principled approach to indirect responses to yes-no questions was signalled by early work on indirect responses in task-oriented domains by Hobbs and Robinson [Hobbs 79a]. In a classic essay, Hobbs and Robinson distinguish three classes of appropriate responses which, though indirect, nonetheless addressed the goals of the question asked: those which indeed answer the question asked indirectly, those which deny a presupposition of the question, and those which address a higher goal of the question. Problems with identifying responses in the first and third categories purely from linguistic evidence lead Hobbs and Robinson to restrict the utterances under consideration to those contained in task-oriented dialogues where speaker and hearer goals can be predetermined; their tokens are drawn from dialogues between an expert and an apprentice trying to repair an air compressor. Example 259 (repeated here) is thus explained by [Hobbs 79a] in terms of particular

E: Have you disconnected the air line?

A: I loosened it.

domain goals: Since Expert and Apprentice both believe that the goal of their discourse is the completion of the air compressor repair, then if A is obeying the Maxim of Quantity, he will give as much information about his progress toward fulfilling this goal as he can. If *to disconnect* means *to cause not to be attached* and *to loosen* means *to cause not to be tightly attached*, then, since E can assume A has done no more toward achieving the goal of

¹⁵⁹XCALIBRE uses this means more simplistically to guide the provision of additional information in responses to *wh*-questions, outputting all the information accessed to compute the direct response.

disconnection than he has stated, he can conclude that, although the air line is not attached tightly it is still attached and the literal answer is *no*. Such answers are thus appropriate because, even if they do not provide the information requested directly, they provide *some* information that allows the questioner to achieve the goal of his/her query.

While [Hobbs 79a]'s account of 259 is intuitively satisfying, it lacks the generality: Domain goals may indeed provide an important clue to the appropriateness of indirect responses, but it is not clear how one might abstract beyond the particular examples provided to derive useful generalizations. However, research more directly concerned with goal inference does show how inferred goals can guide the provision of additional cooperative information. Allen and Perrault [Allen 80] examine the provision of additional information to direct and indirect responses based upon questioners' inferred goals in their. Since they limit possible goals to either taking or meeting a train, however, abstraction to broader classes of indirect responses is again problematic. More recent work by Carberry [Carberry 83] on goal tracking and by Pollack [Pollack 86] on inferring questions which questioners may themselves be unaware of promises a broader range of goal inference. In Section 6.3.2.2 I have proposed how such goal inference might facilitate the calculation of scalar implicatures by providing information on salient expressions and orderings.

The detection and correction, as well as the prevention of misconceptions forms a broad category of domain-independent goals for indirect and modified direct responses. Joshi [Joshi 82] has noted the importance of *squaring away* mutual beliefs in question-answering when respondents perceive misconceptions in an exchange -- lest the questioner's misconceptions be implicitly confirmed. In human-machine interaction even more than in natural discourse, Joshi suggests that people may expect such cooperative behavior, since the possible processing restraints which make human cooperativity difficult may not be expected to apply to more resourceful machines.¹⁶⁰ This assumption underlies Kaplan's work, as well as related work on presumption failure by Mays [Mays 80a, Mays 80b], more recent work on utterance presupposition by Mercer and Rosenberg [Mercer 84], and work on the correction of object-related misconceptions by McCoy [McCoy 85].

Joshi's redefinition of the Maxim of Quality (See Section 2.3.) suggests that, even beyond correcting perceived misconceptions, cooperative speakers have an obligation to anticipate and prevent misconceptions which their otherwise truthful utterances might engender. Preliminary

¹⁶⁰For example, *H* may not expect a human *S* to consider all the possible consequences of his/her utterances, Such consideration would clearly require super-human computational power. But in interaction with a computer, humans may expect more, since they know/believe that machine processing is easier. So, the user of a question-answering system may expect that system to be more cooperative than another human [Joshi 82:191].

studies by Joshi, Webber, and Weischedel [Joshi 84a, Joshi 84b] have examined exchanges in which cooperative respondents are obligated to block misleading inferences. If a questioner Q expects that an expert respondent R would inform him of some proposition p_i related to his/her query if in fact that R knew p_i to be true, then Q may interpret the R's silence regarding p_i as implying $\neg p_i$. So, if R knows p_i to be true, his/her silence may lead to Q's being misled. So, R should block such inferences, as in :

Q: Is Sam an associate professor?

R: Yes, but he doesn't have tenure.

The authors suggest that their phenomenon might be represented in Reiter's [Reiter 80] default logic. The actual characterization of those p_i to be represented, however, awaits further study.

7.1.3. Blocking Potential False Scalar Implicatures

Many of the approaches described above provide useful information for the planning of text to be generated in response to yes-no questions. I will now discuss yet another approach which subsumes some of those described above and complements others.

A large class of indirect and modified direct responses to yes-no questions, exemplified in many of the examples presented throughout this thesis, may be analyzed as attempts by cooperative speakers to block potential false inferences which hearers might otherwise (wrongly) infer to be implicatures arising from a direct response -- while also providing information from which the direct response can be derived. While no claims to cognitive reality are made for this explanation, it does appear to account for the data, it is plausible (particularly if Joshi's claims are correct), and it does provide a principled and linguistically motivated basis on which to support much of the behavior supported by the systems described in Section 7.1.2.1.

Recall that a central claim of work on quantity implicature in general has been that speakers derive considerable information from what is not said in comparison to what is said. Similarly, the appropriateness of indirect responses such as that provided in 277 depends in large part upon the inappropriateness of alternative responses available to B.

(277) A: Has Marvin had his medication?

B: He's taken the exadin.

Suppose that B (a simple question-answering system) can only generate simple direct responses to yes-no questions even when more information is available, as in 278:

Such a respondent forces A to play Twenty Questions to elicit the information provided in a single indirect response in 277. While such limited response capability is obviously tedious, it may have more serious consequences. For, if A does not recognize that B can provide only *yes*

- (278) A: Has Marvin had his medication?
 B: No.
 A: Do you mean he hasn't had any of it?
 B: No.
 A: Well, has he had his cortisone?
 B: No.
 A: Exadin?
 B: Yes.
 A: How much of that has he had?
 B: 5 cc.
 A: Has he taken the niaprin?

or *no* responses, s/he may not ask for clarification after B's initial *no*. So, s/he will not learn that in fact Marvin has taken part of his medication. Thus, simple direct responses may be uncooperative, if technically truthful. And note that, in 277, A may infer the direct response (*No. Marvin has not taken all of this medication.*) via the scalar implicature conventions, if s/he understands that B has said as much as s/he truthfully can.

In addition to facilitating the brief conveyance of relevant information, knowledge of scalar implicature can be used more directly to avoid block certain false inferences I will term POTENTIAL FALSE SCALAR INFERENCEs. These inferences may be identified computationally by calculating the scalar implicatures a speaker (system) might license given that certain expressions and posets are deemed salient in the discourse and by comparing these licenseable inferences to the speaker (system's) knowledge base.

Assume that B believes that Madge has five children. Let X = the set of propositions { \neg BEL(B, Madge has 4 children), \neg BEL(B, Madge has 5 children)}. Then, given the scalar implicature conventions, we know that a simple direct response (279a) to A's query in 279 would

- (279) A: Does Madge have three children?
 a. B: Yes (Madge has three children).
 b. B: She has five.

allow B to implicate X -- if B believes s/he and A mutually believe the cardinals (or some compatible ordering) salient in the discourse. While, by the definition of conversational implicature presented in Chapter 2, B will only implicate X if s/he intends to implicate X, another necessary condition on conversational implicature is B's anticipation of A's ability to 'work out' X. This process is itself defined in Section 2.4.4. Even if B does not intend to convey X, then, s/he may still anticipate that A may infer X, if the conditions under which B might implicate X hold. If B does anticipate that A may infer X -- but if X is inconsistent with B's beliefs, then we may explain B's alternative response of (279b) as the blocking of a potential inference which is inconsistent with B's beliefs.

So, where a simple *yes* or *no* might be interpreted as licensing implicatures inconsistent with *S*'s beliefs, *S* may provide a scalar response to block those inferences. More generally, for any speaker *S*, a sentence p_j represents a potential false implicature of an utterance u_i iff, in the current context C_h , *H* will be entitled to infer that *S*'s u_i licenses p_j as a scalar implicature -- but p_j is inconsistent with *S*'s beliefs.

Of course, even with this admittedly over-simplified and stylized analysis, there may be more than one scalar response which will not license potential false inferences of the sort described above. For example, if Marvin's

- (280) A: Has Marvin taken his medication?
 a. B: He's taken the exadin.
 b. B: He hasn't taken the niaprin.

medication consists of *exadin* and *niaprin*, and if the appropriate mutual beliefs hold between A and B, then (280a) and (280b) will both represent 'cooperative' responses in the sense that they will not license potential false inferences of the sort described above (if an ordering of *medication* with its subparts *exadin* and *niaprin* is salient). B may deny the lowest value *s/he* can truthfully deny or affirm the highest -- in this example lowest and higher are in fact alternate values in the ordering. Such examples illustrate the dangers of a 'single theme' approach to question-answering: While the notion of blocking false potential implicatures can indicate that either (280a) or (280b) is to be preferred over a simple *no* when *parts-of-medication* is salient, it cannot help to select between the two responses.

So, this approach to providing cooperative responses to yes-no questions should be seen only as providing one test of the appropriateness or cooperativity of any response. Utterances that license no potential false scalar inferences may be inappropriate on other grounds -- perhaps they fail to address the user's goals in some way or provide more information than the user desires. Other aspects of the discourse -- prior 'cancelation' for one -- might overrule the inappropriateness of utterances that do license such false inferences. So, in an ideal world, an understanding of scalar implicature should be used for cooperative question-answering only in conjunction with other criteria. With this caveat, I will proceed to describe the QUASI system implemented to demonstrate the calculation of scalar implicature.

