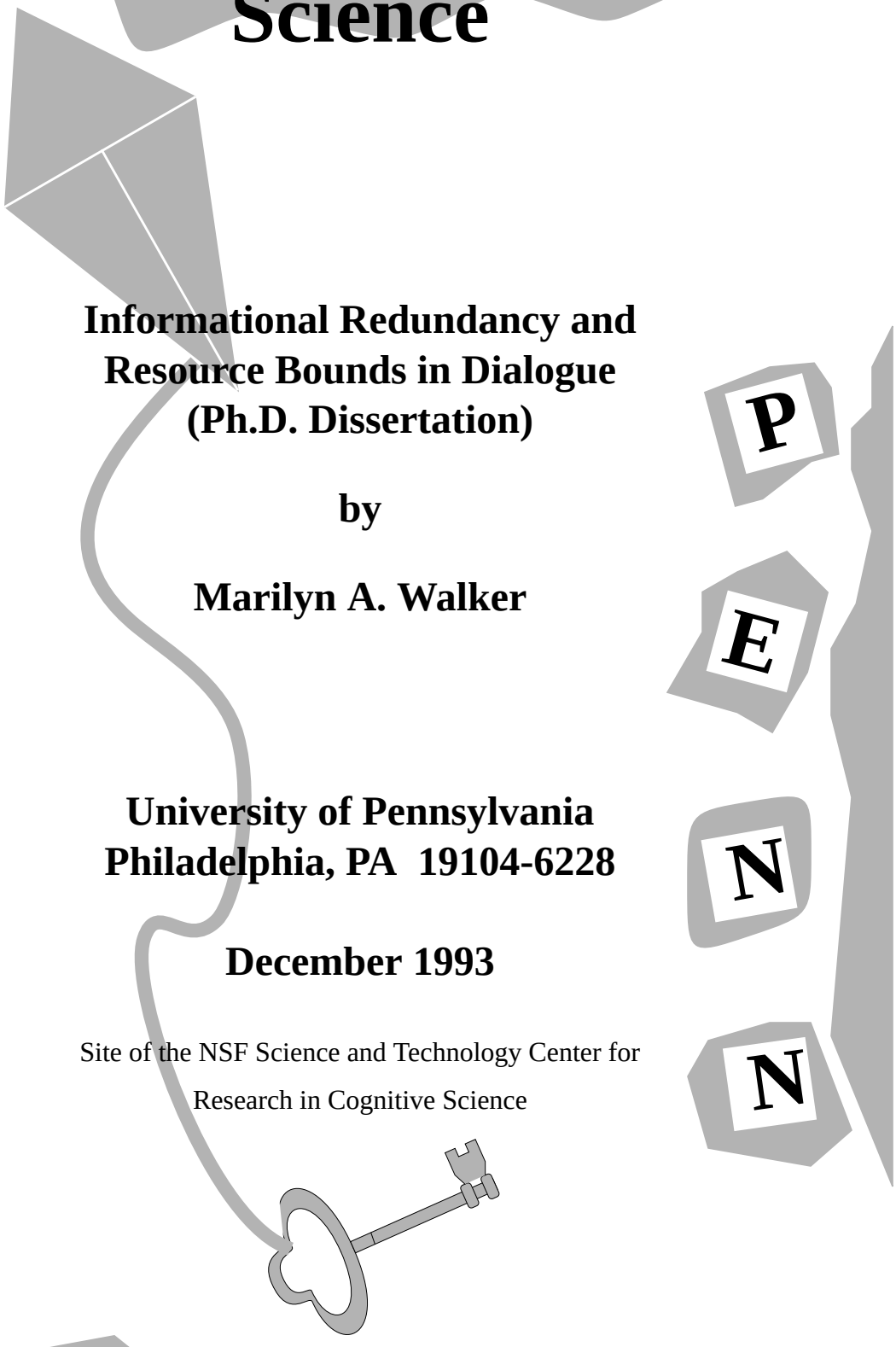
The Institute For Research In Cognitive Science



Informational Redundancy and Resource Bounds in Dialogue (Ph.D. Dissertation)

by

Marilyn A. Walker

University of Pennsylvania
Philadelphia, PA 19104-6228

December 1993

Site of the NSF Science and Technology Center for Research in Cognitive Science

University of Pennsylvania
Founded by Benjamin Franklin in 1740

IRCS Report 93-45

Informational Redundancy and Resource Bounds in Dialogue

Marilyn A. Walker

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

1993

Aravind K. Joshi
Co-Supervisor of Dissertation

Ellen F. Prince
Co-Supervisor of Dissertation

Mark Steedman
Graduate Group Chairperson

This research was partially funded by ARO grants DAAG29-84-K-0061 and DAAL03-89-C0031PRI, DARPA grants N00014-85-K0018 and N00014-90-J-1863, NSF grants MCS-82-19196 and IRI 90-16592 and Fellowship INT-9110856 for the 1991 Summer Science and Engineering Institute in Japan, and Ben Franklin Grant 91S.3078C-1 at the University of Pennsylvania, and by Hewlett-Packard Laboratories.

Acknowledgements

I'd like to thank my advisors, Aravind Joshi and Ellen Prince, for being interested in redundancy, for innumerable discussions over free lunches, for creating the Institute for Research in Cognitive Science, and for being a friend as well as an advisor.

Karen Sparck Jones was the best external possible. I am unendingly grateful to her for her unfailing dedication in reading and commenting on my proposal and for pushing me to find a way to computationally support my theory. I am also grateful to Mark Liberman for his skepticism about theories of discourse, for telling me about psychological theories of attention and memory, for making me think about learning and adaptability, and for teaching me about prosody. Thanks also to Max Mintz for teaching me about statistics and to Bonnie Webber and Scott Weinstein for being on my committee.

My colleagues have made writing this dissertation fun; thanks to the friends at HP and at Penn. There are no better colleagues and friends. Penni Sibun, Ellen Germain and Janet Cahn deserve special mention for reading and commenting on chapters of this dissertation.

Steve Whittaker has read and reread every chapter and discussed every idea. This dissertation would have been much the less without him.

Finally, I'd like to thank my parents, Marilyn and Phil Walker, for always believing in me and for insisting that one should have fun at work.

ABSTRACT

INFORMATIONAL REDUNDANCY AND RESOURCE BOUNDS IN DIALOGUE

Marilyn A. Walker

Aravind K. Joshi and Ellen F. Prince (cosupervisors)

This thesis investigates the relationship between language behavior and agents' resource bounds by examining the use of INFORMATIONALLY REDUNDANT UTTERANCES (IRUs) in problem-solving dialogues. The content of an IRU is a proposition that the conversants already know or could infer. Since communication is a subcase of action, the existence of IRUs is a paradox because IRUs appear to be actions whose effects have already been achieved. The explication of the paradox of IRUs has ramifications for models of action in general and of dialogue in particular.

I argue that IRUs can only be explained by a processing model of dialogue that reflects agents' autonomy and limited attentional and inferential capacity. The central thesis is that the communicative function of IRUs is related to the cognitive properties of resource-limited agents. In order to investigate this interaction the thesis relies on two empirical methods: (1) distributional analysis of IRUs in a large corpus of naturally occurring dialogues, and (2) dialogue simulations in the Design-World environment, which supports the parametrization of the dialogue situation, the task definition, and agents' cognitive properties.

The distributional analysis provides support for the claimed communicative functions by showing that IRUs demonstrate agents' autonomy and are used to support deliberation and inference. IRUs help autonomous agents coordinate on a collaborative task given their resource limits. The Design-World simulations show that discourse strategies that include IRUs are beneficial when agents are attention and inference limited or when the task is fault intolerant, inferentially complex, or requires a high degree of agreement. While some types of IRUs are beneficial simply as a rehearsal strategy, the general result is that IRUs are most beneficial when targeted at specific requirements of the communication situation.

Contents

List of Figures

Chapter 1

Informational Redundancy

1.1 Introduction

While the view of language as action is well-entrenched in theories of discourse, little attention has been paid to the fact that agents' resource-bounds must affect their language behavior just as it does every other type of action in which agents engage. Equally unrecognized is the fact that agents' dialogue behavior reflects their belief that their conversational partners, like themselves, are resource-limited. In addition, surprisingly little work has focused on how agents coordinate in dialogue given these limitations.

This thesis explores the interaction of resource-bounds and language behavior through an examination of the function of INFORMATIONALLY REDUNDANT UTTERANCES, utterances defined informally as those in which a conversant restates information that is already shared. INFORMATIONALLY REDUNDANT UTTERANCES, henceforth IRUs, will be defined more precisely in section 1.2. For now, consider the simple example of one type of IRU in 1 below:

(1) Frieda YOU'RE A PSYCHOLOGIST. What do you think about this case of a ten year old boy kidnapping a two year old? (fg 4/14/93)

In 1, both the speaker and the addressee already believed the proposition that *Frieda is a psychologist* and mutually believed that they did so. Therefore, it might seem that asserting this proposition is pointless.

However, one function of the IRU in 1 is to constrain the set of propositions relevant to answering the question to those related to the belief expressed by the IRU. IRUs also demonstrate that the interpretation of an utterance is highly **context dependent**. The same proposition realized in a

different context can **mean** different things, and, by becoming part of the context itself, can change the meaning of other propositions in that context.[1]

This is because the way that cognitive agents **process** language determines the meanings that an utterance can have. For example, consider the fact that a speaker cannot rationally intend a hearer to interpret an utterance with a meaning that is dependent on an inference if that inference would take the hearer too long to derive. Thus language, its use and structure, and the contribution of these to meaning, reflect processing constraints.

1.1.1 Resource Limits in Attention and Inference

This thesis explores the effects of two types of resource limitations on conversants' behavior in dialogue: limited attentional capacity and limited inferential capacity. It is widely known that human agents are limited in both their attentional capacity and their inferential capacity (Miller, 1956; Tversky and Kahneman, 1982; Sperber and Wilson, 1986; Johnson-Laird, 1991). Attentional capacity is defined by a limit on the number of propositions that can be held at one time in working memory. Propositions and other discourse entities currently held in working memory are SALIENT (Chafe, 1976; Prince, 1981b). While the exact limit on how many items can be salient at one time is not clear and depends on the nature of the material to be remembered (Baddeley, 1986), previous research suggests that the limit is somewhere around 7 primitive propositions, or 1 to 3 sentences (Miller, 1956; Kintsch, 1988). Chapter 3 introduces a model of working memory that will operationalize the concept of discourse salience.

Reasoning can be either time-limited or data-limited. A key assumption of the theory presented here is that the process of deliberating about beliefs, or deriving inferences from previously held beliefs, is constrained to operate on propositions that are currently discourse salient. This relation between discourse salience and inference is encapsulated in the DISCOURSE INFERENCE CONSTRAINT given below:

> DISCOURSE INFERENCE CONSTRAINT:
> Inferences in dialogue are derived from propositions that are currently discourse salient (in working memory).

Previous research suggests that inferences from salient propositions are drawn automatically, while inferences on propositions that must be retrieved from long term memory are the result of more

[1] It is well known that a sentence with indexicals such as *I am here* can mean different things when uttered in different situations. This is a stronger claim, namely that most aspects of interpretation are constrained by the context and that this is partially **because** agents are resource-limited.

strategic inference processes (McKoon and Ratcliff, 1992). Thus, if a belief is not currently salient is required for reasoning, a process must be invoked to retrieve it and thus make it salient. In contrast, inferences from propositions that are already salient are made without any special strategic inference process.

The DISCOURSE INFERENCE CONSTRAINT means that attentional capacity provides a major constraint on which facts are used in deliberation and in reasoning, in addition to the commonly assumed limits based on time or the number of derivation steps (Konolige, 1986). Thus one function of IRUs is to manipulate context by bringing propositions back into the listener's working memory, hence making them available for inferences.