7.2. The QUASI System

QUASI is a question-answering system which proposes cooperative responses to yes-no questions about a mail system from its users. Using the representation and algorithms described in previous chapters, QUASI calculates the potential scalar implicatures which simple direct or alternative responses might license when certain expressions and posets are salient. It tests these possible inferences for consistency with its own knowledge base, and, when inconsistency is detected, proposes alternate responses that do **not** license inconsistent scalar implicatures. The system has been implemented in CProlog on a VAX 11-785. In this section, I will provide an overview of the system, examine its major constituents, and illustrate particular system features.

The goal of this implementation is to demonstrate how the formalization of scalar implicature can aid in the calculation of licenseable implicatures and to suggest a simple use for such information. Not all aspects of the theory presented in the thesis have been implemented. First, in the absence of a user model, QUASI makes the simplifying assumption that the user believes the system cooperative and that the user can 'work out' any inferences the system can calculate. Since QUASI's goal is **not** the licensing of particular inferences -- but rather the generation of responses that will not license false inferences -- these assumptions are actually conservative; the system will err on the side of caution if it anticipates inferences the user would not be in a position to draw. Second, QUASI does not adopt an independent strategy for determining when and which expressions and posets are salient, but must rely upon interaction with the user to select from expressions appearing on posets known in the system. Third, QUASI does not maintain or assume information from a discourse history, which might modify its decisions on when a false inference is likely -- i.e., when the user has already demonstrated knowledge of some licenseable inference. However, the core of the thesis, the scalar implicature conventions and algorithms for the calculation of scalar implicatures when one or more expressions are salient, has been incorporated into QUASI.

7.2.1. Overview of the System

The implemented system is depicted schematically in Figure 7-1: Major system components are identified by capital letters, information available in the system by small capitals, user inputs by italics, outputs by bold italics, and information passed through the system by normal font.

QUASI's semantic translation module creates a semantic representation of a user query which its direct response retrieval module uses both to retrieve and to represent the direct response in terms of system beliefs. This module also determines whether or not the query is a yes-no

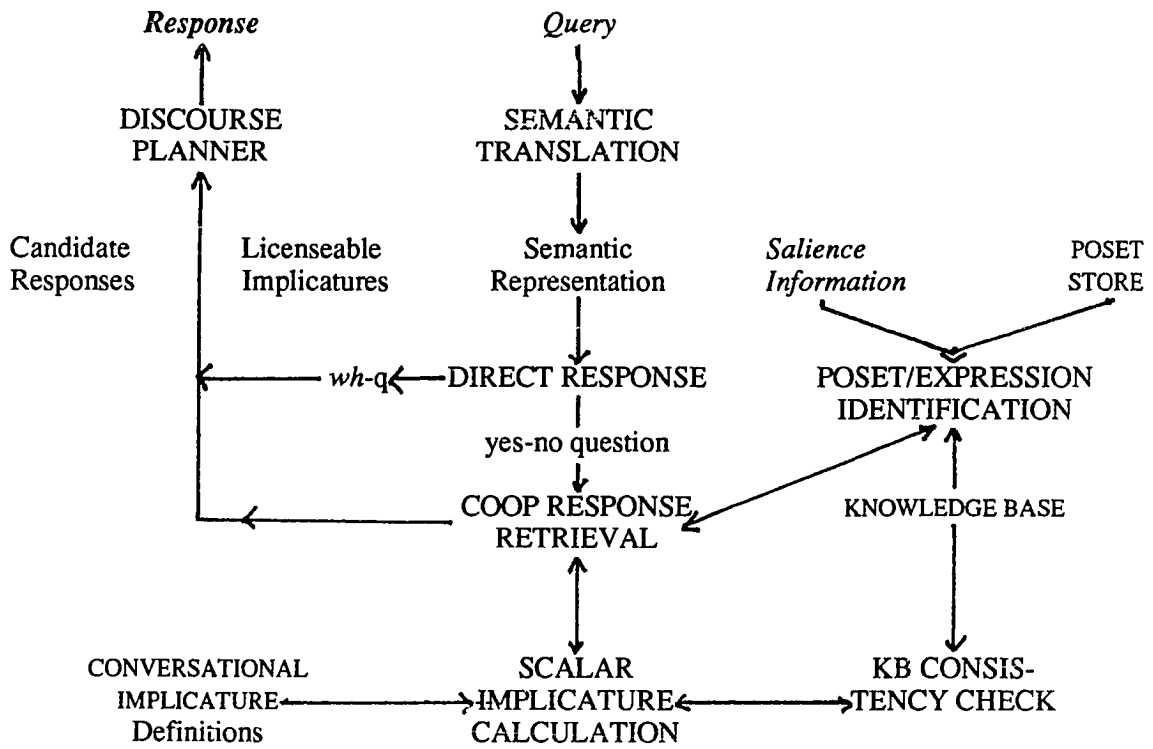


Figure 7-1: The QUASI System

question. If so, it is passed to another module which guides the generation of one or more candidate cooperative responses. Representations of responses to *wh*-questions are passed directly to the discourse planner.

The cooperative response retrieval module (CRR) first acquires information about salient expressions and posets from the poset/ expression identification module, which provides the user with a list of possible expression/ poset pairs from which to choose and returns those choices to CRR. CRR passes this information and the representation of the direct response to a scalar implicature calculator, which determines licenseable scalar implicatures from it. If implicatures are licenseable, CRR sends them to the KB consistency checker. If this check discovers licenseable implicatures which are inconsistent with the system KB, CRR will propose alternate responses which will themselves be checked for consistent implicatures. When all true responses which license only scalar implicatures which are consistent with the KB have been found, these are output to the discourse planner as candidate responses.

QUASI also provides some help facilities as well as tracing, which will be demonstrated in examples below.

7.2.2. Domain

The domain chosen to demonstrate the QUASI system was a mail system currently in use at the University of Pennsylvania.¹⁶¹ Currently, the system KB only includes information about boolean switches in this mail system and how they are used, including the mail commands which can set them, the arguments these commands may take, default switch values, and values that are associated with each switch. For example, the switch 'auto-archive' has the default setting 'off' and takes an associated value which is the name of an archive file; the default setting for this value is the filename 'archive.msg'. Both the switch value and that of its associated values can be set with 'setswitch' or 'msetswitch'. The latter command cannot be used in mailinit files (i.e., noninteractively) and is employed for temporary switch-setting.

This domain was chosen because, while it is small and well-defined, its conceptual model naturally reflects a wide variety of partial orderings discussed in Chapter 5, including sets, whole/ part relations, entity/ attribute orderings, and simple taxonomic hierarchies, as well as the canonical orderings, as well as process or prerequisite relationships. Although not all these

¹⁶¹This system was written by Sharon E. Perl.

orderings have in fact been represented in the current implementation,¹⁶² the potential for extensibility is clear. Also, while QUASI currently operates only on information about switches in the mail domain, the KB could readily be extended to other aspects of the mail system. Additionally, while switches make up an important component of this mail system, they are little understood by the mail user community. So, a question-answering system in this domain has potential for practical use.

7.2.3. Knowledge Representation

7.2.3.1. Epistemic Representation

Because of Prolog's closed-world assumption, it was necessary to make a major simplification of the epistemic force of licensed implicatures. Since Prolog, like most systems, assumes 'negation as failure', ignorance cannot be represented in it; it supports only a standard two-valued logic. Everything provable in the KB is true and all else is assumed false. This assumption is easily translated in terms of system beliefs as: the system believes everything provable in the KB and it believes that it believes -- otherwise it believes things false.

```
bel(system, Sent1) :- Sent1.
bel(system, ~Sent1) :- \+Sent1163.
```

Since ignorance is not representable in the system KB, the representation of scalar implicatures as epistemic disjunctions of belief or ignorance must be simplified to simple belief. Where scalar implicatures are represented as disjunctions in a three-valued logic -- S believes $p_i \vee S$ does not know whether p_i or S believes $\neg p_i \vee S$ does not know whether $\neg p_i$ -- they are in effect reduced to their first conjuncts in a two-valued logic. So $\sim\text{bel}(\text{Speaker}, \text{Sent1})$ is true iff Sent1 is false -- just as $\text{bel}(\text{Speaker}, \sim\text{Sent1})$, and $\sim\text{bel}(\text{Speaker}, \sim\text{Sent1})$ is true iff Sent1 is true -- as $\text{bel}(\text{Speaker}, \text{Sent1})$. While a more expressive system would permit a more accurate reflection of the theory, in fact, little is lost in this translation: QUASI returns representations of system beliefs which then may be simplified in their interpretation for a given system -- or not. Too, since few systems support a three-valued logic,¹⁶⁴ it seems reasonable to test the theory in a more restricted environment.

¹⁶²Of the canonical orderings, modals and ordinals are not represented. Also, there is no way to represent temporal orderings, and no special support for the representation of entity/ attribute or whole/ part orderings.

¹⁶³ Sent1 is not provable.

¹⁶⁴See [Levesque 84] for a recent proposal.

7.2.4. Representing Posets

Although a new domain was modeled for QUASI, the system as a whole has been designed to operate on an existing KB. The chief expense involved in transferring QUASI to other domains is the need to define particular ways in which posets may be derived from a particular knowledge representation. For standard Prolog databases, however, such redefinition will be minimal, since QUASI includes its own representations of the canonical orderings and derives posets from relations explicit or implicit in the domain by its own poset derivation rules. Although some meta-level translation must take place when, for example, existing KB's include more sophisticated knowledge representations which will facilitate poset inference. Also, if any of the canonical orderings are not supported by a new KB, it will be misleading to include these among user options when salient expressions and posets are selected.

7.2.4.1. Canonical Orderings

Posets are represented explicitly in a domain-independent poset store or are derived from the domain KB by rules represented in that store. Explicitly represented posets include the canonical orderings such as the quantifiers, cardinals,¹⁶⁵ and logical connectives, as described in Chapter 5. These orderings are represented as 4-tuples: a poset label, which is used to identify the ordering uniquely for system purposes and for interaction with the user when appropriate posets are chosen during the query analysis; an arbitrarily complex specification of the set of referents in the poset; the metric partially ordering this set of referents, which must itself be defined in the system; and a keyword specifying the method of derivation for the poset. For predefined canonical orderings this is simply 'canonical'. Posets are accessed and manipulated with predicates that 'specify' (or test) the membership of expressions in their set of referents, their ordered pairs, and the incomparability of items in the poset. These definitions are used to define higher, lower, and alternate sentences as described in Chapter 4.

7.2.4.2. Derived Posets

However, the approach to scale outlined in Section 3.3 should support domain independence; in particular the scale derivation rules proposed there should allow interface to existing knowledge bases without extensive recoding. That is, no specific knowledge representation should be required for this extension. Relationships and values present in the knowledge base should be interpretable as scales as required by a given query. While initial translation from such relationships to scales might cut response time for given queries, it seems

¹⁶⁵Limited here arbitrarily to 1-5.

likely that considerations of space and initial cost will indicate that a lazy evaluation of scales will be preferable.

Posets are derived from the KB by means of a set of poset derivation rules stored in the poset store. The essence of poset derivation is simply the identification of partial orderings from relations represented explicitly or implicitly in the KB. While only some poset derivation rules have been implemented in QUASI, I will discuss others which might easily be added.

In any representation, certain distinguished predicates must always define relations which, as I have noted in Chapter 5, represent posets. The predefined predicate *setof* in most Prolog implementations will always define posets ordered by set inclusion. In the mail switches KB used in this application, an *isa* predicate identifies classification hierarchies, which also represent partial orderings of their members. Similar meta-level predicates might be defined for other knowledge representations and added to the poset derivation rules in the poset store.

Posets can be derived from any KB predicate of two or more arguments as follows: If a predicate P defines a relation on any two of its arguments that is irreflexive, asymmetric, and transitive or that is reflexive, antisymmetric, and transitive, then P defines a poset on these arguments. Posets defined by set inclusion can also be derived from the range of values a single argument of a predicate which does not satisfy these conditions may take on. Boolean combinations of predicates for which the ranges of some argument intersect may also define similar orderings.

7.2.4.3. Storing Identified Posets

Once a poset implicit in the KB has been identified, its representation tuple is stored in the poset store for later access -- to minimize redundant poset inference. Modules which access the poset store look first at posets which have already been derived. It seems likely too that orderings deemed salient should be so marked by the discourse planner in a more sophisticated application, for future identification of salient posets. This is not done in QUASI, since salience information is provided by the user. The availability and compatibility of posets is testable in QUASI but results are not used in the absence of a user model.

7.2.5. Input and Semantic Processing

QUASI accepts either wh or yes-no questions in English as input, although only the latter can generate 'cooperative' responses. The ability to handle wh-questions was provided to support more natural behavior in the system.

Legal queries are given a semantic representation in first-order predicate calculus using a definite-clause grammar extended from that presented in [Pereira 80].¹⁶⁶ At present, the grammar considerably limits the range of queries that can be handled by QUASI: for example, conjunction is only available at the sentence level.

The parse produces a simple first-order semantic representation of the query which corresponds to the notion of SENTENCE discussed in Chapter 4. So, a query like (281a) will be represented as (281b).

- (281)
 a. Are any switches boolean?
 b. exists(X, isa(switch, X) & boolean(X))

In trace mode, QUASI produces the following:

|: |: **Are any switches boolean?**

The semantic representation of your query is:

exists(_247, isa(switch, _247) & boolean(_247))

Singular definites and indefinites are represented as existentials and plural definites are universally quantified, following [Pereira 80, Warren 81] and cardinals are also treated as determiners, following [Jackendoff 68]. Also, the first successful representation for each query is accepted: So, for example, no scope ambiguities are recognized; in effect, the first quantifier in the surface order of the input query will have scope over subsequent quantifiers.

7.2.6. Direct Response Retrieval

Once the semantic representation of a query has been determined, QUASI proceeds to determine a direct response to the query and then to describe that response in terms of the system's beliefs, e.g., '*BEL(system, p_i)*', where *p_i* is the direct response to the user's query. For *wh*-questions, QUASI simply returns this information to the discourse planner. For yes-no questions,¹⁶⁷ QUASI proceeds to identify a cooperative response, passing the direct response to a cooperative-response retrieval module. With the direct response plus information about salient expressions and posets, this module controls the generation and testing of possible alternate responses. Saliency information is acquired interactively.

¹⁶⁶This type of grammar permits the generation of a semantic representation directly, without constructing an intermediate parse tree.

¹⁶⁷Since *how* and *why* queries are not currently supported by the grammar, QUASI treats non-*wh*-questions as yes-no questions.

7.2.7. Identifying Salient Expressions and Posets

From the semantic representation of the query, QUASI identifies all the legal subexpressions of the query's semantic representation. As I noted in Section 4.1.2.2, a subexpression of a sentence p_i can be any constant, predicate, logical operator (including the epistemic operators), connective, or quantifier symbol of p_i or any wff contained in p_i that is not identical to p_i . So, the following subexpressions are contained in the sentence 'exists(Y, isa(switch, Y) & boolean(Y))', and will be so identified

```
exists
isa(switch, Y) & boolean(Y)
&
isa(switch, Y)
isa
switch
boolean(Y)
boolean
```

by QUASI's poset/expression identification component. In effect, the module excludes the sentence itself, variables and non-wffs other than predicates, constants, and logical symbols from the set of possible subexpressions of a sentence.

Once subexpressions are identified, QUASI checks each for membership in the set on which some partial ordering explicit or implicit (See Section 7.2.4.) in the poset store or KB is defined. For the above query, QUASI will currently identify the following expression/ poset pairs:

Please choose salient expressions and a salient ordering for each from the following list, forming a list of lists (e.g., [[exists, quantifiers],[3, ucardinals]]):¹⁶⁸

Expression	Poset
&	[andor]
exists	[quantifiers]

|: |: [[exists, quantifiers]].

QUASI currently asks the user to select one or more expressions from those identified as associable with posets known to the system. For each selected expression, the user must choose a single salient poset, as above, forming a list of lists. This list of lists will be passed on to the scalar implicature calculation module as the list of salient expression/ poset pairs. Salient

¹⁶⁸'Ucardinals' designate the cardinals ordered by '<'.

expression/ poset pair lists might also be input from the discourse planner if QUASI were incorporated into a sophisticated natural-language processing system.

7.2.8. Calculating Cooperative Responses

The generation of cooperative responses is driven by the calculation of potential scalar implicatures and the testing of these implicatures against the system KB. Where a candidate response might license inferences about system beliefs which are inconsistent with the system's KB, alternate responses are identified and themselves tested for the scalar inferences they might license. A simple example of this process of testing and retesting with accompanying trace information from QUASI follows:¹⁶⁹

```
|: TRACE.
|: |: ARE ANY SWITCHES BOOLEAN?
```

The semantic representation of your query is:

```
exists(_247, isa(switch, _247) & boolean(_247))
```

The direct response is: true

Please choose salient expressions and a salient ordering for each from the following list, forming a list of lists (e.g., [[exists, quantifiers], [3, ucardinals]]):

Expression	Poset
&	[andor]
exists	[quantifiers]

```
|: |: [[EXISTS, QUANTIFIERS]].
```

If the following expressions and orderings are salient,

Expression	Poset
exists	[quantifiers]

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, exists(_247, isa(switch, _247) & boolean(_247)))
```

¹⁶⁹User input has been altered to upper case for clarity.

are:

```
[~bel(system, all(_247, isa(switch, _247) => boolean(_247)))]
```

The following licenseable implicature is inconsistent with the system KB:

```
~bel(system, all(_247, isa(switch, _247) => boolean(_247)))
```

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, all(_247, isa(switch, _247) => boolean(_247)))
```

are:

```
[]
```

This assertion is consistent with the system KB.

Output this cooperative response to the query:

```
bel(system, all(_247, isa(switch, _247) => boolean(_247)))
```

In this sample query analysis, the user first puts QUASI into trace mode and then enters a query. QUASI first returns the semantic representation of the input query and then identifies the direct response. It then asks the user to choose a list of salient expressions and posets from those expressions identified in the query for which QUASI can locate posets from the poset store or poset derivation rules. Here the user chooses the quantifiers as salient. By accessing the scalar implicature conventions in its definition of conversational implicature, QUASI determines that, if the quantifiers are salient, the scalar implicatures that could be licensed by the assertion of the direct response to the user's query (that the system does indeed believe that some switches are boolean), would be the implicature that the system does not believe that all switches are boolean. QUASI then determines that this belief is inconsistent with its KB -- for, in fact, it does believe that all switches are boolean. So, it tests an alternate response, derived by replacing the salient expression in the response it is currently testing -- *exists* -- with another value in the salient ordering -- *all*. It finds that no scalar implicatures whatsoever will be licensed by the assertion that the system believes all switches to be boolean, given that only the quantifiers are salient; *all* is in fact the highest value in this ordering, and affirmation of a value can license scalar implicatures only about still higher values. QUASI finds that asserting 'bel(system, all(_247, isa(switch, _247) => boolean(_247)))' is consistent with the KB; it is true itself and it licenses no (and thus, no false) scalar implicatures. So, it proposes this cooperative response to the discourse planner.

Note that, if no expressions are chosen as salient, the cooperative response will be identical with the direct response, as in the following interaction where the same query elicits a different 'cooperative' response:

|: |: ARE ANY SWITCHES BOOLEAN?