1.1.2 The Effect of Resource Limits on Mutuality

An additional focus of this thesis is the effect of resource limits on the process of achieving the **mutuality** of certain beliefs and intentions in dialogue. Dialogue can be characterized as a coordination game between two resource-limited agents (Schelling, 1960; Lewis, 1969; Axelrod, 1984) who wish to achieve a level of mutuality that is dependent on their goals in the dialogue, but who must do so with a limited amount of resources. Agents do not assume that mutuality is automatically achieved. Because other agents are autonomous, an agent's assertions and proposals are not automatically accepted. This is reflected in the following ATTITUDE ASSUMPTION:

> ATTITUDE ASSUMPTION: Agents deliberate whether to ACCEPT or REJECT an assertion or proposal made by another agent in discourse.

That mutuality is not automatic means that agents have to work at achieving mutuality. This is represented by COORDINATION ASSUMPTION 1:

> COORDINATION ASSUMPTION 1:
> Achieving mutuality of beliefs and intentions is a coordination problem for conversants.

The required level of mutuality or coordination varies according to the type of dialogue and the individual purposes of the agents involved. Again, in contrast to agents' physical actions, which are observable, the mental objects of another agent's beliefs, preferences and intentions are not observable. Because of this, mutuality is achieved primarily when agents look for and give public evidence of the effects of utterance actions on their beliefs and intentions. Whether particular strategies are required to achieve this mutuality is dependent on the likelihood that agents can make **different** inferences and on the level of agreement required to achieve the purposes of the dialogue.

1.1.3 The Communicative Functions of IRUs

The motivation for IRUs is based on the two factors discussed above: (1) the role of the attentional and inferential limitations in deriving meaning, and (2) the need to achieve mutuality. These factors give rise to IRUs with three communicative functions:[2]

- **Communicative Functions of IRUs**:
  - Attitude: to provide evidence supporting beliefs about mutual understanding and acceptance
  - Attention: to manipulate the locus of attention of the discourse participants by making a proposition salient.
  - Consequence: to augment the evidence supporting beliefs that certain inferences are licensed

Example 1 was an Attention IRU. The IRU in 1 made the proposition salient that *Frieda is a psychologist* and making this proposition salient set the context for the interpretation of the question that followed. Below I will briefly give examples of some of the other types of IRUs; each communicative function given above is represented by a number of subtypes that will not be represented by these examples. The following chapters will discuss a range of examples of each type.

An Attitude IRU, said with a falling intonation typical of a declarative utterance, is given in 2-27, where M repeats what H has said in 2-26.[3]

(2) (24) H: That is correct. It could be moved around so that each of you have two thousand.
(25) M: I see.
(26) H: *Without penalty.*
(27) M: WITHOUT PENALTY.
(28) H: Right.
(29) M: And the fact that I have a, an account of my own from a couple of years ago, when I was working, doesn't affect this at all.

The IRU provides evidence supporting beliefs about mutual understanding and acceptance because M's repetition shows that she heard exactly what H said. In addition, according to the account

[2]There are also cases of REPAIR IRUs where it is clear that the speaker is reattempting the same action because some evidence has been provided by the addressee that the utterance has not been understood. While there are repairs IRUs in the corpus, their analysis seems fairly straightforward and can be handled directly by the account of context incrementation using defaults given in Chapter 6.

[3]This example is from a talk show for financial advice which was taped from a live radio broadcast and originally transcribed by Julia Hirschberg and Martha Pollack (Pollack, Hirschberg, and Webber, 1982). I am grateful to Julia Hirschberg for providing me with the tapes of the original broadcast.

that will be supported in Chapter 6, M's response supports the default inference that she accepts and therefore believes what H has asserted, because she provides no evidence to the contrary at this point in the dialogue. Attitude is mainly motivated by the need to achieve mutuality and is defined to be directed at demonstrating to other conversants what is assumed to be mutually understood or accepted.[4]

Consequence and Attention IRUs are motivated by both processing constraints and the need to achieve mutuality. This is because processing constraints introduce uncertainty about what will be retrieved from memory or retained in working memory (Attention) or what will be mutually inferred (Consequence). Consequence and Attention IRUs reduce this uncertainty by directly addressing these processing limitations. A Consequence IRU is given in 3, where 3-15 provides a biconditional inference rule, 3-16 instantiates one of the premises of this rule, and 3-17 realizes an inference that follows from 3-15 and 3-16, for the particular tax year of 1981, by the inference rule of MODUS TOLLENS:

(3) (15) H: Oh no. *I R A's were available as long as you are not a participant in an existing pension*
(16) j. Oh I see. Well I did work, *I do work for a company that has a pension*
(17) H: ahh. THEN YOU'RE NOT ELIGIBLE FOR EIGHTY ONE

The Consequence IRU in 3-17 augments the evidence supporting beliefs that certain inferences are licensed by making the inference explicit in the context.

Each communicative function proposed above is related to a noncontroversial cognitive property of human agents. The need to achieve goals related to Attitude follows from the fact that agents are autonomous with their own preferences, beliefs, and goals. If addressees always understood and accepted whatever a speaker asserted, then there would be no need for Attitude IRUs. However, addressees do not necessarily believe what they are told or adopt the goals that are suggested to them. Chapter 6 will discuss Attitude in more detail.

The need to achieve goals related to Consequence follows from the fact that agents are not logically omniscient. They might not have time to make all the relevant inferences and might not know all

[4]A distinction is made here between utterances who purpose is to **demonstrate** one's own mental state, which contributes to achieving mutual belief (Clark and Schaefer, 1989; Brennan, 1990), and utterances whose purpose is to affect another's mental state. Attitude IRUs are defined to be primarily about demonstrating one's own mental state, but cases of Attitude IRUs which, say, demonstrate that a certain inference was made, can simultaneously function to ensure that another agent made the same inference. At times the prosodic realization indicates that the speaker treats the information as part of the context, and so is merely demonstrating his/her understanding (Walker, 1993c). However, in many cases our understanding of the role of prosody is not far enough advanced to provide definitive arguments about utterance function (Pierrehumbert and Hirschberg, 1990; McLemore, 1991).

the relevant inference rules (Konolige, 1985). Time and processing/retrieval effort are key factors, given the heavy planning and processing requirements of producing and interpreting speech in real time (Clark and Brennan, 1990). Thus agents may make relevant inferences explicit, indicating to other agents that a certain inference was made. Agents also can provide support for their conversational partner's inference process, thus increasing the likelihood that another agent will make the desired inference. Chapter 8 will discuss Consequence in more detail.

Since human agents have limited attentional capacity, the need to manipulate Attention also follows from agents' resource-bounds. One aspect of agents' coordination in dialogue is tracking and manipulating the locus of attention of other agents. IRUs can be used to set the context for interpretation and reasoning and ensure that the agents involved are jointly attending to the same concepts. Chapter 7 will discuss Attention in more detail.

While there is a distinction between Consequence and Attention, the DISCOURSE INFERENCE CONSTRAINT means that making propositions salient that are premises for inferences also supports inferential processes. This illustrates a relationship between Consequence and Attention. One reason an agent can have for making a proposition salient is that it is a premise for an inference. If all the premises for an inference are discourse salient, then that inference is easily derived. Furthermore, if an inference is made explicit, then its content is salient, which means that other inferences can be derived from it. I will distinguish between Consequence and Attention by the use of the distributional parameters discussed in chapter 4 but the reader should keep in mind the fact that they are strongly related.

Another hypothesis about the function of IRUs is that they are a rehearsal mechanism, i.e. agents repeat propositions to help themselves and their conversational partners remember them. In addition, it is always possible that an IRU makes a proposition salient. Even if the proposition is already relatively salient, we might have many degrees of salience. The possibility that IRUs can function for rehearsal and salience is considered for each class of IRUs in the following chapters. The model of Attention/Working memory that will be presented in chapter 3 will ensure that the Design-World simulations will show rehearsal and salience benefits whenever they exist.

This demonstrates a general fact about the interpretation of IRUs and utterances in general: utterances can serve multiple functions simultaneously and any account of utterance function must allow for this possibility. The relationship between discourse function and IRUs is not claimed to be isomorphic: IRUs do not necessarily realize only one discourse function and the discourse functions discussed here are not realized solely by IRUs. This will become clearer in the following chapters, where each function of IRUs will be discussed. The relationship between Consequence and Attention will be discussed further in both Chapters 7 and 8.

1.1.4 Dialogue situations in which IRUs are Beneficial

While the functions of IRUs discussed above are very general, the occurrence of IRUs in naturally-occurring dialogues are highly **situation specific**. The communicative functions of IRUs proposed above predict which dialogue situations should lead to an increased frequency of IRUs due to a greater need to achieve these communicative functions. These are:

- tasks with high inferential complexity
- situations in which conversants have reduced or low attentional capacity
- situations in which there is uncertainty about whether assertions will be accepted by the other conversants
- situations in which there is uncertainty about whether understanding will be achieved

In the first two of these cases, IRUs are used to help other agents simplify processing, whereas, in the last two, IRUs are used as demonstrations of acceptance and understanding. One way of testing the theory of the function of IRUs is by manipulating the discourse situation in a computational simulation environment, and by demonstrating that IRUs are beneficial in these situations. The Design-World simulation environment is developed specifically for the purpose of providing support for the theory proposed here, and will be discussed in detail in chapter 5. Another way of testing the theory is to carefully examine the situations in which IRUs occur in the corpus, and see whether the distribution of IRUs supports arguments about their function. The distributional factors that are used to examine the distribution of IRUs in the financial advice corpus will also be discussed in chapter 4. The combination of these two methods will provide support for the claimed functions of IRUs elaborated in the remainder of the thesis.

This section has provided a sketch of the theory of the role of IRUs in discourse which will be elaborated and supported in the following chapters. First, however, Section 1.2 will define the concept of an INFORMATIONALLY REDUNDANT UTTERANCE more precisely. Chapter 2 will review previous analyses of IRUs in discourse. Finally, Section 1.3 then presents an overview of the remainder of the thesis.

1.2 Informationally Redundant Utterances

An IRU consists of information that is already shared between conversants. Information can be assumed to be shared because it was discussed in the current discourse, because it was discussed in

a previous discourse or because it is generally assumed to be commonly known. Prince calls these types of shared information HEARER OLD information. Thus IRUs consist of information that is HEARER OLD.

To explain how IRUs fit with the representation of HEARER OLD information in general, section 1.2.1 first discusses the representation of the discourse situation and how this determines the discourse model. Then in section 1.2.2 I discuss distinctions that have been made in the literature as to the different information statuses of entities in the discourse model. Finally, in section 1.2.3 I show that IRUs are are at one end of a continuum of HEARER OLD information that differs from other types in consisting of the assertion of complete propositions. I conjecture that if the functions of other classes of HEARER OLD propositions were studied, they might be similar to the functions of IRUs. Finally, in section 1.2.4 I present a brief taxonomy of types of IRUs that will be developed further in section 2.2.

1.2.1 The Discourse Model

One type of HEARER OLD information is what has been discussed in the current discourse. This information is usually stored in the DISCOURSE MODEL. The discourse model mediates between text and the real world. Information gets added to a discourse model through the occurrence of UTTERANCE EVENTS.[5] A discourse situation $\mathcal{S}$ is a sequence of utterance events, $\mathcal{S} = \mathrm{U}_1 \ldots \mathrm{U}_n$. An utterance event U is represented as a 4-tuple ($\mathrm{P} \times \mathrm{P} \times \mathrm{I} \times \Sigma$), where P is the population of conversants, I is the set of natural numbers representing the sequential locus of the utterance in a discourse, and Σ is the set of utterance strings. Other indices such as the time and place of the utterance will not be used here although they must be available for semantic interpretation. In addition, I am assuming dialogic discourses, but it would be simple to extend the treatment to audiences of more than one addressee. For convenience, selector functions are defined on U: Speaker(U) = U[1]; Addressee(U) = U[2]; U_n = U[3]; Σ(U) = U[4].