The semantic representation of your query is:

exists(_4211, isa(switch, _4211) & boolean(_4211))

The direct response is: true

Please choose salient expressions and a salient ordering for each from the following list, forming a list of lists (e.g., [[exists, quantifiers],[3, ucardinals]]):

Expression	Poset
&	[andor]
exists	[quantifiers]

|: |: [].

If the following expressions and orderings are salient,

Expression	Poset
------------	-------

The scalar implicatures that could be licensed by the assertion of:

bel(system, exists(_4211, isa(switch, _4211) & boolean(_4211)))

are:

[]

This assertion is consistent with the system KB.

Output this cooperative response to the query:

bel(system, exists(_4211, isa(switch, _4211) & boolean(_4211)))

Since the user chooses no expression/ poset pairs from the proffered list, no scalar implicatures can be calculated from the assertion of the direct (affirmative) response. Licenseable implicatures are, then, trivially, consistent with the KB, so the direct response is deemed cooperative by QUASI.

7.2.8.1. Responses When Multiple Expressions are Salient

When multiple expression/ poset pairs are selected as salient, QUASI identifies scalar implicatures as described in Section 6.4.1, as illustrated in the following exchange:

|: DO ANY COMMANDS TAKE 2 ARGUMENTS?

The semantic representation of your query is:

```
exists(_11127, isa(command, _11127) & xnumberof(_11141,
isa(arg_type, _11141) & take(_11127, _11141), 2))
```

The direct response is: true

Please choose salient expressions and a salient ordering for each from the following list, forming a list of lists (e.g., [[exists, quantifiers], [3, ucardinals]]):

Expression	Poset
2	[dcardinals, ucardinals]
&	[andor]
exists	[quantifiers]

|: |: [[EXISTS, QUANTIFIERS], [2, UCARDINALS]].

If the following expressions and orderings are salient,

Expression	Poset
exists	[quantifiers]
2	[ucardinals]

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, exists(_11127, isa(command, _11127) &
xnumberof(_11141, isa(arg_type, _11141) & take(_11127,
_11141), 2)))
```

are:

```
[~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
2))), ~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
3))), ~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
```

```
4))), ~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
5))), ~bel(system, exists(_11127, isa(command, _11127) &
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
3))), ~bel(system, exists(_11127, isa(command, _11127) &
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
4))), ~bel(system, exists(_11127, isa(command, _11127) &
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
5)))]
```

The following licenseable implicature is inconsistent with the system KB:

```
~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
2)))
```

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
2)))
```

are:

```
[~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
3))), ~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
4))), ~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
5)))]
```

The following licenseable implicature is inconsistent with the system KB:

```
~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
3)))
```

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
3)))
```

are:

```
[~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
4))), ~bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
5)))]
```

This response is consistent with the system KB.

Output this cooperative response to the query:

```
bel(system, all(_11127, isa(command, _11127) =>
xnumberof(_11141, isa(arg_type, _11141) & take(_11127, _11141),
3)))
```

The questioner selects two expression/ poset pairs as salient for the analysis of this query, the quantifiers and the cardinals ordered by '<', termed here the 'ucardinals'. QUASI first identifies all the implicatures licensable when both *exists* and 2 are treated as salient expressions. Checking these potential implicatures against its KB, QUASI finds the first inconsistent implicature '*~bel(system, all(_6278, isa(command, _6278) => xnumberof(_6292, isa(arg_type, _6292) & take(_6278, _6292), 2)))*' and attempts to find a more cooperative response. This process of checking potential responses and finding an inconsistent scalar implicature is carried out once more for the candidate response '*bel(system, all(_6278, isa(command, _6278) => xnumberof(_6292, isa(arg_type, _6292) & take(_6278, _6292), 2)))*'. Finally, QUASI identifies a candidate response which licenses no false scalar inferences, '*bel(system, all(_6278, isa(command, _6278) => xnumberof(_6292, isa(arg_type, _6292) & take(_6278, _6292), 3)))*'. So, QUASI outputs this candidate cooperative response to the discourse planner.

Note that changing the order in which expression/ poset pairs are considered produces no change in QUASI's output:

|: DO ANY COMMANDS TAKE 2 ARGUMENTS?

The semantic representation of your query is:

```
exists(_18636, isa(command, _18636) & xnumberof(_18650,
isa(arg_type, _18650) & take(_18636, _18650), 2))
```

The direct response is: true

Please choose salient expressions and a salient ordering for

each from the following list, forming a list of lists (e.g., [[exists, quantifiers], [3, ucardinals]]):

Expression	Poset
2	[dcardinals, ucardinals]
&	[andor]
exists	[quantifiers]

|: |: [[2, UCARDINALS], [EXISTS, QUANTIFIERS]].

If the following expressions and orderings are salient,

Expression	Poset
2	[ucardinals]
exists	[quantifiers]

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, exists(_18636, isa(command, _18636) &
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
2)))
```

are:

```
[~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
2))), ~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
3))), ~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
4))), ~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
5))), ~bel(system, exists(_18636, isa(command, _18636) &
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
3))), ~bel(system, exists(_18636, isa(command, _18636) &
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
4))), ~bel(system, exists(_18636, isa(command, _18636) &
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
5)))]
```

The following licenseable implicature is inconsistent with the system KB:

```
~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
```

2)))

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
2)))
```

are:

```
[~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
3))), ~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
4))), ~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
5)))]
```

The following licenseable implicature is inconsistent with the system KB:

```
~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
3)))
```

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
3)))
```

are:

```
[~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
4))), ~bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
5)))]
```

This response is consistent with the system KB.

Output this cooperative response to the query:

```
bel(system, all(_18636, isa(command, _18636) =>
xnumberof(_18650, isa(arg_type, _18650) & take(_18636, _18650),
3)))
```

The scalar implicatures identified in the above exchange and the candidate responses proposed are identical to those of the previous exchange. This demonstrates that the calculation of scalar implicatures is not sensitive to the ordering of salient expressions.

7.2.8.2. Changing Salient Expression Changes Output

Although the ordering of salient expressions does not change the calculation of scalar implicatures, the selection of one expression as salient over another certainly does alter the implicatures that will be identified and the cooperative responses suggested by QUASI. Consider:

|: DO ANY COMMANDS TAKE 2 ARGUMENTS?

The semantic representation of your query is:

```
exists(_21330, isa(command, _21330) & xnumberof(_21344,
isa(arg_type, _21344) & take(_21330, _21344), 2))
```

The direct response is: true

Please choose salient expressions and a salient ordering for each from the following list, forming a list of lists (e.g., [[exists, quantifiers],[3, ucardinals]]):

Expression	Poset
2	[dcardinals, ucardinals]
&	[andor]
exists	[quantifiers]

|: |: [[EXISTS, QUANTIFIERS]].

If the following expressions and orderings are salient,

Expression	Poset
exists	[quantifiers]

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, exists(_21330, isa(command, _21330) &
xnumberof(_21344, isa(arg_type, _21344) & take(_21330,
_21344), 2)))
```

are:

```
[~bel(system, all(_21330, isa(command, _21330) =>
xnumberof(_21344, isa(arg_type, _21344) & take(_21330,
_21344), 2)))]
```

The following licenseable implicature is inconsistent with the system KB:

```
~bel(system, all(_21330, isa(command, _21330) =>
xnumberof(_21344, isa(arg_type, _21344) & take(_21330,
_21344), 2)))
```

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, all(_21330, isa(command, _21330) =>
xnumberof(_21344, isa(arg_type, _21344) & take(_21330,
_21344), 2)))
```

are:

[]

This assertion is consistent with the system KB.

Output this cooperative response to the query:

```
bel(system, all(_21330, isa(command, _21330) =>
xnumberof(_21344, isa(arg_type, _21344) & take(_21330,
_21344), 2)))
```

When a quantifier ordering is salient, QUASI will suggest 'bel(system, all(_21330, isa(command, _21330) => xnumberof(_21344, isa(arg_type, _21344) & take(_21330, _21344), 2)))' as a cooperative response. However, when the cardinals are salient, QUASI will calculate a different cooperative response:

! : DO ANY COMMANDS TAKE 2 ARGUMENTS?

The semantic representation of your query is:

```
exists(_26492, isa(command, _26492) & xnumberof(_26506,
isa(arg_type, _26506) & take(_26492, _26506), 2))
```

The direct response is: true

Please choose salient expressions and a salient ordering for each from the following list, forming a list of lists (e.g., [[exists, quantifiers], [3, ucardinals]]):

Expression	Poset
2	[dcardinals, ucardinals]
&	[andor]
exists	[quantifiers]

! : ! : [[2, UCARDINALS]].

If the following expressions and orderings are salient,

Expression	Poset
2	[ucardinals]

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, exists(_26492, isa(command, _26492) &
xnumberof(_26506, isa(arg_type, _26506) & take(_26492,
_26506), 2)))
```

are:

```
[~bel(system, exists(_26492, isa(command, _26492) &
xnumberof(_26506, isa(arg_type, _26506) & take(_26492,
_26506), 3))), ~bel(system, exists(_26492, isa(command, _264 92)
& xnumberof(_26506, isa(arg_type, _26506) & take(_26492,
_26506), 4))), ~bel(system, e xists(_26492, isa(command, _26492)
& xnumberof(_26506, isa(arg_type, _26506) & take(_264 92,
_26506), 5)))]
```

The following licenseable implicature is inconsistent with the system KB:

```
~bel(system, exists(_26492, isa(command, _26492) &
xnumberof(_26506, isa(arg_type, _26506) & take(_26492,
_26506), 3)))
```

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, exists(_26492, isa(command, _26492) &
xnumberof(_26506, isa(arg_type, _26506) & take(_26492,
_26506), 3)))
```

are:

```
[~bel(system, exists(_26492, isa(command, _26492) &
```

```
xnumberof(_26506, isa(arg_type, _26506) & take(_26492,
_26506), 4)), ~bel(system, exists(_26492, isa(command, _26492)
& xnumberof(_26506, isa(arg_type, _26506) & take(_26492,
_26506), 5))))]
```

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, exists(_26492, isa(command, _26492) &
xnumberof(_26506, isa(arg_type, _26506) & take(_26492,
_26506), 3)))
```

are:

```
[~bel(system, exists(_5451, isa(command, _5451) &
xnumberof(_5465, isa(arg_type, _5465) & take(_5451, _5465),
4)), ~bel(system, exists(_5451, isa(command, _5451) & xnum
berof(_5465, isa(arg_type, _5465) & take(_5451, _5465), 5)))]
```

This assertion is consistent with the system KB.

Output this cooperative response to the query:

```
bel(system, exists(_26492, isa(command, _26492) &
xnumberof(_26506, isa(arg_type, _26506) & take(_26492, _26506),
3)))
```

7.2.8.3. When Several Responses are Cooperative

In some cases, more than one response to a query may be cooperative, in the sense that QUASI predicts it will license no scalar implicatures inconsistent with the KB. One such case is illustrated below:¹⁷⁰

|: IS MSETSWITCH NONINTERACTIVE AND IS SETSWITCH NONINTERACTIVE?
The semantic representation of your query is:

```
noninteractive(msetswitch) & noninteractive(setswitch)
```

The direct response is: false

Please choose salient expressions and a salient ordering for each from the following list, forming a list of lists (e.g., [[exists, quantifiers],[3, ucardinals]]):

¹⁷⁰Where '#' designates logical disjunction and '&', conjunction.

Expression	Poset
&	[andor]
[msetswitch]	[switch_commands]
[setswitch]	[switch_commands]

! : ! : [[&, andor]].

If the following expressions and orderings are salient,

Expression	Poset
&	[andor]

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, ~(noninteractive(msetswitch) &
noninteractive(setswitch)))
```

are:

```
[~bel(system, ~(noninteractive(msetswitch) #
noninteractive(setswitch)))]
```

This response is consistent with the system KB.

The scalar implicatures that could be licensed by the assertion of:

```
bel(system, noninteractive(msetswitch) #
noninteractive(setswitch))
```

are:

```
[~bel(system, noninteractive(msetswitch) &
noninteractive(setswitch))]
```

This response is consistent with the system KB.

Choose one of these candidate cooperative responses:

```
bel(system, noninteractive(msetswitch) #
noninteractive(setswitch))
```

```
bel(system, ~(noninteractive(msetswitch) &
noninteractive(setswitch)))
```

In this exchange, where the user has chosen the logical connectives as salient, QUASI finds that two candidate responses license no misleading scalar implicatures and outputs both. In each case, QUASI presents those implicatures that are licenseable via the utterance of each.

7.2.9. Outputs

Examples of system output when QUASI is in trace mode have been presented above. When not in trace mode, QUASI returns only a set of candidate cooperative responses to the terminal.¹⁷¹ The first example, when QUASI is not in trace mode, thus runs as follows:

```
|: ARE ANY SWITCHES BOOLEAN?
```

```
Please choose salient expressions and a salient ordering for
each from the following list, forming a list of lists (e.g.,
[[exists, quantifiers],[3, ucardinals]]):
```

Expression	Poset
&	[andor]
exists	[quantifiers]

```
|: |: [[EXISTS, QUANTIFIERS]].
```

```
bel(system, all(_218, isa(switch, _218) => boolean(_218)))
```

Given the expressions and posets selected as salient, each member of this set satisfies the criterion of **not** licensing misleading implicatures. From the returned set of candidate appropriate responses plus information external to QUASI, it is intended that the discourse planner select an actual response. The system also outputs the scalar implicatures which each response (if output) might license. If QUASI were embedded in a larger natural-language question-answering system, the choice of one of these outputs should trigger the marking of the corresponding set of licenseable implicatures as communicated to the user implicitly.

7.2.10. Future Research

As noted above, some aspects of the theory presented in previous chapters are not implemented in QUASI. I have mentioned some of these in the discussion above. Too, although classification hierarchies are represented in the switches domain, no attempt has been

¹⁷¹Other information is of course available to the discourse planner.

made to limit salient hierarchies using the 'entry level' notion discussed in Section 6.3.2.1, in large part because the hierarchies modeled here are not infinite and because there is no empirical evidence allowing us to identify entry levels in an electronic mail domain. At present, QUASI generates only semantic representations of candidate responses, since this module is not actually suggested as a stand-alone natural-language interface. However, for demonstrational purposes, it might be desirable to return actual English output.

7.2.11. Summary

The system described above demonstrates that the theory of scalar implicature presented in this thesis is both computational and of some potential use for natural-language generation. Whether or not the approach to response generation suggested above has any cognitive reality, it provides a convenient fiction which accounts for naturally occurring data in a reasonable way. However, the real computational benefit which results from the calculability of scalar implicature is a more general one: the calculation of what may be conveyed and what may be inferred -- beyond the literal propositional content of generated or interpreted text.

In this application, I have chosen to focus on the use of a theory of scalar implicature in natural-language generation and have limited this application to the generation of cooperative responses to yes-no questions. However, the formalism and, indeed, the core of the implementation may easily be adapted to other uses in generation -- and to the interpretation of user input. The anticipation of potential scalar implicatures licensable via system output exactly parallels the analysis of implicatures possibly licensed by user input. Even in systems that do not support natural-language interaction, the notion of scalar implicature proposed here should prove useful, since these inferences are analyzed from formal language representations of natural language input.

CHAPTER VIII

Conclusions

In this thesis I have outlined a computational approach to the study of conversational implicature. I have identified one type of generalized conversational implicature, scalar implicature, and demonstrated how these implicatures can be calculated. I have further shown how the calculation of conversational implicatures can be applied to a particular problem in natural-language question-answering, the specification of cooperative responses to yes-no questions.

In Chapter 2, I described obstacles to a definition of conversational implicature, but argued for a set of necessary and sufficient conditions on this phenomenon based upon speaker intention, mutual belief in speaker cooperativity and in hearer ability to 'work out' an inference, as well as the cancelability, nondetachability, and reinforceability of that inference. In Chapter 3, I proposed a reinterpretation of the class of GENERALIZED CONVERSATIONAL IMPLICATURES, which include what have previously been identified as generalized quantity implicatures. I then examined attempts to define such quantity implicatures by Horn, Harnish, and Gazdar and the obstacles which have stymied them. I showed that p_j represents a quantity implicature when the 'working out' of p_j is dependent upon speaker's and hearer's perception of some salient metric ordering speaker's utterance with other potential utterances. Finally, I introduced scalar implicature and distinguished it from these earlier definitions of quantity implicature. In particular, I showed how scalar implicatures are supported by perception of a wider variety of orderings, including set inclusion, whole/ part, entity/ attribute, and type/ subtype, as well as previously noted 'canonical' orderings, such as the quantifiers, cardinals, and modals. Scalar implicature is defined not only for the affirmation of values in such an ordering, but for the denial or assertion of ignorance regarding such values. Scalar implicatures may represent inferences about higher, lower, or alternate values to those mentioned. From an intuitive description of this phenomenon, I proposed a set of scalar implicature conventions which, for a given utterance and given salient orderings, permit the anticipation or interpretation of scalar implicatures.

In Chapters 4-6, I showed how these conventions may be used to calculate licensable scalar implicatures. Chapter 4 dealt largely with representational issues: I adapted Gazdar's syntactic approach to the ranking of utterances via the presence in their semantic representations

of certain rankable expressions and discussed the epistemic force appropriately assigned to licensed implicatures. In Chapter 5, I first described the varied types of ordering relations that can support the ranking of utterances for the licensing of scalar implicature. I next argued that all and only those relations that can be modeled as partial ordering relations or POSETS support scalar implicature. Partial orderings of higher, lower, and alternate expressions can indirectly order semantic representations of utterances similarly, so that the appropriate scalar implicature convention can be accessed.

In Chapter 6 I related the conventional aspects of scalar implicature to the interpretation of particular inferences in context. I proposed a definition of ordering COMPATIBILITY and suggest how this concept may be used to relax the requirement that speaker believe some partial ordering is mutually believed salient for speaker and hearer. To identify salient expressions and salient partial orderings in a discourse, I proposed use of FOCUS/ TOPIC/ CENTERING information and suggested how current work on these phenomena might be adapted/ extended to this end. I examined certain strategies for marking and identifying focus, including syntactic, prosodic, and semantico-pragmatic cues. I also proposed ways to limit potentially infinite orderings to locally salient orderings using concepts of BASIC and ENTRY LEVEL for classification hierarchies and ways to distinguish between duals when one is believed salient. I also suggested possibilities for future research in the identification of salient orderings. I extended the scalar implicature formalism to accommodate scalar implicatures arising from utterances in which more than one expression is salient.

In Chapter 7, I described an application of the theory defined in previous chapters to the calculation of indirect responses to yes/ no questions. I described QUASI, a natural-language interface to a database which generates indirect and qualified direct responses to yes/ no questions about switches in a computer mail system. Given a semantic representation of the desideratum of a yes/ no question, QUASI retrieves direct responses from the knowledge base and calculates the scalar implicatures that could be licensed by their generation. Where any such inferences are inconsistent with the system's knowledge base, QUASI calculates alternative responses that will not license such false inferences.

While I have demonstrated one practical use for the calculation of of one type of conversational implicature, the undemonstrated uses are clearly even more significant. If we can calculate conversational implicature, then we can access a heretofore inaccessible level of utterance meaning computationally. For natural-language understanding systems, this means that systems can understand far more from user input than they have previously understood. For natural-language generation systems, it means they can anticipate far more accurately the effect that system output will have on its audience. And, in less practical terms, it means that we are one step closer to understanding utterance meaning.