The situation $\mathcal{S}$ merely represents the sequence of events. What each conversant believes the discourse situation INDICATES maps $\mathcal{S}$ to each conversant's discourse model (Lewis, 1969). Conversants' models may vary; conversants' views of what is shared between the models depends on whether assertions are ACCEPTED or REJECTED and conversants' beliefs about dialogue conventions. A convenient fiction is that there is one discourse model; this is often referred to as **the** discourse model. For convenience, I will talk about a single discourse model in the remainder of

[5]The representation could be extended to include information from the physical environment, but that is beyond the scope of this work.

this section. In chapter 6, I will return to the discussion of potential differences in conversants' discourse models.

The discourse model minimally consists of the set of propositions, $\mathcal{P}$, that have been discussed and accepted (either explicitly or implicitly) and a set of discourse referents, $\mathcal{D}$ (Stalnaker, 1978; Karttunen, 1976; Webber, 1978; Heim, 1982). When an assertion is accepted, a predication about an entity in $\mathcal{D}$ is STORED in the model. Subsequent access to an entity or an accepted belief requires RETRIEVAL from the discourse model.

In addition, the discourse model includes (1) a function Struct-$\mathcal{D}$ on discourse referents defining an accessibility relation on discourse referents that is used for constraining potential antecedents for anaphors (See for example (Grosz, 1977; Sidner, 1979; Joshi and Weinstein, 1981; Walker, Iida, and Cote, 1993).); and (2) a function Struct-$\mathcal{P}$ defining a structure on the propositions in the discourse model. Struct-$\mathcal{P}$ restricts the domain of relevant propositions and constrains the availability of propositions used for inferences. Given the current discourse model, Struct-$\mathcal{P}$ returns a set of propositions that are available to constrain interpretation and inference. Thus the discourse model is a 4 tuple ($\mathcal{P}$, $\mathcal{D}$, Struct-$\mathcal{D}$, Struct-$\mathcal{P}$).

It is possible that Struct-$\mathcal{D}$ and Struct-$\mathcal{P}$ are determined by the same set of discourse structure constraints, but I leave this an open issue. Current dynamic semantics accounts assume that Struct-$\mathcal{P}$ simply corresponds to the sequence of utterances that realize $\mathcal{P}$ (Groenendijk and Stokhof, 1991). The function Struct-$\mathcal{P}$ is probably at least partially determined by the discourse structure, but the determination of discourse structure is an open problem (Hobbs, 1979; Polanyi and Scha, 1984; Reichman, 1985; Grosz and Sidner, 1986; Mann and Thompson, 1987; Roberts, 1993a). Struct-$\mathcal{P}$ will be defined here based on an operationalization of discourse salience and I will mainly be concerned with the relationship between Struct-$\mathcal{P}$ and what an utterance INDICATES (Lewis, 1969). This will be discussed in chapter 3.

In order to clarify the relationship of IRUs to other entities in the discourse model, the next section discusses the different information statuses of entities in the discourse model.

1.2.2 Information Status: Hearer Old and Salient

Entities in the discourse model can be classified according to two orthogonal distinctions: whether they are HEARER OLD and whether they are SALIENT (Prince, 1981b; Prince, 1992). HEARER OLD describes the **known** status of an entity: for referents this means whether or not a 'file card' for the entity has already been created in the discourse model. The SALIENCE status of an entity determines whether a previously existing file card is currently in working memory.

SALIENCE status for propositions is the same as for referents: whether or not the proposition is in current working memory. However, HEARER OLD and SALIENCE are dependent when applied to discourse referents, because it is difficult to imagine a referent being SALIENT unless it exists in the model. In contrast, as Horn (1986) noted, HEARER OLD status for a proposition reflects whether it is being treated by the participants as factive and for propositions this is completely independent of salience.[6]

Usually an utterance consists of some HEARER OLD information, along with some NEW information. For example, an utterance may be composed of an open proposition, the HEARER OLD information, and a discourse entity, the NEW information, that instantiates the variable in that open proposition (Prince, 1986). For example 5a introduces the open proposition given in 4:

(4) *λx.[go agt:Barbara loc:Grand Canyon with:x tns:past].*

(5) a. Who did Barbara go to the Grand Canyon with?

b. A friend from Slovenia.

b'. Jana.

Utterance 5b instantiates this open proposition with a new discourse entity *a friend from Slovenia.*

Because 5a presupposes that Barbara went to the Grand Canyon with someone and 5b doesn't take issue with this presupposition, the open proposition is treated as HEARER OLD information, in 5b. The answer given in 5b introduces a new discourse referent into the discourse model and provides the new information that *a friend from Slovenia* instantiates the variable in the open proposition. The answer given in 5b' presupposes that the addressee already has a discourse referent for *Jana* in the discourse model. In this case, the new information provided is only that *Jana* is the entity that instantiates the variable in the open proposition.

It is also possible for the hearer old information in an utterance to be a single discourse referent that is already in the discourse model such as *Barbara* in 6 below:

(6) Barbara went to the Grand Canyon.

Here the predication about the discourse entity *Barbara* is new, but the entity *Barbara* is already in the discourse model.

[6]It is possible to strongly believe a proposition, so that it is nondefeasible without it being currently salient. On the other hand, it is possible for a proposition, which is not strongly believed, to be under consideration and currently salient.

1.2.3 Defining Informationally Redundant Utterances

While the most common situation may be that the hearer old information in an utterance is either an open proposition or a discourse referent as in 5 and 6 above, a complete proposition can be hearer old information as well. Consider the proposition in 7a which could be realized as 7b:

(7) a. *[invite agt:phil theme:clare tns:past]*

b. *Phil invited Clare*

The examples in 8 show some of the many ways that this proposition can be realized in a discourse.

(8) a. *Phil invited Clare.*

b. That *Phil invited Clare* must have surprised you.

c. *Phil inviting Clare* balanced our table.

d. *Phil's invitation to Clare* means that we need an extra place setting.

e. If *Phil invited Clare* we need an extra place setting.

f. Sarah invited Kate because *Phil invited Clare.*

g. Sarah invited Kate and *Phil invited Clare.*

h. Sarah invited Kate but *Phil invited Clare.*

i. Grace didn't realize that *Phil invited Clare.*

Complete propositions can be **asserted** as in 8a, 8g and 8h. They may also be simply evoked, by nominalized finite and nonfinite clauses such as those in 8b and 8c, by nominalizations as in 8d, or in subordinate clauses as in 8e and 8f. A particularly interesting and complex class of cases is exemplified by 8i, in which a proposition is an argument of another predicate which expresses an attitude toward the proposition. Section 2.1 will discuss in more detail a subclass of these attitude expressing predicates which treat the propositional argument as presupposed (Kiparsky and Kiparsky, 1970; Karttunen, 1973; Gazdar, 1979).

This thesis examines only those cases where a proposition is both re-evoked and asserted as in 8a, 8g and 8h.[7] These propositions are HEARER OLD. A study of all the ways of re-evoking a

[7]The inclusion of conjoined propositions makes sense because the logical treatment of propositions realized by a sequence of utterances in discourse is as conjunction (Groenendijk and Stokhof, 1991), and because in natural speech very long sequences of utterances may be explicitly conjoined.

proposition as demonstrated in 8 is beyond the scope of this work. The definition of IRU is given in 9:[8]

(9) **Definition of Informational Redundancy**

An utterance u_i is INFORMATIONALLY REDUNDANT in a discourse situation $\mathcal{S}$

1. if u_i expresses a proposition p_i, and another utterance u_j that entails p_i has already been said in $\mathcal{S}$.
2. if u_i expresses a proposition p_i, and another utterance u_j that presupposes or implicates p_i has already been said in $\mathcal{S}$.

Presuppositions and implicatures are two types of non-logical inferences that may be communicated as part of the non-truth-conditional meaning of an utterance (Grice, 1967; Gazdar, 1979; Levinson, 1983). I will discuss the basis for these inferences in section 2.2.

Just as 8e, 8f, 8g and 8h explicitly relate the proposition expressed by *Phil invited Clare* to another proposition, and just as 8i is an argument of another predicate, IRUs may participate in relations of causality or explanation with other propositions, or can have an attitude of the speaker predicated of them via their intonational realization (See chapter 6). If a more exhaustive examination of the re-evocation of propositions in all the forms shown in 8 were carried out, it might well show that propositions that are re-evoked can have similar functions to those proposed for IRUs in section 1.1.

The definition of IRU given in 9 above depends on the logico-semantic account of the discourse model in which the function of utterances is to reduce the number of possible worlds consistent with what has been said. In this account, (1) an assertion reduces the context set by eliminating worlds which are not consistent with the newly added information; (2) propositions are added one at a time and a sequence of propositions is treated as the logical conjunction of individual propositions; and (3) all the inferences deriving from the most recently added proposition in combination with all the previously communicated propositions are automatically derived and added to the discourse model (Stalnaker, 1978; Gazdar, 1979; Barwise, 1988b; Groenendijk and Stokhof, 1991).[9] While this model is used to define utterances as informationally redundant, the remainder of the thesis is devoted to exploring a richer model of the context for a dialogue in which IRUs are not **communicatively redundant**. The claim here is that the discourse representation must take into account conversants' resource limitations.

[8] The first part of the definition is a variation on the definition of redundant given by Hirschberg (1985) and used in her theory of scalar implicature.

[9] The automatic derivation of all inferences is an unavoidable consequence of the possible worlds model (Partee, 1982; Levesque, 1984; Konolige, 1985).

The definition of an IRU is based solely on the informational component of an utterance, using a possibly narrow definition of information. Utterances consist of a string that is said, the proposition that is realized by that string in a particular context, and an associated utterance intention, which consists of the role of that utterance in the overall structure of the discourse and the conversants' intentions. Informational redundancy only refers to the propositional content of an utterance. An IRU does not necessarily realize the same utterance intention as the utterance that originally added the propositional content of the IRU to the discourse model.

1.2.4 Types of Informationally Redundant Utterances

In what follows, it will be useful to have a term to refer to the utterance(s) that originally added the propositional content of the IRU to the discourse situation. Following work on referential discourse entities, I will call this the IRU's ANTECEDENT. The term antecedent will be used to refer to both the prior utterance and the proposition realized by that prior utterance, but this should not cause any confusion. In the dialogue excerpts given here, IRUs will be marked with CAPS whereas their antecedents will be given in *italics*. An IRU may be explicitly related to its antecedent, e.g. a repetition which replicates the surface structure of the antecedent, or implicitly related to its antecedent, e.g. inferrable from its antecedents by modus ponens or other deductive inference rules. The types of IRUs examined here are shown in figure 1.1.

TYPE of IRU	EXPLICIT Relation to Antecedent(s)	IMPLICIT Relation to Antecedent(s)
Entailment	Repetitions Paraphrases	Logical Inferences Mathematical calculations
Presupposition		Existential Presuppositions Factive Presuppositions
Conversational Implicature		Scalar Implicatures

Figure 1.1: Types of Informationally Redundant Utterances

Thus there are three basic types of IRUs: entailments, presuppositions and implicatures. From a logical perspective there is no substantive difference among the different types of entailments, but from a cognitive and communicative perspective they may be more or less 'available'.[10]

[10]One motivation for distinguishing repetition from other entailments is that lexical repetition has often been taken as a key parameter for predicting prosodic deaccenting of given information (Cruttenden, 1986; Walker, 1993c).

One remaining issue with the notion of entailment, and specifically related to the concept of paraphrase, is how propositions are represented. For example if one speaker says 10a and the other says 10b, it seems that the propositional representation should allow us to say that 10b is an IRU because *truly* and *wonderfully* are 'close enough' in meaning. Thus the notion of entailment required here may be closer to what Resnik has called PLAUSIBLE ENTAILMENT based on measures derived from lexical collocation indices (Resnik, 1993; Sparck-Jones, 1964).

(10) a. Juliet is a truly beautiful girl.

b. Wonderfully beautiful girl is Juliet.