BIBLIOGRAPHY

- [Allen 80] Allen, James F. and C. Raymond Perrault. Analyzing Intention in Utterances. *Artificial Intelligence* 15:143-178, 1980.
- [Allwood 72] Allwood, J. Negation and the strength of presuppositions. *Gothenburg: Logical Grammar Report 2*, 1972.
- [Anglin 79] Anglin, J. *The Growth of Word Meaning*. MIT Press, Cambridge MA, 1979.
- [Aqvist ??] Aqvist, Lennart. A New approach to the logical theory of interrogatives. University of Uppsala.
- [Berlin 69] Berlin, B. and Kay, P. *Basic Color Terms: Their Universality and Evolution*. Berkeley University Press, Berkeley CA, 1969.
- [Bock 77] Bock, J. Kathryn. The Effect of a pragmatic presupposition on syntactic structure in question-answering. *Journal of Verbal Learning and Verbal Behavior* 16:723-734, 1977.
- [Bolinger 78] Bolinger, D. Yes-No Questions Are Not Alternative Questions. In Hiz, H. (editor), *Questions*, pages 87-105. Reidel, Dordrecht (Neth), 1978.
- [Brownell 78] Brownell, J. J. *Picture perception and semantic memory*. PhD thesis, Johns Hopkins University, 1978.
- [Carberry 83] Carberry, Sandra. Tracking User Goals in an Information-Seeking Environment. In *Proceedings of the Third National Conference on Artificial Intelligence*, pages 59-63. 1983.
- [Carbonnell 70] Carbonnell, Jaime R. and Collins, Allan M. *Mixed-Initiative Systems For Training and Decision-Aid Applications*. Technical Report ESD-TR-70-373, Electronics Systems Division, Laurence G. Hanscom Field, US Air Force, Bedford MA, November, 1970.
- [Chafe 76] Chafe, W. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, C. (editor), *Subject and Topic*. Academic Press, New York, 1976.
- [Chomsky 72] Chomsky, N. *Studies on semantics in generative grammar*. Mouton, The Hague, 1972.
- [Clark 81] Clark, H., and Marshall, C. Definite reference and mutual knowledge. In Joshi, A. K., Webber, B., and Sag, I. (editors), *Elements of discourse understanding*. Cambridge University Press, Cambridge, 1981.
- [Cohen 79] Cohen, P. R. and C. R. Perrault. Elements of a Plan-Based Theory of Speech Acts. *Cognitive Science* 3:177-212, 1979.

[Conklin and McDonald 82]

Conklin, E.J., and McDonald, D.D. Salience: The Key to the Selection Problem in Natural Language Generation. In *Proceedings of the 20th Annual Meeting*, pages 129-135. Association of Computational Linguistics, University of Toronto, Toronto, Canada, June, 1982.

[Corella 84] Corella, F. Semantic Retrieval and Levels of Abstraction. In L. Kerschberg (editor), *Proceedings of the First Int'l Workshop on Expert Database Systems*, pages 397-420. Kiawah Island SC, October, 1984.

[Culicover 83] Culicover, Peter W., and Michael Rochemont. Stress and focus in English. *Language* 59(1):123-165, March, 1983.

[Fauconnier 75] Fauconnier, G. Polarity and the Scale Principle. In *Papers of the Eleventh Regional Meeting*, pages 188-199. Chicago Linguistic Society, 1975.

[Fauconnier 79] Fauconnier, G. Implication Reversal in a Natural Language. In Guenther, F. and Schmidt, S. J. (editors), *Formal Semantics and Pragmatics for Natural Language*, pages 289-301. D. Reidel Publishing Co., Dordrecht, 1979.

[Fillmore 83] Fillmore, C., Kay, P., and O'Connor, C. Let Alone. 1983.Draft.

[Firbas 65] Firbas, Jan. A note on transition proper in functional sentence analysis. *Philologica Pragensia* 8:170-176, 1965.

[Fodor 81] Fodor, J. *Representations*. MIT Press, Cambridge MA, 1981.

[Fodor 83] Fodor, Jerry A. *The Modularity of Mind*. MIT Press, Cambridge MA, 1983.

[Fogelin 67] Fogelin, R. *Evidence and meaning*. Humanities, New York, 1967.

[Gazdar 79a] Gazdar, G. *Pragmatics: Implicature, Presupposition, and Logical Form*. Academic Press, New York, 1979.

[Gazdar 79b] Gazdar, G. A Solution to the Projection Problem. In Oh, C.-K. and Dinneen, D. (editors), *Syntax and Semantics*. Academic Press, New York, 1979.

[Gazdar 80] Gazdar, Gerald. Pragmatics and logical form. *Journal of Pragmatics* 4:1-13, 1980.

[Geis 71] Geis, M. and Zwicky, A. On invited inferences. *Linguistic Inquiry* 2(4):561-66, Fall, 1971.

[Gelman and Baillargeon 83]

Gelman, R., and Baillargeon, R. A review of some Piagetian concepts. In Flavell, J. H., and Markman, E. M. (editor), *Handbook of child psychology*, chapter 4th edition. Wiley, New York, 1983.

[Grice 65] Grice, H. P. The Causal theory of perception. In Shwartz, R. (editor), *Perceiving, Sensing, and Knowing*. Doubleday, New York, 1965. First published in 1961 in the *Proceedings of the Aristotelian Society*, Supplementary volume 35:121-152.

[Grice 68] Grice, H. Paul. Logic and conversation. Mimeo. 1968 Quoted in [Kempson 75].

- [Grice 69] Grice, H. P. Utterer's Meaning and intentions. *Philosophical Review* 68(2):147-177, 1969. Originally delivered at Oberlin College, April, 1968.
- [Grice 75] Grice, H. P. Logic and Conversation. In Cole, P. and Morgan, J.L. (editors), *Syntax and Semantics*. Academic Press, New York, 1975. From 1967 lectures.
- [Grice 78] Grice, H.P. Further notes on logic and conversation. In Cole, P. (editor), *Syntax and Semantics*, pages 113-128. Academic Press, New York, 1978.
- [Grosz 77] Grosz, B. *The Representation and Use of Focus in Dialogue Understanding*. Technical Report 151, SRI International, Menlo Park CA, 1977.
- [Grosz 83] Grosz, B., Joshi, A.K., and Weinstein, S. Providing a Unified Account of Definite Noun Phrases in Discourse. In *Proceedings of the 21st Annual Meeting*, pages 44-50. Association for Computational Linguistics, Cambridge MA, June, 1983.
- [Grosz and Sidner 85] Grosz, B. and Sidner, C. Discourse Structure and the Proper Treatment of Interruptions. In *Proceedings of the 1985 Joint Conference on Artificial Intelligence*. IJCAI85, Los Angeles, Ca., August, 1985.
- [Hamblin 71] Hamblin, C. L. Mathematical models of dialogue. *Theoria* 37(2):130-155, 1971.
- [Harnish 79] Harnish, Robert M. Logical form and implicature. In Bach, Kent, and Harnish, Robert M. (editors), *Linguistic Communication and Speech Acts*, pages 313-391. MIT, Cambridge MA, 1979.
- [Herskovits 82] Herskovits, Annette. *Space and the Prepositions in English*. PhD thesis, Stanford University, 1982.
- [Hintikka 62] Hintikka, J. *Knowledge and Belief*. Cornell University Press, Ithaca, 1962.
- [Hintikka 78] Hintikka, J. Answers to Questions. In Hiz, H. (editor), *Questions*, pages 279-300. Reidel, Dordrecht (Neth.), 1978.
- [Hirschberg 84a] Hirschberg, J. *Scalar Implicature: Generating Cooperative Responses to Yes-No Questions*. Technical Report MS-CIS-84-9, University of Pennsylvania, April, 1984.
- [Hirschberg 84b] Hirschberg, J. Toward a Redefinition of Yes/No Questions. In *Proceedings of COLING-84*. COLING, Stanford CA, July, 1984.
- [Hirschberg 84c] Hirschberg, J. Scalar Implicature and Indirect Responses to Yes/No Questions. In *Proceedings*. CSCSI-84, London, Ontario, May, 1984.
- [Hirschberg 84d] Hirschberg, J. Anticipating False Implicatures: Cooperative Responses in Question-Answer Systems. In L. Kerschberg (editor), *Proceedings of the First Int'l Workshop on Expert Database Systems*, pages 456-466. Kiawah Island SC, October, 1984.
- [Hobbs 79a] Hobbs, J. and Robinson, J. Why Ask? *Discourse Processes* 2, 1979.
- [Hobbs 79b] Hobbs, J. Coherence and coreference. *Cognitive Science* 3(1):67-90, 1979.
- [Hobbs ??] Hobbs, J. On the coherence and structure of discourse. SRI International.

[Hoepelman 83] Hoepelman, J. On questions. In Kiefer, F. (editor), *Questions and answers*, pages 191-227. D. Reidel, Dordrecht, 1983.

[Hoeppler 83] Hoeppler, W., et al. Beyond domain-independence. In *Proceedings of IJCAI-83*, pages 588-594. IJCAI, Karlsruhe, 1983.

[Hoeppler 84a] Hoeppler, W., Morik, K., and Marburger, H. *Talking it Over: The Natural Language Dialog System HAM-ANS*. Report ANS-26, University of Hamburg, May, 1984.

[Hoeppler 84b] Hoeppler, W. and Marburger, H. *Dialogue sequences with HAM-ANS*. Memo ANS-16, University of Hamburg, March, 1984. English version.