For all types of entailment, a diagnostic of whether the propositional content of an IRU can plausibly be denied can be used to test whether the information is already available in the discourse situation (Stalnaker, 1978). This diagnostic cannot be used for implicatures since these inferences are in fact defeasible.[11]

In an examination of redundancy via corpus analysis, what counts as an entailment or as inferrable in general must be strictly defined for operational reasons. This is so that the selection criteria are replicable. In analyzing the radio talk show corpus, I assume that lexical knowledge is shared if it is not domain-specific. Otherwise, all the information that an entailment depends on must be made public in the dialogue, although I will assume knowledge of standard deduction schemas such as modus ponens, modus tollens, proof by elimination, conditionalization, etc.

On the other hand, when the use of IRUs is examined in empirical situations such as task-oriented dialogues collected experimentally or through simulations, it is possible to state in principle what is already known and what is inferrable. This leads to the possibility of a broader characterization of IRUs in these circumstances, which will be discussed in more detail when I discuss these experimental environments in chapter 5.

1.3 Overview of the Thesis

Section 1.1 sketched the account of INFORMATIONALLY REDUNDANT UTTERANCES (IRUs) that will be developed in the remaining chapters. I argue that there are three main classes of IRUs: Attitude, Attention and Consequence. These functions of these classes of IRUs can be explained by a processing model of dialogue that reflects the fact that agents are autonomous, and have limited attentional and limited inferential capacity. Section 1.2 defines the set of IRUs analyzed here. I

[11] Reasons for examining these defeasible inferences are discussed in section 2.2.

haven't said how the theory of IRUs will be supported and tested: this is explained in chapters 3, 4 and 5.

Chapter 2 will review relevant previous work on redundancy in discourse.

Chapter 3 presents the theory of agents attentional and inferential systems that supports the analysis of the function of IRUs given here. This account builds on the work on resource-bounded agents of Bratman, Israel and Pollack, on Gallier's theory of belief revision, and on Lewis's shared environment model of mutual belief (Bratman, Israel, and Pollack, 1988; Pollack and Ringuette, 1990; Galliers, 1991b; Galliers, 1991a; Lewis, 1969; Clark and Marshall, 1981; Barwise, 1988a).

Chapters 4 and 5 describes the two empirical methods that will be used to support the theory proposed here: corpus based analysis and computational modeling. Corpus based distributional analysis provides support for claims about the relationship of context to function, and computational modeling supports claims about processing. Chapter 4 discusses the factors used in the corpus analysis of IRUs and chapter 5 presents the Design-World simulation environment used to test the processing claims of the theory. Then each of the following chapters will present evidence for the theory from these empirical bases.

Chapter 6 elaborates on the function of Attitude IRUs. This chapter proposes an inferential account of how understanding and acceptance are achieved; this account is based on the model of mutual beliefs presented in chapter 3. The treatment of Attitude distinguishes between understanding, acceptance and rejection of the previous utterance, and discusses cases of IRUs that support the inference of mutual beliefs about understanding and acceptance. A model of this inference process is proposed and the class of examples explained by this model is demonstrated. This model is then used to explain how agents make plans together. Design-World simulations of some types of Attitude IRUs are presented in section 6.6. The main result of these experiments is that Attitude IRUs can help agents avoid making mistakes such as putting invalid steps in their plans.

Chapter 7 on Attention IRUs discusses three distributional classes of of Attention IRUs: Open Segment, Close Segment and Deliberation IRUs. The simulation shows that when both agents know exactly what the structure of the task is, Open Segment and Close Segment statements alone have little benefit. It is only when these Open Segment and Close Segment statements include other IRUs that a benefit for these can be demonstrated. In these cases the function of Attention IRUs overlaps that of Consequence IRUs. Deliberation IRUs are particularly beneficial in reducing the retrieval of propositions from memory, even when agents are logically omniscient, because they obviate the need for search.

Chapter 8 on Consequence IRUs discusses two types of Consequence IRUs: Inference Explicit and

and Affirmation. The distributional analysis supports the claim that the premises for inferences are restricted to those that are currently salient. Design-World simulations for Consequence IRUs are carried out in two situations: (1) where agents are logically omniscient and (2) where agents are inference limited. Consequence IRUs are of some benefit even when agents are logically omniscient, in cases of tasks with high inferential complexity. In situations where agents are inference-limited, Consequence IRUs can provide major benefits as would be expected.

Finally Chapter 9 discusses the ramifications of the analysis and proposes future work.

Chapter 2

Previous Research

This chapter reviews previous research related to the treatment of IRUs or work that has specifically addressed the discourse functions of IRUs. Section 2.1 discusses assumptions of previous work as to the functions of IRUs and some issues that these assumptions raise. Section 2.2 discusses why IRUs present a problem for the Gricean distinction between entailment and implicature. Section 2.3 discusses proposed analyses of subclasses of IRUs in previous work.

2.1 Introduction

There has been little analysis of IRUs in formal theories of dialogue in linguistics, philosophy, and computational linguistics, and the analyses proposed have been rather uneven. Indeed, most logic-based theories of dialogue have rules that forbid the use of IRUs. For example, Hamblin's formal system for dialogue includes a rule that forbids the speaker to say anything that is already part of the common ground (Hamblin, 1971). Stalnaker states that *to assert something that is already presupposed is to attempt to do something that is already done* (Stalnaker, 1978). And Grice's QUANTITY maxim: *Do not make your contribution more informative than is required* has often been interpreted to mean that asserting the same proposition twice is infelicitous (Grice, 1967; Horn, 1991).

The view of language as an action on another's mental state has meant that theories of rational behavior can be applied to language (Austin, 1965; Grice, 1967; Searle, 1975). This program is best exemplified by the application of planning paradigms to language and has produced many useful insights (Bruce, 1975; Cohen, 1978; Allen, 1979; Sidner and Israel, 1981; Allen and Perrault, 1980; Sidner, 1985; Cohen and Levesque, 1985; Litman, 1985; Litman and Allen, 1990). However, these treatments of language haven't considered cases where there might be a reason to communicate

a proposition that is already believed by the hearer. Cohen's speech act generation axioms rule out the generation of INFORM utterances whose propositional content has already been conveyed because a precondition for an INFORM is that the hearer doesn't already believe the content of the utterance (Cohen, 1978). Similarly, speech-act plan-inference heuristics disprefer the recognition of a plan whose effect has already been achieved and thus disprefer recognizing an IRU as an INFORM (Allen, 1983; Litman and Allen, 1990).

It seems that it is not uniformly recognized that IRUs occur, except for certain special types of IRUs such as tautologies, whose treatment will be discussed below. Rules that forbid IRUs stem from both types of formal analyses of language: (1) the logico-semantic account in which the purpose of utterances is to describe the world and an utterance is informative only if it reduces the number of possible worlds consistent with what has been said; and (2) the plan-based view, in which utterances are actions on the addressee's mental state, but in which IRUs are actions whose purpose has already been achieved because the hearer already believes the proposition that is realized by the utterance.

Both of these analyses can treat IRUs as felicitous if they may be interpreted nonliterally. In the logico-semantic analysis this means that the meaning of the utterance bears some indirect relation to its content, which must be derived by access to other information, but the details of this derivation have yet to be specified. One account in the plan-based tradition is that the nonliteral interpretation comes about from the recognition of the utterance as an 'indirect speech act' (Perrault, 1990). On this account, the recognition that the utterance is redundant means that it violates the felicity conditions for an INFORM: *the hearer doesn't already believe the content of the utterance*(Cohen, 1978). Based on the assumption of the speaker's cooperativity, this recognition then triggers a process which attempts to infer a different speaker intention such as a REQUEST. If these nonliteral interpretations require that the alternate content or speech act is a complex function of the content of the IRU and the context of utterance, then these other accounts could end up being quite similar to the one given here. Both of these ways of determining an alternate interpretation will be discussed further below.

Similar, but not identical, to the nonliteral meaning analysis, is one in which IRUs are felicitous as long as they add implicatures to the common ground (Gazdar, 1979). This is the basis of Gazdar's treatment of tautological utterances such as 11a, which adds the two implicatures in 11b and 11c:

(11) a. *She will either come or she won't.*

b. *It is possible that she will come.*

c. *It is possible that she won't come.*

On this view, depending on what can count as an implicature, IRUs can always add implicatures; they trivially add 'relevance' implicatures based on Grice's Relevance Maxim:

> *Make your contributions relevant*

The Relevance Maxim is based on the assumption that the speaker must have had a reason for saying the utterance. If these reasons include metalinguistic propositions such as 'the speaker really believes P' or 'the speaker has reasons to want to make sure that the hearer doesn't forget that P', then what I would characterize as functions of IRUs with a cognitive basis might possibly be treated as types of relevance implicatures. However, this seems contrary to the spirit of Grice's account, which was primarily concerned with content-based inferences.

Note also that by any account where IRUs mean something other than their literal meaning, IRUs are informative via some inferential process. Yet this does not distinguish IRUs as a special case since most utterances communicate more than their literal content (Karttunen and Peters, 1979; Gazdar, 1979; Grosz and Sidner, 1986; Hobbs, 1979). How can we distinguish the standard case where the realization of any proposition adds inferences from its content to the context and the case where IRUs are only felicitous when they do so?

One way to approach this is by determining whether or not the interpretation and inferred function of the IRU depends on **recognizing** the IRU as redundant, and then a special inference process is triggered because the utterance is redundant. One problem with this approach is to say whether or not this recognition needs to be **conscious**, i.e. whether the addressee can say, upon introspection, that the reason the utterance functioned as it did was because it was recognized as already believed.

This is still a problem however because it is difficult to distinguish cases of conscious recognition from differential processing due to redundancy. For example, hearers react more quickly to facts that they already know, rehearsing a fact increases its accessibility, and facts that are already believed may not need to be verified through additional processing (Baddeley, 1986; Landauer, 1975; Hintzmann and Block, 1971; Shepard, 1967; Tulving, 1967; Ratcliff and McKoon, 1988). If propositions realized as IRUs are more quickly accessed and processed, the hearer may perceive that arguments that include IRUs are more coherent, without necessarily consciously recognizing their redundancy. The question of whether recognition of redundancy is an important part of the function of IRUs will be discussed for each class of IRUs in the following chapters. The following section will discuss the relationship of IRUs to the Gricean program, which in the main has assumed that IRUs don't exist. Section 2.3 will then discuss previous research on the function of IRUs in discourse.

2.2 A Classification of IRUs according to Grice

2.2.1 The Reinforceability and Defeasibility Diagnostics

The Gricean program greatly simplified the job of providing a semantics for natural language by distinguishing those aspects of meaning that are truth-conditional (entailments) from those that arise only in particular conversational contexts (conversational implicatures). According to the classical Gricean view, there are two logical properties that distinguish presuppositions and implicatures from entailments (Levinson, 1983; Sadock, 1978):

- Reinforceability: whether the inference can be made explicit without redundancy.
- Defeasibility: whether the inference can be defeated by additional information or depending on the discourse situation when the utterance is made.

The way in which the types of IRUs given in Figure 1.1 are classified by these properties is shown in Figure 2.1 (Gazdar, 1979; Bridge, 1991).

	Reinforceability	Defeasibility
Entailment	no	no
Presupposition	order dependent	by context
Conversational Implicature	yes	yes

Figure 2.1: Properties of different types of Information Antecedents

It is clear that IRUs are a problem for the Gricean program, since felicitous reinforceability is the main diagnostic for distinguishing implicatures from entailments. In the remainder of this section, I will present examples that exemplify the claims of the classical Gricean view and contrast them with examples from my corpus that are counterexamples to this view, but which however seem perfectly felicitous.

2.2.2 IRUs include Entailments

As an example of the infelicity of attempting to reinforce or defeat entailments, consider 12a and 12b:

(12) a. # My sister is older than I am and I am younger than my sister.

b. # My sister is older than I am but I'm not younger than my sister.

Example 12a conjoins two clauses where the second is entailed by the first; 12b conjoins two clauses where the second attempts to defeat the first. Because entailments are neither reinforceable nor defeasible, 12a and 12b are anomalous. The anomaly occurs independent of the order in which these clauses are stated. However, now consider 13-9, extracted from a longer naturally-occurring dialogue.