[Hoeppler 84c] Hoeppler, W., and Marburger, H. *Dialogue Sequences with HAM-ANS: Commented Examples of the System's Performance*. Memo ANS-16, University of Hamburg, March, 1984. English version.

[Horn 72] Horn, L. R. *On the Semantic Properties of Logical Operators in English*. PhD thesis, University of California at Los Angeles, 1972.

[Horn 73] Horn, L. Greek Grice: A brief Survey of proto-conversational rules in the history of logic. In *Papers of the Ninth Regional Meeting*. Chicago Linguistic Society, Chicago, 1973.

[Horn 84] Horn, L. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In Schiffrin, D. (editor), *Meaning, form, and use in context: linguistic applications*. Georgetown University Round Table 1984, Washington, D.C., 1984.

[Hungerland 60] Hungerland, I. Contextual Implication. *Inquiry* 3(4):211-258, 1960.

[Hutcheon 70] Hutcheon, E. G. An investigation into stimulus classification under varying instructions. Master's thesis, University of Dundee, 1970.

[Inhelder and Piaget 64] Inhelder, B., and Piaget, J. *The early growth of logic in the child*. W.W. Norton, New York, 1964.

[Intraub 81] Intraub, H. Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance* 7(3):604-610, 1981.

[Jackendoff 68] Jackendoff, R. S. Quantifiers in English. *Foundations of Language* 4:422-442, 1968.

[Jolicoeur 84] Jolicoeur, P., Gluck, M., and Kosslyn, S. M. Pictures and names: making the connection. *Cognitive Psychology* 16:243-275, 1984.

[Joshi 81] Joshi, A.K. and Weinstein, S. Control of Inference: Centering. In *Proceedings of the 1981 Meeting*, pages 385-387. International Joint Conf. on Artificial Intelligence, Vancouver, Canada, August, 1981.

[Joshi 82] Joshi, A.K. The Role of Mutual Beliefs in Question-Answer Systems. In Smith, N. (editor), *Mutual Belief*. Academic Press, New York, 1982.

[Joshi 84a] Joshi, A.K., Webber, B., and Weischedel, R. Preventing False Inferences. In *Proceedings of COLING-84*. COLING-84, Stanford CA, July, 1984.

- [Joshi 84b] Joshi, A., Webber, B., and Weischedel, R. Living Up To Expectations: Computing Expert Responses. In *Proceedings of AAAI-84*. AAAI-84, Austin, TX, August, 1984.
- [Kalita 84] Kalita, J. K., Colbourn, M. J., and McCalla, G. I. A response to the need for summary responses. In *Proceedings of COLING-84*, pages 432-436. COLING, Stanford, 1984.
- [Kaplan 79] Kaplan, S. Jerrold. *Cooperative Responses from a Portable Natural Language Data Base Query System*. PhD thesis, Univ. of Pennsylvania, 1979.
- [Karttunen 75] Karttunen, L. and S. Peters. Conventional implicature. In Cogen, C., et. al. (editor), *Proceedings of the First Annual Meeting*. Berkeley Linguistics Society, Berkeley CA, 1975.
- [Karttunen 79] Karttunen, L. and Peters, S. Conventional implicature. *Syntax and Semantics* 11:1-55, 1979.
- [Katz 64] Katz, J. J., and Postal, P. M. *Research Monographs*. Volume 26: *An Integrated theory of linguistic descriptions*. MIT Press, Cambridge MA, 1964.
- [Katz 76] Katz, J. J. and D. T. Langendoen. Pragmatics and presupposition. *Language* 52:1-17, 1976.
- [Kempson 75] Kempson, Ruth M. *Cambridge Studies in Linguistics*. Volume 15: *Presupposition and the delimitation of semantics*. Cambridge University Press, Cambridge, Eng., 1975.
- [Kiefer 80] Kiefer, F. Yes-No Questions as WH-Questions. In Searle, J., Kiefer, F., and Bierwisch, J. (editors), *Speech Act Theory and Pragmatics*, pages 97-119. Reidel, Dordrecht (Neth), 1980.
- [Knuth 73] Knuth, Donald E. *The Art of Computer Programming*. Addison-Wesley Publishing Company, Reading, MA, 1973.
- [Ladd 80] Ladd, D. R. *The Structure of Intonational Meaning*. Indiana University Press, Bloomington, Ind., 1980.
- [Ladusaw 80] Ladusaw, W. A. Affective Or, Factive Verbs, and Negative-Polarity Items. In *Papers of the Sixteenth Regional Meeting*. Chicago Linguistic Society, 1980.
- [Lakoff 71] Lakoff, R. Ifs, ands, and buts about conjunction. In Fillmore, C. and Langendoen, D. T. (editor), , pages 115-150. , 1971.
- [Lakoff, G 72] Lakoff, G. Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. In *Papers of the Eighth Regional Meeting*. Chicago Linguistic Society, 1972.
- [Lehnert 77] Lehnert, W. A conceptual theory of question answering. In *Proceedings of IJCAI-77*. IJCAI, Cambridge, MA, 1977.
- [Levesque 84] Levesque, H. A Functional Approach to Representation. *Artificial Intelligence* 23(2):155-212, July, 1984.
- [Levinson 83] Levinson, Stephen C. *Pragmatics*. Cambridge University Press, Cambridge, 1983.

- [Lieberman 77] Liberman, Mark, and Alan Prince. On stress and linguistic rhythm. *Linguistic Inquiry* 8(2):249-336, 1977.
- [Mann 84] Mann, W. C. Discourse Structures for Text Generation. In *Proceedings of COLING84*, pages 367-375. Association for Computational Linguistics, Stanford CA, July, 1984.
- [Mann and Thompson 83] Mann, W. C., and Thompson, S. A. *Relational Propositions in Discourse*. Technical Report ISI/RR-83-115, ISI/USC, November, 1983.
- [Markman 83] Markman, E.M., and Callanan, M.A. An analysis of hierarchical classification. In R. Sternberg (editor), *Advances in the Psychology of Human Intelligence*, 1983.
- [Mays 80a] Mays, Eric. Correcting Misconceptions about Data Base Structure. In *Proceedings of the Conference*. Canadian Society for Computational Studies of Intelligence, 1980.
- [Mays 80b] Mays, E. Failures in natural language systems: application to data base query systems. In *Proceedings of the First National Conference*. Conference on Artificial Intelligence (AAAI), Stanford CA, August, 1980.
- [McCawley 78] McCawley, James D. Conversational implicature and the lexicon. In Cole, P. (editor), *Syntax and Semantics*, pages 245-259. Academic Press, New York, 1978.
- [McCawley 81] McCawley, J. D. *Everything linguists have always wanted to know about logic but were ashamed to ask*. University of Chicago Press, Chicago, 1981.
- [McCoy 85] McCoy, K. *Correcting Object-Related Misconceptions*. PhD thesis, University of Pennsylvania, December, 1985.
- [McKeown 82] McKeown, K. . *Generating Natural Language Text in Response to Questions About Database Structure*. PhD thesis, University of Pennsylvania, May, 1982.
- [McKeown 85] McKeown, K., Wish, M. and Matthews, K. Tailoring Explanations for the User. In *Proceedings of the 1985 Conference*. International Joint Conference on Artificial Intelligence, Los Angeles CA, August, 1985.
- [Mercer 84] Mercer, R. and Rosenberg, R. Generating Corrective Answers by Computing Presuppositions of Answers, not of Questions. In *Proceedings of the 1984 Conference*, pages 16-19. Canadian Society for Computational Studies of Intelligence, University of Western Ontario, London, Ontario, May, 1984.
- [Morgan 78] Morgan, J. L. Toward a rational model of discourse comprehension. In Waltz, D. (editor), *Proceedings of TINLAP-2*. TINLAP, Urbana, Ill., 1978.
- [O'Hair 69] O'Hair, S. G. Implications and Meaning. *Theoria* 35:38-54, 1969.
- [Pereira 80] Pereira, F.C.N., and Warren, D.H.D. Definite clause grammars for language analysis. *Artificial Intelligence* 13:231-278, 1980.
- [Perrault 80] Perrault, C. Raymond and James F. Allen. A Plan-Based Analysis of Indirect Speech Acts. *American Journal of Computational Linguistics* 6:167-182, 1980.

- [Perrault 81] . It's for your own good: a note on inaccurate reference. In Joshi, A. K., Webber, B., and Sag, I. (editors), *Elements of discourse understanding*. Cambridge University Press, Cambridge, 1981.
- [Pierrehumbert 80] Pierrehumber, Janet B. *The Phonology and phonetics of English intonation*. PhD thesis, Massachusetts Institute of Technology, September, 1980.
- [Pinkal 83] Pinkal, Manfred. Questions of believing. In Kiefer, F. (editor), *Questions and Answers*, pages 241-156. D. Reidel, Dordrecht, 1983.
- [Pollack 82] Pollack, M. E., Hirschberg, J., and Webber, B. *User Participation in the Reasoning Processes of Expert Systems*. Technical Report MS-CIS-82-9, University of Pennsylvania, July, 1982. A shorter version appears in the AAAI Proceedings, 1982.
- [Pollack 86] Pollack, Martha E. *The Inference and Evaluation of Plans in Question-Answering*. PhD thesis, University of Pennsylvania, 1986. In preparation.
- [Prince 81a] Prince, E.F. Toward a taxonomy of given-new information. *Radical Pragmatics*. Academic Press, New York, 1981.
- [Prince 81b] Prince, E.F. Topicalization, Focus-Movement, and Yiddish-Movement: A Pragmatic Differentiation. In Alford, D., et al. (editor), *Proceedings of the Seventh Annual Meeting*, pages 249-264. Berkeley Linguistics Society, 1981.
- [Prince 82a] Prince, E.F. Grice and universality: a reappraisal. In *Proceedings of the Penn Linguistics Colloquim*. Penn Linguistics Colloquium, Philadelphia, PA, 1982.
- [Prince 82b] Prince, E.F., Frader, J. and Bosk, C. On Hedging in Physician-Physician Discourse. In De Pietro, R. (editor), *Linguistics and the Professions*, pages 83-97. Ablex, Norwood (N.J.), 1982.
- [Reiter 80] Reiter, R. A Logic for Default Reasoning. *Artificial Intelligence* 13:81-132, 1980.
- [Rosch 73] Rosch, E. Natural Categories. *Cognitive Psychology* 4:328-350, 1973.
- [Rosch 75] Rosch, E. and Mervis, C. B. Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology* 7:573-605, 1975.
- [Rosch 76] Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. and Boyes-Braem, P. Basic Objects in Natural Categories. *Cognitive Psychology* 8:382-439, 1976.
- [Russell 04] Russell, B. On Denoting. *Mind* 14:479-93, 1904.
- [Sadock 77] Sadock, J. M. Modus Brevis: The Truncated Argument. In *Papers from the 13th Regional Meeting*. Chicago Linguistic Society, 1977.
- [Sadock 78] Sadock, Jerrold M. On testing for conversational implicature. In Cole, P. (editor), *Syntax and Semantics*, pages 281-297. Academic Press, New York, 1978.
- [Schmerling 75] Schmerling, Susan F. Asymmetric conjunction and rules of conversation. In Cole, P. and Morgan, J. L. (editors), *Syntax and Semantics*, pages 211-231. Academic Press, New York, 1975.