(13) (8) H: you can stop right there: *take your money.*
(9) J: TAKE THE MONEY.
(10)H: absolutely.....

In 13, 9 is said with falling intonation, the H*+LLL% declarative pattern (Pierrehumbert, 1980). The contour indicates that J is **not** questioning (h)'s assertion due to beliefs to the contrary. Indeed, just prior to 13-8 J had asked H **whether** she should take the money from her pension plan (as a lump sum payment) or take an annuity. This question, under standard circumstances, should convey that *J doesn't know whether p.* Thus neither J's intonation nor the prior context gives any reason to believe that J is expressing disbelief or surprise at the advice.

Entailments also include inferences made via classic inference rules, such as modus ponens and modus tollens. Condition (1) of the definition of IRU in 9 means that if 14a and 14b are in the common ground, then an utterance that realizes the proposition in 14c will be an IRU since it follows from 14a and 14b via modus tollens:

(14) a. You can buy an I R A[1] if and only if you do NOT have an existing pension plan.

b. You have an existing pension plan.

c. You cannot buy an I R A.

This structure is illustrated by 15, where 15-15 realizes the proposition in 14a, 15-16 realizes the proposition in 14b, and 15-17 makes the inference explicit that is given in 14c for the particular tax year of 1981.

(15) (15) H: Oh no. *I R A's were available as long as you are not a participant in an existing pension.*
(16) J: Oh I see. Well I did work, *I do work for a company that has a pension.*
(17) H: ahh. THEN YOU'RE NOT ELIGIBLE FOR 81.

[1] An Individual Retirement Account, which was a way of deferring tax on income until retirement.

Both 13 and 15 seem completely felicitous. The contrast between these examples and the infelicitous reinforcement given in 12a is explained by the theory proposed in the following chapters.

2.2.3 IRUs include Presuppositions

Like entailments, presuppositions are supposed to be not reinforceable (Karttunen, 1973; Gazdar, 1979). Presuppositions are similar to implicatures in that they are not part of the truth-conditional meaning of an utterance. There are two types of presuppositions: existential and factive. These will be discussed in turn below.

Existential presuppositions are often carried by the use of definite expressions. For example, an utterance in which I speak of *my sister* will carry an existential presupposition that *I have a sister*.

(16) a. I have a sister and my sister is older than I am.

b. # My sister is older than I am and I have a sister.

c. I don't have a sister so my sister isn't older than I am.

In 16a, the presupposition is in the clause before the one that presupposes it. Since presuppositions can be reinforced as long as the reinforcement comes before the clause that introduces the presupposition, 16a is not anomalous. However 16b puts the presupposition after the clause that presupposes it and is anomalous; 16c shows that presuppositions can be defeated by entailments that are already part of the context (Gazdar, 1979).

Unlike entailments and implicatures, existential presuppositions survive negation. Note that the existential presupposition given in 17c, is presupposed by both 17a and 17b.

(17) a. The charges will be excessive.

b. The charges won't be excessive.

c. There are charges.

In contrast with the infelicity of 16b, consider the affirmation of *there are charges* in 18:

(18) (22) B: Are there ah .. I don't think *the ah brokerage charge* will be ah that excessive
(23) H: No *they're* not excessive but THERE ARE CHARGES

In 18, B starts out with a question as to whether there are brokerage charges or not, but modifies his utterance to presuppose that there are brokerage charges by using the definite referring expression *the brokerage charge*. In 18-23, H further presupposes the existence of said brokerage charges, by referring to them with a personal pronoun and then predicating of them that they are in fact not excessive, and then, in the second conjunct, he affirms their existence.

Factive presuppositions are introduced into the context as the argument of a factive predicate, e.g. *unfortunate* in 19:

(19) It's unfortunate that you failed.

The emotive factive predicate *unfortunate* in 19 presupposes the proposition *you failed*. This proposition should not be reinforceable according to the classic account. However, the affirmation in 20 is perfectly felicitous (Ward, 1985; Horn, 1991; Ward, 1990):[2]

(20) It's unfortunate that *you failed*, but YOU DID.

Factive predicates, first noted in (Kiparsky and Kiparsky, 1970), include *odd, strange, surprising, realize, regret, manage*.

2.2.4 IRUs include Conversational Implicatures

Generalized conversational implicatures, such as the inference from *some* to *not all*, are both reinforceable without anomalous redundancy, as in 21a, and defeasible, as in 21b:

(21) a. I ate some of the cookies but I didn't eat all of them.

b. I ate some of the cookies and in fact I ate all of them.

Since such conversational implicatures are reinforceable, a possible account of IRUs that have a conversational implicature as an antecedent is that the IRU is just a reinforcement of the implicature. However, there are two reasons to examine implicatures. First, such reinforcements appear to occur in contexts similar to IRUs whose antecedents are entailments. Second, all the work on the properties of implicatures, such as reinforceability and defeasibility, has been based on **single utterances** such as those in 21, where the defeating or reinforcing proposition is immediately adjacent to the implicata and said by the same speaker. Reinforcements that are not adjacent to their

[2]Ward's and Horn's analyses of IRUs of this type are based on the concept of RHETORICAL CONTRAST, which will be discussed further in sections 2.3 and 8.2.

antecedents may primarily function as Attention IRUs rather than at the logico-semantic level of reinforcement.

An IRU with an implicature as antecedent is shown in 22. In 22-9, D uses the term *girlfriend*, which implicates, by the Quantity Maxim (Grice, 1967; Horn, 1972; Hirschberg, 1985), that the woman in question is not his wife.[3]

(22) (9) D: Uh *my girlfriend* and I bought a house in December. And uh we're both on the mortgage. And we're not sure as to how to handle the mortgage, or excuse me, how to handle the deductions.
Uh every example I see in tax books usually cites a married couple though and WE'RE NOT MARRIED.
And we uh, I'm afraid that this year she will not qualify, for uh, to itemize. I will – I'll be well over it.
(10) H: Who made the payments? ....

Note that as D continues the description of his problem, he makes this implicature explicit, by actually stating that they are not married.[4]

In sum, while the Gricean program depends on the non-reinforceability of entailments, examples from the corpus show that all types of entailments are reinforceable. The infelicity of some reinforcements is explained by what follows. Section 2.3 reviews previous analyses of some subclasses of IRUs which will be extended in the following chapters.

2.3 IRUs in Previous Work

In section 2.1 I discussed the assumption of 'no redundancy' incorporated in formal theories of dialogue, then in section 2.2, discussed Gricean distinctions between different types of propositions that are based on the 'no redundancy' assumption. This review gives a somewhat negative characterization of redundancy. This section reviews proposals for the function of various types of IRUs (Schiffer, 1972; Levinson, 1979; Heritage and Watson, 1979; Sperber and Wilson, 1981; Schiffrin, 1982; Schegloff, 1982; Levinson, 1983; Ward, 1985; Finin, Joshi, and Webber, 1986; Schiffrin, 1987; Cohen, 1987; Whittaker and Stenton, 1988; Clark and Schaefer, 1989; Tannen, 1989; Ward, 1990; Walker and Whittaker, 1990; Horn, 1991; Ward and Hirschberg, 1991).

[3] I find this inference to be stronger than an implicature, since I find it infelicitous to say *My girlfriend, and in fact we're married,...*

[4] The use of sentential conjunction by D demonstrates the difficulties in determining when a proposition linked to others by conjunction could count as a separate utterance.

2.3.1 Schiffer's Reminders

Schiffer's definition of speaker meaning anticipated the class of Attention IRUs by defining mutual beliefs to be those that are mutually activated (Schiffer, 1972):

> S meant that p by (or in) uttering x if S uttered x intending thereby to realize a certain state of affairs $\mathcal{E}$ which is (intended by S to be) such that the obtainment of $\mathcal{E}$ is sufficient for S and a certain audience A **mutually knowing* (or believing*)** that $\mathcal{E}$ obtains and that $\mathcal{E}$ is conclusive (very good or good) evidence that S uttered x with the primary intention
>
> 1. that there be some ρ such that S's utterance of x causes in A the activated belief that $p/\rho(t)$[5] ;
> 2. satisfaction of (1) to be achieved at least in part by virtues of A's belief that x is related in a certain way R to the belief that p;
> 3. to realize $\mathcal{E}$.

By Schiffer's definition, utterances realize states of affairs that provide evidence for the active inferences of an audience. This evidence has a conventional basis in 'common knowledge'. Schiffer includes the notion of an activated belief, rather than just any belief in order to account for cases of reminding, pointing out, etc, in which a speaker presumably tells the hearer something that they already know. Schiffer had in mind utterances such as 23b:

(23) A: Now what was that girl's name?

B: Rose.

Thus Schiffer distinguishes between known facts and their accessibility or salience in terms of activation. There are no examples like this in the corpus and the types of Attention IRUs are much more varied than Schiffer's characterization would suggest. However it seems that Schiffer is making the distinction between HEARER OLD information and that which is both HEARER OLD and SALIENT, at issue in the analysis of Attention IRUs,

Another point of interest is that Schiffer claimed the actual saying of 23B is the only thing required for the satisfaction of the speaker's intention. Presumably, this follows from the fact that saying an utterance makes its content salient, and since the proposition realized by the utterance is already believed, this is all that is required. No active processing is needed because, as long as the literal

[5] p is the response with reason(s) ρ, and that these are truth supporting reasons is denoted by $\rho(t)$

content of the utterance is understood, it achieves the intention of making the proposition it realizes salient.

2.3.2 IRUs are used in Arguments

Levinson foreshadows the analysis of Consequence and Attention IRUs in noting that IRUs can be important in argument structure. In discussing various types of uninformative utterances including those in classroom speech, he observes that the most natural kind of discourse in which it is *appropriate, and perhaps necessary, to state things that are already known in a certain order or sequence is in the presentation of an argument* (Levinson, 1979). This is exemplified in 24, from a courtroom interrogation between an attorney for the defense (A) and a girl (G). In 24-11, the girl's age is asked, even though the facts of the case such as her age are already known to all parties:

(24) (1) A: You have had sexual intercourse on a previous occasion, haven't you?
(2) G: Yes
(3) A: On many previous occasions?
(4) G: Not many.
(5) A: Several?
(6) G: Yes
(7) A: With several men.
(8) G: No
(9) A: Just one.
(10)G: Two
(11) A: Two. AND YOU ARE SEVENTEEN AND A HALF?
(12) G: Yes.

Levinson argues that the point of getting the witness to state what is known to everyone present is not in the question itself or in its answer, but in its JUXTAPOSITION with what has gone before. This juxtaposition manages to suggest that a girl of seventeen who has already had relations with two men is not a woman of good repute.

Levinson's notion of juxtaposition anticipates the analysis here in that the interpretation of a proposition is highly dependent on context and that inferences must be derived from currently salient propositions. As Levinson observes, the function of the question in 24-19 above is to extract from the witness one of a **series** of answers that contribute to forming a 'natural' argument for the jury. Given the fact that the context contributes the meaning of this utterance and that no

two contexts are exactly alike, it is surprising that there would be any hard constraints on the re-evocation of a proposition in discourse.[6]

Cohen's analysis of argumentative discourse also notes the occurrence of IRUs in argument and claims that using IRUs in an argument reduces hearers' processing (Cohen, 1987). This is because, in her model, a hearer must infer whether an utterance is a claim, or evidence for a claim, in order to infer the structure of the discourse. Since, according to Cohen, an IRU cannot be used as a claim, it is easy for hearers to recognize that an IRU must be the evidence for a claim.

One type of Consequence IRU, also related to argument structure, is related to a verb-preposing construction called Proposition Affirmation (Ward, 1985; Ward, 1990).[7] An example of PROPOSITION AFFIRMATION (henceforth PA) is shown in 25:

(25) Tchaikovsky was one of the most tormented men in musical history. In fact, one wonders how *he managed to produce any music at all. BUT PRODUCE MUSIC HE DID.* [WFLN Radio, Philadelphia] (Ward's 96, (Ward, 1990))

Ward argues these IRUs are felicitous because the affirmed proposition is doubtful or contrary to expectation. Horn (1991) points out that the 'surprise' condition is not necessary, since 26 is perfectly felicitous:

(26) It's unfortunate that *it's cloudy in San Francisco this week*, but CLOUDY IT IS – so we might as well go listen to the LSA papers.

Since everyone knows what San Francisco weather is like, there can be nothing surprising or unexpected about this weather report.

Horn proposes replacing Ward's 'surprising' or 'unexpected' condition with a constraint of RHETORICAL CONTRAST, noting that whenever the affirmed clause can be introduced by *but*, affirmation is possible. Furthermore, the basis for contrast can be an opposition in argument structure. This class of IRUs and Horn's analysis will be discussed in more detail in section 8.2.

Schiffrin also notes that the cue word *but* often cooccurs with IRUs. She presents a schemata for the use of *but* with IRUs (Schiffrin, 1982; Schiffrin, 1987), formulated as:

- **Schiffrin's schemata for IRUs with 'but'**

[6] Ralph Weischedel (p.c. 1992) points out that the use of repetition in classic poetry depends on the way the context changes throughout the poem so that the same line repeated in multiple stanzas is interpreted differently each time it is said.

[7] Cases where the affirmed proposition is entailed by the context are a subset of all cases of preposed VPs discussed in (Ward, 1985; Ward, 1990).

1. *information X*, disclaimer of X, but INFORMATION X
2. *information X*, information Y, but INFORMATION X

Note that I have used my notation of italics for the antecedent of an IRU and upper-case for the IRU in Schiffrin's schemas. The only constraint on *information Y* in schema 2 is that it be 'related' information. Example 27 illustrates schema 1 with the clause *We're not the one to judge* as a disclaimer.

(27) Debby: Would you say this is a friendly block?
Jack: *Fairly friendly.* Wouldn't you say?
We're a bit prejudiced, I think. Ah because uh we've been here so long that we don't even remember the original groups that were here.
So we're bad to judge.
We're not the one to judge.
But I WOULD SAY FAIRLY, FAIRLY FRIENDLY.

Example 28 illustrates the second schema. According to Schiffrin, the IRU in 3 is simply 'related information'. However while 3 isn't a disclaimer of the speaker, since disclaimers are always a type of related information, it is similar to one in reflecting the speaker's reasoning about the truth of the proposition asserted. The IRU in 4 is affirmed in the face of this related information.

(28) (1) See this one right here?
(2) *He's smart.*
(3) He himself don't think he's smart,
(4) but HE'S SMART.
(5) *He came in first* in plumbing,
(6) out of a hundred thirty five,
(7) He was the only Jewish kid.
(8) HE CAME IN FIRST.

The IRU shown in 28-8 belongs to a class that Schiffrin identifies as 'intensifiers'. The intensifying effect of 28-8 is because 28-7 emphasizes the uniqueness of coming in first. This example will also be discussed in section 8.2.

2.3.3 IRUs have a Social Purpose

Tannen (1989) analyzes 'repetition' in casual dinner table conversation and emphasizes that IRUs can be used for social reasons: she argues that the overlapping or closely sequential repetition, of part or all of others' utterances is used to maintain participation in conversation. This use of IRUs may be one that is peculiar to the discourse situation that Tannen studied; however it is also possible that the general function of Attitude IRUs, demonstrating that a belief is mutual, would also explain this 'participatory' function. A cognitive function of IRUs could be conventionalized so that, whereas the IRU may reduce uncertainty about mutuality in one situation, the same utterance in another situation can function to reaffirm the mutuality of a belief and thus achieve a social purpose.

2.3.4 IRUs can increase Reliability in Communication

In Finin, Webber and Joshi's (1986) work on advice-giving they note that paraphrase is used in conversation for three reasons, all of which are related to reliable communication. First, if the hearer notices an ambiguity or vagueness in the speaker's utterance, the hearer can produce multiple paraphrases of the utterance, each corresponding to and highlighting a different alternative. Second, even without noticing an ambiguity or vagueness, the listener may attempt to paraphrase the speaker's utterance to show how it has been understood, looking for confirmation or correction from the speaker. Finally, the hearer may paraphrase the speaker's utterance in order to confirm his or her belief that there is a common understanding between them.

In the first case, I would expect the paraphrase to be produced with a phrase final rise, while the second and third cases are examples of Attitude IRUs. Note that without prosodic marking, the second and third alternatives may be indistinguishable from the point of view of the speaker.

Clark and Schaefer also present a theory of dialogue in which repetition and paraphrase provide positive evidence of understanding(Clark and Schaefer, 1987; Clark and Schaefer, 1989; Brennan, 1990). According to their account there are five types of evidence of understanding:

- continued ATTENTION, e.g. via gaze,
- IMPLICIT acceptance, ie. going on with next turn,
- ACKNOWLEDGEMENTS such as *ok, uh-huh*,
- REPEAT of all or part of what the other said
- PARAPHRASE, demonstrate what you understood

The repetitions and paraphrases that Clark and Schaefer discuss are treated as Attitude IRUs here, and the analysis presented in chapter 6 draws on their analysis. Heritage and Watson (1979) also analyze a class of utterances they call 'reformulations' which are in some cases informationally redundant. Their analysis is that these utterances are the participants' way of demonstrating to one another what it is they think they are doing and their level of understanding of the evolving conversation .

Both of these analyses is consistent with the analysis of Attitude IRUs presented here. The need for participants to demonstrate to one another what they are doing follows from the fact that conversants are not omniscient, are autonomous, and cannot read one another's minds, so whether in fact beliefs are shared is a problem that must be managed among the conversants. By demonstrating their view of what is going on in the conversation, conversants facilitate the coordination of their discourse models.

2.3.5 IRUs can be used for Irony

Sperber and Wilson present an analysis of IRUs that are used ironically, based on the distinction between the 'use' of a proposition and its 'mention' (Sperber and Wilson, 1981). This is based on the standard use/mention distinction between talking about the entity denoted by a string in the language (use) and talking about the string itself (mention). In S&W's analysis, a proposition is used when it is asserted, and it is mentioned whenever the speaker utters it but does not assert it. This distinction is the basis of their analysis of irony as a type of echoic mention. It seems that the basis of this analysis is that speakers can predicate an attitude toward a proposition P by saying an utterance that realizes P in a certain way, in a context in which other interpretations (simpler interpretations) of the utterance are not felicitous. This function is distinct from that of IRUs discussed here since none of the IRUs in my corpus are used ironically.

2.3.6 IRUs manage Interaction

Whittaker and Stenton argued that utterances with 'no new information' function at the level of CONTROL of initiative in advisory dialogues (Whittaker and Stenton, 1988). On their account, prompts, adjacent repetitions, paraphrases and summaries can all be analyzed as utterances in which a speaker indicates that s/he does not wish to maintain the initiative in a mixed-initiative dialogue. In some cases, e.g. adjacent repetitions, this leaves the current speaker in control, while, in other cases, e.g. end of segment summaries, these utterances result in the current initiator

abdicating initiative and thus closing the current discourse segment. Walker and Whittaker analyzed a range of dialogue types using W&S's framework and investigate the relationship between control and other indicators of discourse structure such as the distribution of anaphora (Walker and Whittaker, 1990). W&W show that IRUs correlate strongly with other indicators of discourse structure. The theory of the function of both Attitude and Attention IRUs presented here draws on and expands these analyses.

2.3.7 Tautologies and Prompts

The Gricean account of tautology is that tautologies get their communicative import from the obvious flouting of the Quantity maxim: *Do not make your utterance less informative than is required* (Grice, 1967; Levinson, 1983). Since this requires speakers to be informative, some informative inference must be made to preserve the assumption that the speaker is cooperating. Thus, in the case of 29a, Levinson suggests that the inference might be that in 29b:

(29) a. War is war.

b. Terrible things always happen in war. That's its nature, and its no good lamenting that particular disaster.

An important factor in the interpretation of tautologies is that the subject matter of the tautology must already be up for discussion. Consider 30:

(30) She will either come or she won't.

This makes sense only if the participants are currently discussing **whether** she will come. Levinson suggests that a more general interpretation for most tautologies is that there is nothing more to be said about the topic. This can arise as a natural inference because anybody who says 30 clearly doesn't know whether or not she will come, so no more information can possibly be forthcoming.

The 'nothing more to be said' function of tautologies is similar to the close segment IRUs described in (Whittaker and Stenton, 1988), to be discussed in chapter 7. However, it is not clear, without access to a number of naturally-occurring tautologous utterances in their context of utterance, whether in fact tautologies have a conventional use of closing a segment. This must be left to future work.

Ward and Hirschberg (1991) provide an account in which the meaning of a tautology is derived as a denial of the truth or relevance of alternative predications that a speaker might have made

of an entity . W&H suggest that these denied alternatives $\mathcal{B}$ will often be readily available in the discourse context at the point where the tautology is said. For example, 29 is an EQUATIVE tautology schematically characterized as *a is a*. The speaker's uttering the equative tautology is interpreted in the light of a set of alternative utterances defined as *a is b* or *some a are b* for a $b \in \mathcal{B}$. None of the IRUs studied here seem at all similar to this denial function.

Another type of informationally redundant utterance are PROMPTS such as *uh huh*, which add no new propositional content to the common ground. Like tautologies, prompts don't have an antecedent in the dialogue; no matter what has been said previously, a prompt can add no new information. These are said to be 'continuers', i.e. they indicate a choice by the speaker to pass up a turn while prompting the current speaker to continue talking (Schegloff, 1982), and they share a number of properties with Attitude IRUs as to their function in dialogue (Whittaker and Stenton, 1988; Walker and Whittaker, 1990; Hockey, 1991). This shared function has been noted by classifications in which adjacent repetitions are analyzed as backchannels, e.g. the utterance of *take the money* in example 13. Also like tautologies and close-segment IRUs, a prompt adjacent to another prompt can function as a coordinated closing of a segment (Schegloff and Sacks, 1977). The commonality between prompts and Attitude IRUs will be briefly discussed in section 6.3.

2.4 Summary

This chapter has reviewed previous arguments as to the discourse functions of IRUs. While none of the previous work has focused on the relationship of IRUs to resource-bounds, many of the discourse functions of IRUs have previously been noted. The following chapters will extend and elaborate on these discourse functions and argue that they are related to resource limitations.

Chapter 3

Limited Attention and Limited Reasoning

3.1 Introduction

In chapter 1, I introduced the hypotheses that Attention IRUs are the result of agents' limited attentional capacity, that Consequence IRUs are the result of agents' limited inferential capacity and that Attitude IRUs reflect the process of coordinating the conversants discourse models given these limitations.

This chapter presents the theoretical framework for the analysis of IRUs in the following chapters. The framework is based on an architecture for resource-limited reasoning (Bratman, Israel, and Pollack, 1988), a model of belief revision (Gardenfors, 1988; Harman, 1986; Galliers, 1990; Gardenfors, 1990), and a model of limited attention (Landauer, 1975). The point of the different components of this framework is to model agents as having the resource limitations related to IRUs: limited attention and limited inference. Within this resource-limited model, the cognitive function of IRUs can be demonstrated.

First section 3.2 presents the model of limited attention/working memory. The limited attention model operationalizes the concept of SALIENT information, and provides the basis for the function of both Attention IRUs and some types of Consequence IRUs. The model of belief revision presented in section 3.3.2 represents the different kinds of HEARER OLD information needed to support the analysis of Attitude and Consequence IRUs. Both of these components will be integrated into the IRMA architecture presented in section 5.4.

3.2 Limited Attention

Chapter 1 discussed the distinction between the information that is hearer old (known) and information that is salient. The important distinctions in hearer old information are represented by the model of belief presented in section 3.3.2. This section describes a model of discourse salience, used as the basis for Struct-$\mathcal{P}$, the structure of propositions in the discourse model. I leave open the question of whether Struct-$\mathcal{P}$ is defined by the same constraints as Struct-$\mathcal{D}$, the structure of accessible discourse entities for anaphora resolution and domains of quantification.