- [Schuster 82] Schuster, E. *Explaining and Expounding*. Technical Report MS-CIS-82-49, University of Pennsylvania, 1982.
- [Searle 69] Searle, John R. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge UK, 1969.
- [Segui 68] Segui, J. and Fraisse, P. Le temps de reaction verbale. III reponses specifiques et reponses categorielles a des stimulus objects. *L'Annee psychologique* 68:69-82, 1968.
- [Sidner 79] Sidner, C. *Towards a computational theory of definite anaphora comprehension in English discourse*. PhD thesis, MIT, 1979.
- [Sidner 83] Sidner, C. L. Focusing in the Comprehension of Definite Anaphora. In M. Brady (editor), *Computational Models of Discourse*. MIT Press, Cambridge MA, 1983.
- [Smith 78] Smith, E. E., Balzano, G. J., and Walker, J. Nominal, perceptual, and semantic codes in picture categorization. In Cotton, J. W. and Klatzky, R. L. (editor), *Semantic factors in cognition*. Erlbaum, Hillsdale, NJ, 1978.
- [Soames 82] Soames, S. How presuppositions are inherited: a solution to the projection problem. *Linguistic Inquiry* 13(3):483-545, 1982.
- [Sperber 83] Sperber, D., and Wilson, D. Relevance. 1983. Manuscript.
- [Strawson 50] Strawson, P. F. On referring. *Mind* 59:320-344, 1950. New series.
- [Sugarman 83] Sugarman, S. *Children's early thought: Developments in classification*. Cambridge University Press, Cambridge MA, 1983.
- [Thomason 72] Thomason, R. H. A Semantic Theory of Sortal Correctness. *Journal of Philosophical Logic* 1:209-258, 1972.
- [Thomason 73] Thomason, R. H. Semantics, pragmatics, conversation and presupposition. 1973. Cited in Gazdar 79b.
- [Walker 75] Walker, Ralph C. S. Conversational implicatures. In Blackburn, S. (editor), *Meaning, reference and necessity: new studies in semantics*. Cambridge University Press, Cambridge, Eng., 1975.
- [Ward 83] Ward, G. A Pragmatic Analysis of Epitomization: Topicalization it's Not. *Papers in Linguistics* 17, 1983.
- [Ward 85a] Ward, Gregory, and Hirschberg, Julia. Implicating Uncertainty: The Pragmatics of Fall-Rise Intonation. *Language* 61(4), December, 1985. Forthcoming.
- [Ward 85b] Ward, Gregory. *The Semantics and Pragmatics of Preposing*. PhD thesis, University of Pennsylvania, August, 1985.
- [Warren 81] Warren, D.H.D., and Pereira, F.C.N. An efficient easily adaptable system for interpreting natural language queries. In *Proceedings. IJCAI-81, Vancouver, Can., 1981*.
- [Waxman 85] Waxman, Sandy. *Hierarchical relations in language and classification: evidence from preschool children*. PhD thesis, University of Pennsylvania, December, 1985.

[Wright 75] Wright, Richard A. Meaning_{gn} and conversational implicature. In Cole, P. and Morgan, J. L. (editors), *Syntax and Semantics*, pages 363-382. Academic Press, New York, 1975.

[Yu 79] Yu, Paul. On the Gricean program about meaning. *Linguistics and Philosophy* 3:273-88, 1979.

INDEX

- Affirming the consequent 88
 All-implication 90
 Alternates 122
 Answerhood 175
- BEL 10
 Belief 9
 Blocking 28, 60, 182, 183
 BMB 10
- Cancellation 18, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36,
 37, 38, 44, 54, 86, 92
 Cardinals 16, 44, 47, 48, 51, 92, 93, 94, 119, 183
 lexicalized 93
 Causatives 46
 CAUSE 10
 Centering 121, 142, 148, 150
 Classification 147, 156
 basic/entry level 157
 Communicative dynamism 152
 Conditional 88
 Conditionals 44
 Conjunction 7, 14, 25, 28, 41, 43, 70, 87, 88, 90, 91
 Connectives 47
 Consistency 25
 Context 9
 Context dependence 9, 16, 21, 29, 43, 140
 Context-dependence 7, 9
 Conventional force 8, 11, 24, 29, 30, 31, 34, 35, 36, 37, 92
 Conventions
 scalar implicature 65, 82, 127, 173
 COOP 178
 Cooperative Principle 11, 14, 15, 16
 Cycle 122
- Default reasoning 44
 Definites/indefinites 96
 Denying the antecedent 88
 Detachability/nondetachability 24, 30, 31, 33, 34, 35, 37, 38, 54
 Disjunction 87, 88, 89
 Distance from pole 47, 48, 75
 Domain of sortal applicability 69
- Entailment 1, 24, 25, 26, 28, 33, 36, 38, 47, 49, 50, 51, 52, 55, 56,
 57, 62, 69, 84, 115, 120
 Entailments 27
 Epistemic force 75
 Epitomization 121
 Expression 69, 71, 126, 141
 Expression alternate
 simple 150
 Expression alternative 72
 Expressions alternative
 simple 73
- FELICITOUS 32
 Focus 121, 142, 143, 146, 151, 153
 Functors 72, 73
- Given/new 147
 Goal inference 180
- HAM-ANS 178, 179
 Hedge 28
 Hedges 99
 approximators 100
 shields 100
- Illocutionary force 1
 Implicate 8
 Implicatum 8
 Implicature
 as disjunction 24, 79, 80, 81
 clausal quantity 44
 conventional 9, 24, 26, 27, 28, 31, 32, 33, 34, 36, 37, 38, 55
 disjunctive 188
 generalized conversational 41, 42, 43, 44, 45, 46, 52
 generalized quantity 41, 47
 multiple 17
 particularized conversational 42, 43, 46
 quantity-quality 47, 51
 scalar predication 47, 48, 53, 68
 scalar quantity 47, 68
 Implicature simpliciter 15
 INTEND 10
 Intention 9, 17, 18, 19, 27, 44
 Intonation 48, 58, 96
 contrastive stress 49, 121, 143, 145, 146, 152
 fall-rise 49, 121, 143
 It-clefts 142
- KNOW 10
- LICENSE 21
 Logical operators 47
- Maxims of cooperative conversation
 Manner 33, 70
 Quality 181
 Quantity 41, 89
 Quantity-Quality 86
 Maxims of cooperative conversation 11, 14, 63
 clash 15
 flouting 14
 Manner 11, 13, 14, 43
 opting out 15
 other 11
 Quality 11, 12, 15, 20
 Quality_{Joah} 12, 13
 Quantity 11, 13, 14

- Quantity-Quality 51
- Relation 11, 13, 15
- violating 15
- meaning_{nn}* 17, 18
- Modals 47, 71, 87, 115
 - deontic 87
 - epistemic 87
- Modus brevis 177
- Mutual belief 9, 11, 12, 21

- Negation 47, 72

- Ordering
 - available 139
 - compatible 138
- Orderings 122
 - duals 165
- Ordinals 47, 95

- Partially ordered sets 125, 137, 189
- Perlocutionary effect 1
- Polarity 121
- Presumption 179, 181
- Presupposition 1, 26, 27, 28, 29, 33, 36, 38, 69, 179, 181
- Presuppositions 32
- Principle of Charity 15
- Principle of extra effort 46
- Projection problem 68
- Pseudo-clefts 142

- QUALM 178
- Quantifier 85, 87
- Quantifiers 47, 84, 85, 86, 115

- REALIZE 10
- Redundance 36, 37, 38
- Reflexivity 122
- Reinforceability 36, 37, 38, 54
- Relevance 54, 65
- Rounding 93

- Saliency 65, 128, 133, 140, 148, 155
- SAY 10
- Scale 47, 50, 69
- Scripts 42, 166
- Semantic interpretation 70
- Semantic representation 70
- Semantics 6, 9, 27, 34
- Sentence 71, 74, 172
- Sortal correctness 120
- Specificity of assertion 91
- Strength of claim 13, 51, 52, 54
- Surface form 71, 89
- Surface order 27
- Suspension 28, 117
- Symmetry 122, 124

- There-insertion 142
- Topic 142, 143
- Topicalization 121, 142
- Transitivity 122, 124

- Utterance 10

- XCALIBRE 180
- Yes-no questions 60, 145, 153, 175, 176, 182