One reason that it is important to have a model of discourse salience is because this thesis argues for the DISCOURSE INFERENCE CONSTRAINT, repeated here from chapter 1:

> DISCOURSE INFERENCE CONSTRAINT: Inferences in dialogue are derived from propositions that are currently discourse salient (in working memory). [1]

The idea that some beliefs are more salient than others is well supported in the literature. It is well known that human agents have limited attentional capacity (Anderson and Bower, 1973; Miller, 1956; Landauer, 1975). Furthermore, several previous computational accounts of reasoning and inference have proposed that only a subset of beliefs is used in reasoning at any one time (Joshi, 1978; Webber and Joshi, 1982; Joshi, Webber, and Weischedel, 1984; Fagin and Halpern, 1985). The role of this subset of beliefs is a reduction in computation that results from reasoning over a smaller set of facts.

The attentional limit on inference plays a crucial role in explaining a pernicious problem in computational models of reasoning. It is well known that human agents are not logically omniscient, but which factors determine a restricted set of inferences is still an open question. The proposal that inference is restricted by limited attention makes intuitive sense and is supported by empirical results (Kintsch, 1988).[2]

However, the observation that attentional capacity is limited doesn't provide us with an operationalization of what exactly the limits are and how beliefs become salient and lose their salience. Therefore, in order to test the hypothesized DISCOURSE INFERENCE CONSTRAINT without solving this difficult open issue, section 3.2.1 proposes a simple model of limited attention/working memory (AWM). Attentional capacity can be parameterized in this model, thus supporting testing hypotheses about the effects of different limits.

[1] The distributional analysis of CONSEQUENCE IRUs presented in chapter 8 will provide support for the claimed relationship between salience and inference.

[2] It may also be related to results on the complexity of epistemic reasoning, namely that operators that pull together two facts from memory and juxtapose them via and-introduction or or-introduction rules increase the complexity of the underlying model (Vardi, 1989).

3.2.1 Landauer's Attention Working Memory Model

The attention/working memory model, AWM, is a very simple model, adapted from Landauer's 'garbage can' model (Landauer, 1975). In this model, AWM consists of a three dimensional space in which propositions are stored in chronological sequence according to the location of a moving memory pointer. Storage of propositions depends on the chronology of events in the world and propositions that are encountered multiple times are stored multiple times. Retrieval processes start from the current location of the moving memory pointer and search through memory in a spreading search pattern, with the effect that items stored more recently are found with less search. Landauer suggests that a garbage can is a useful metaphor, because items are stored there in chronological sequence with the effect that coffee grounds and orange peels may be found near one another simply because they are eaten together at breakfast. The details of the storage and retrieval operations will be discussed in the following sections.

While the AWM model is extremely simple, Landauer showed that it could be parameterized to fit many empirical results on human memory and learning (Hellyer, 1962; Landauer, 1969; Oldfield and Wingfield, 1965; Collins and Quillian, 1969; Sternberg, 1967; Tulving, 1967; Anderson and Bower, 1973). A key aspect of the model for studying IRUs is that it incorporates both recency and frequency effects. The multiple copy aspect means that the model predicts the spaced-practice memory effect: the frequency of rehearsal of a fact improves the subject's ability to recall it. This prediction is one of the reasons that the model is attractive for studying the effects of IRUs.

Sections 3.2.1.1 and 3.2.1.2 discuss the details of the storage and retrieval operations. Section 3.6.1 will compare AWM with more sophisticated models of Attention in discourse. The relation of the DISCOURSE INFERENCE CONSTRAINT to other theories of discourse will be discussed in section 3.6.2.

3.2.1.1 Storing Beliefs in AWM

The sequence of memory loci used for storage constitutes a random walk through memory with each loci a short distance from the previous one. This is depicted in figure 3.1. The current memory pointer is the memory location at the end of the path shown in the figure. Thus propositions which tend to co-occur in events in the world will also tend to be stored near one another in memory, but the selection of the exact memory location is stochastic. If items are encountered multiple times, they are stored multiple times. Memory is also wrap-around as shown by the path of the pointer in the figure.

Additional assumptions not made specific in Landauer's original model are that propositions are stored (rather than 1s and 0s) and that there is no overwriting. If the path of the memory pointer

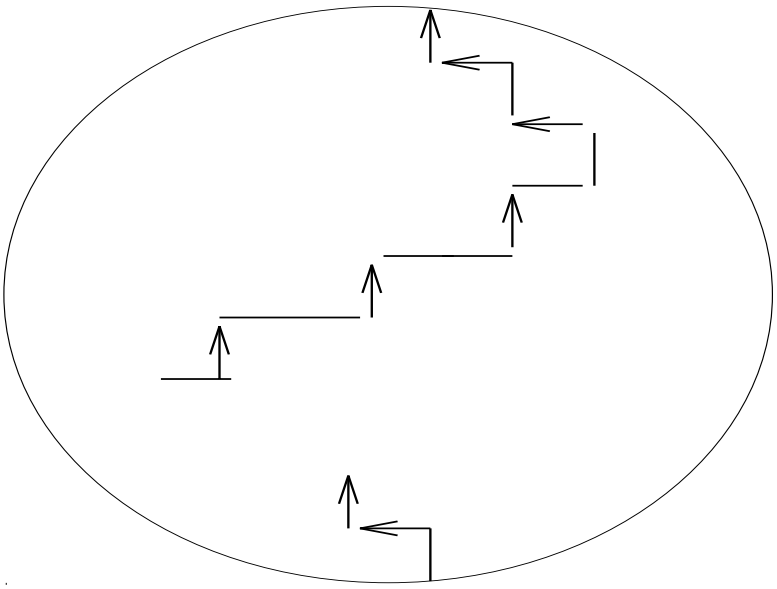


Figure 3.1: Memory Structure: Three Dimensional Store, Random Storage

retraces its steps so that the current memory loci already has something stored in it, the new item is simply added. Thus memory capacity is unbounded, and it is possible for two completely unrelated items to become associated accidentally by virtue of being stored in the same loci.

3.2.1.2 Retrieving Beliefs in AWM

When an agent retrieves items from memory, search starts from the current pointer location and spreads out in a spherical fashion. See figure 3.2. Search is restricted to a particular search radius: radius is defined in Hamming distance. For example if the current memory pointer loci is (0 0 0), the loci distance 1 away would be (0 1 0) (0 -1 0) (0 0 1) (0 0 -1) (-1 0 0) (1 0 0).[3] This limit on the radius defines the capacity of attention/working memory. Since memory is wrap-around, searches that start near the 'edge' of memory, also wrap-around as shown in the figure. Because search starts from the current pointer location, items that have been stored most recently are more likely to be retrieved, predicting recency effects. Because items that are stored in multiple locations are more likely to be retrieved, the model predicts frequency effects (Landauer, 1975).

In addition, in the version of AWM developed here, search can be exhaustive within a radius value, or agents can abort a search as soon as a relevant belief is found. Whether search is exhaustive or not depends on the agent's goal at the time of searching memory and the concomitant memory retrieval pattern. One of the parameters that can be measured is the number of memory loci searched during a particular retrieval or over the course of performing a task.

[3]The actual locations are calculated modulo the memory size.

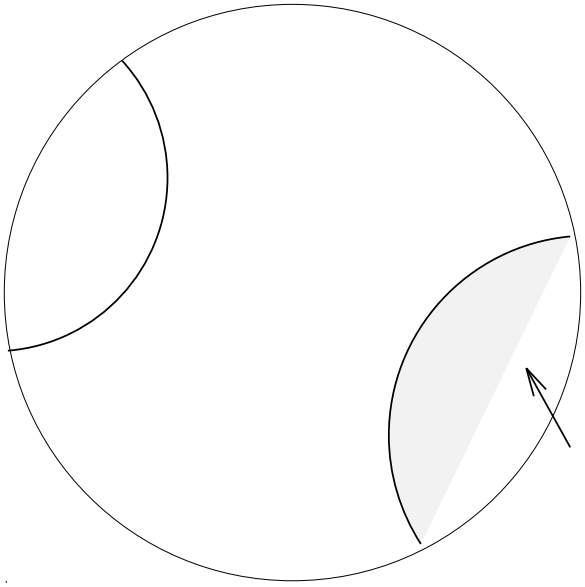


Figure 3.2: Searching Memory: Spreading Search for a Fixed Radius

The point of an explicit model for AWM is to provide an operationalization of discourse salience. Various assumptions about discourse salience can be tested by varying the radius of the search sphere. The radius of the search sphere is the main parameter for the agents' resource-bound on attentional capacity. This is a fundamental parameter that is varied in the Design-World simulation environment to be discussed in section 5.1. Note that the model is simple in that recency is the primary determinant of SALIENT. Other models of attention in discourse will be discussed in section 3.6.1.

3.2.1.3 Example of Retrieval in AWM

Figure 3.3 illustrates a simple example of the model in operation. Note that three dimensions are collapsed onto two dimensions in the figure. Let's say that propositions P,Q,R,S,T,U,V,W,X,Y,Z are stored in sequence in memory. The storage locations for a particular run are, in sequence: (12 5 11), (12 6 11), (13 7 11), (12 6 11), (11 6 11) (10 7 11), (10 8 10), (10 8 9), (11 8 10), (12 8 10), (12 7 9). Note that each memory loci is within Hamming distance 2 of the previous (two city blocks). At the end of this sequence of storage operations, the agent's memory pointer is (12 7 9). A retrieval always starts from the current memory pointer location. If AWM radius is set to 2, the propositions retrieved are Z Y. If AWM radius is set to 3, the propositions retrieved are Z Y S Q X W R. At AWM radius of 4, all of the propositions are retrieved. However of course since the model is stochastic, this pattern can vary from one run to another. On another run, at AWM radius of 4, only propositions W X Y Z were retrieved.

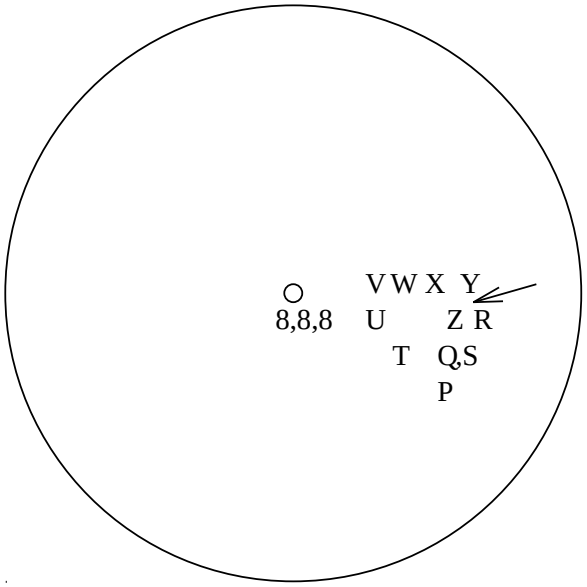


Figure 3.3: Example of Searching Memory: Spreading Search for a Varying Radii

3.2.1.4 Storing Beliefs as a Result of Cognition

As a departure from Landauer's original formulation, the version of AWM here treats cognitive events in which propositions are retrieved and reasoned about as though they were events in the world. Propositions retrieved and reorganized by reasoning are re-stored in memory at the current memory pointer location. Thus retrieval and reasoning locally reorganizes propositions in memory. Using the garbage can metaphor, allowing retrieval and reasoning to re-store propositions is analogous to going through the garbage can with a specific search criteria, copying items, and putting them back in the can at the top. Allowing this additional way of organizing and storing items in memory provides a plausible explanation for associative links in memory (Anderson and Bower, 1973; Collins and Quillian, 1969).

3.2.2 Summary

This section has presented the AWM model of attention and working memory. AWM is a simple model that fits a range of empirical results on memory and learning. The motivations for AWM are:

1. AWM operationalizes the notion of discourse salience.
2. AWM is parameterizable so that the degree to which attentional capacity is limited can be varied.

3. AWM demonstrates frequency effects. Because the benefits of rehearsal in learning are well known, it is necessary to see whether some of the benefits of redundancy are rehearsal benefits.

4. AWM's storage and retrieval model is flexible enough to support the development and testing of additional memory related strategies.

5. AWM is implementable.

Section 3.6 will compare AWM with other views of attention in discourse. In section 3.3.2, I will discuss in more detail the relationship between the representation of beliefs and AWM. Then in section 5.8, once the Design-World simulation domain has been described in enough detail, I will show that the AWM model produces a main effect for attention limitations in Design-World. This model will support testing the effects of different communication strategies.

3.3 Belief and Intention Deliberation

Deliberation is the process by which an agent explicitly or implicitly evaluates a set of alternates in order to decide what to believe and what course of action to pursue (Doyle, 1992).[4]

There are two deliberation relations: SUPPORT and WARRANT. The SUPPORT relation holds between two beliefs when believing one is a reason for believing the other; this relationship can hold at various endorsement levels to be discussed below. The WARRANT relation holds between a belief and intention when the belief is a reason for adopting or having the intention, e.g. the belief that you will make a 15% profit may provide a WARRANT for an intention to purchase Hewlett Packard stock.[5] These relations are dependent on the right kind of implication, entailment or utility relation holding or being plausible in the domain. The use of these relations in what follows is reflected in the following COHERENCE ASSUMPTION:

- COHERENCE ASSUMPTION: Beliefs and intentions that are subject to deliberation are evaluated for their coherence with other beliefs via the relations of SUPPORT and WARRANT.

Agents deliberate about **whether** as well as **how** to revise their beliefs and intentions as they receive new information; this is the ATTITUDE assumption (Galliers, 1991a; Walker, 1992a):

- ATTITUDE: Agents deliberate whether to ACCEPT or REJECT an assertion or proposal made by another agent in discourse.

[4]Evaluation functions (utilities) for beliefs have a different basis than those for intentions.

[5]The SUPPORT relation might be an abstraction of the two RST 'presentational' relations of evidence and justify. The WARRANT relation might be an abstraction of the two RST relations of motivation and concession. See (Hobbs, 1979; Mann and Thompson, 1987; Moore and Paris, 1989) for more detailed sets of relations.

Thus communication is a strategic process aimed at getting other agents to revise their beliefs. Empirical analyses of dialogue informs an account of deliberation because dialogue provides an explicit protocol of which facts agents believe will affect the ACCEPTANCE or REJECTION of an assertion or proposal. An analysis of IRUs in problem-solving dialogues shows that the process of deliberating about beliefs depends on the type of evidence supporting a belief, and that one of the primary functions of IRUs is to upgrade the strength of the evidence supporting beliefs (Walker and Whittaker, 1990; Walker, 1992b). The process of deliberating about intentions also depends on evidence supporting beliefs that the intention is based on, which can contribute to a perception of 'risk'. However, there is an additional **independent** factor that contributes to deliberating about intentions: the utility of the resulting course of action (Pollack, 1990). IRUs function communicatively to support both deliberative processes.

Section 3.3.1 will discuss intention deliberation and section 3.3.2 discussed belief deliberation.

3.3.1 Intention Deliberation

The general rule for intention deliberation is that *if one course of action, A, produces greater benefits than another course of action, B, pursue A over B.* Thus intention deliberation relies on way to evaluate different courses of action. A theory of intention deliberation should predict when an agent will reject a proposal made by another agent in discourse.

Decision theory is the most widely used theory of how agents evaluate alternate courses of action, and is operationalized through probability and the notion of expected utility. The key tenets of decision theory are that (1) all possible outcomes are available to be evaluated (2) a single evaluation function can be applied to these outcomes, and (3) agents use the results of the evaluation function to select the best course of action.

The problems with this formulation are that agents may not know what all the possible outcomes are, it may be difficult to estimate the probability of the different outcomes, there may be incompatible competing evaluation functions, and the agent may not have the resources to evaluate all the possibilities. How agents should choose a course of action given these problems is an active area of research (Bratman, Israel, and Pollack, 1988; Pollack and Ringuette, 1990; Dean and Boddy, 1988).

The analysis of dialogue presented in the following chapters contributes to a theory of intention deliberation by providing a protocol of factors considered in deliberating intentions. The class of Deliberation IRUs dicussed in section 7.2 show that IRUs are often used to support deliberating about intentions and beliefs. The class of Affirmation IRUs discussed in section 8.2 provide evidence

that agents take many types of information into account when deliberating, that there are multiple competing evaluation functions, and thus many different types of WARRANTS. Thus the analysis contributes to a theory of intention deliberation and supports a simple constraint on intention deliberation: the DISCOURSE INFERENCE CONSTRAINT. This constraint posits that deliberation is constrained by an agent's attentional capacity.

Design-World will test the DISCOURSE INFERENCE CONSTRAINT and its effect on deliberation. The Design-World task simplifies the problem of multiple competing evaluation functions by associating scores with each course of action so that propositions about scores provide the WARRANTS for actions.

3.3.2 Belief Deliberation

The account of belief revision used here relies on four assumptions from current theories of belief revision:

- There are three Attitudes with which beliefs can be held: Accepted, Rejected or Indetermined.[6]
- Agents reason autonomously about whether to accept or reject incoming information. This is the ACCEPTANCE ASSUMPTION.
- Decisions abut acceptance or rejection depends on evaluating sets of beliefs according to a measure of coherence.
- Coherence of belief sets is determined by relationships among supporting beliefs. Types of support are reflected in endorsements on supporting beliefs. Endorsements are based on source or hearer-old information type.

These assumptions are supported by current theories of belief revision. In particular, the first three are part of Gardenfors' theory of Epistemic Entrenchment and the last three are part of Galliers' theory of Autonomous Belief Revision (ABR) (Harman, 1986; Gardenfors, 1988; Gardenfors, 1990; Galliers, 1990; Galliers, 1991b; Galliers, 1991a; Cawsey et al., 1992).

The general process of reasoning about beliefs involves the EXPANSION, CONTRACTION or REVISION of an existing belief set. For example, imagine that an agent A is entertaining a belief that P. A may decide either to accept or reject P. EXPANSION of A's beliefs occurs when P is simply added to A's belief database. CONTRACTION occurs when A decides that he no longer believes $\neg$ P. REVISION

[6]This perspective focuses on agents' attitudes towards beliefs rather than whether the belief is true or false in a model.

is the process of first contracting a belief set, and then expanding it: say for instance if P implies Q, and A accepts that P, A must first contract his beliefs by ¬ Q and then expand his beliefs by P and Q.

Typically there is more than one way that a set of beliefs can be revised. Consider a simple example: suppose you believe P → Q, where P is *it is raining* and Q is *Oscar wears a hat.* In addition, you believe that ¬ P → R, where R is *Oscar might wear a hat.* If you believe it is raining, then you will come to believe that Oscar is wearing a hat. However, suppose now that you see Oscar and he is not wearing a hat. Imagine that no other beliefs are related to these. Clearly, in order to maintain consistency in your beliefs, you must either give up the belief that P, or give up the belief that P → Q. What is required is a way of choosing among alternate revisions. [7]

Galliers argues that the ENDORSEMENT on a belief is a factor that determines its corrigibility, i.e. how easily it will be given up in a contraction. This is supported by the use of IRUs in the corpus. For example, perhaps P is strongly believed because it has been asserted by a number of other agents and P → Q was inferred by default inference. Then it will be easier to give up the belief in P → Q than to give up the belief that P.

Endorsement types are discussed in section 3.3.2.1. Section 3.3.2.2 discusses how the endorsements on premises affect the endorsements on beliefs derived from those premises. Deliberation must also take discourse salience into account. This will be discussed in section 3.3.2.4 where a simple model of evaluation of belief sets is proposed based on these factors.

3.3.2.1 Endorsements

In Galliers' theory, endorsements reflect both the source of belief and agents preference about what they want to believe. The general idea is that beliefs are tagged with an endorsement type when they are first formed and stored in memory, and that these endorsements contribute to the degree to which a belief is epistemically entrenched, i.e. endorsements provide a qualitative way of distinguishing between beliefs that are defeasible and those that an agent would rarely change. The types of endorsements used in this thesis are based on the logical types of propositions in discourse identified in chapters 1 and 2 and motivated by the distinctions necessary to explain the function of IRUs. These are:

[7] Gardenfors proposes that choices among alternate revisions are constrained by a set of logical consistency postulates and an ordering on beliefs, ≤, called EPISTEMIC ENTRENCHMENT. A ≤ B should be read as 'B is at least as epistemically entrenched as A'. Epistemic entrenchment is meant to reflect how easy a belief is to give up and beliefs that are more epistemically entrenched are harder to give up. Gardenfors shows that the postulates and the ordering provide a way of choosing among alternate revisions and that the ordering also provides desirable properties of a theory of belief such as being able to say that one belief is a reason for believing another however he leaves open the question of which factors determine the degree to which a belief is epistemically entrenched (Gardenfors, 1990).

ORDERING ON EPISTEMIC ENDORSEMENT TYPES:

HYPOTHESIS < DEFAULT < ENTAILMENT < LINGUISTIC < ABSOLUTE.[8]

The type HYPOTHESIS is the weakest possible endorsement. It is used to endorse assumptions that have no evidence supporting them at all, and DEFAULT is used for defeasible inferences such as implicatures and the interactive defaults discussed in section 6.2 (Joshi, Webber, and Weischedel, 1986). Entailments and presuppositions that have not been explicitly discussed are endorsed as ENTAILMENT. The type LINGUISTIC endorses assumptions that have been made explicit in the dialogue and ABSOLUTE refers to assumptions that are incontrovertible as though from divine authority. In addition, the same assumption can have multiple endorsements; it may be both entailed by what has been said as well as said explicitly, i.e. it has endorsements of both types ENTAILMENT and LINGUISTIC.

The role of the ordering on the types of endorsements reflects the **relative** defeasibility or corrigibility of different assumptions: an assumption endorsed as a DEFAULT may be defeated by LINGUISTIC information. For example, suppose that the belief is that *Madison can swim.* This is based on two assumptions: (1) *Madison is a dog*, (2) *Dogs can swim.* Let's say assumption (1) is something that an agent can see with his own eyes, so it might be strongly endorsed, e.g. ABSOLUTE. In the absence of other evidence, the agent makes a default inference about dogs and their ability to swim. Assumption (2) is thus endorsed as a DEFAULT. Since (2) is only a default, it can be easily defeated. If Madison's owner comes along and tells the agent that *Madison can't swim*, they are likely to abandon their belief that Madison can swim, i.e. the belief that *Madison can swim* is defeated.

Endorsements on assumptions can also be upgraded, and thereby made less defeasible. For instance, Madison's owner might come along and tell the agent that *Madison is a very good swimmer.* The belief that *Madison can swim* now has an endorsement type of LINGUISTIC.

3.3.2.2 Combining Endorsements

An open issue is the specification of how to combine the endorsements on assumptions. In other words, if I believe P and P $\rightarrow$ Q, what is the endorsement on Q? It seems that this should depend on the endorsements on P and P $\rightarrow$ Q. In previous work (Walker, 1992b), I adopt a simplifying rule that a chain of reasoning is only as strong as its weakest link.

[8] Endorsement types have also been called BASES for belief in (Lewis, 1969; Schiffer, 1972; Clark and Marshall, 1981). See also the ranking on HEARER OLD information proposed in (Prince, 1981b; Clark and Marshall, 1981) and a larger set of endorsement types proposed in (Galliers, 1991b; Cawsey et al., 1992)).

WEAKEST LINK RULE: The endorsement of a belief P depending on a set of underlying assumptions $a_i, ...a_n$ is MIN(endorsement $(a_i, ...a_n)$)

This WEAKEST LINK RULE seems intuitively plausible and means that the endorsement of a belief depends on the endorsements of the underlying assumptions. It also means that for all inference rules that depend on multiple assumptions, the endorsement of an inferred belief is the weakest of the supporting beliefs. Since the inference rule itself is one of the supporting beliefs, and it has an associated endorsement, this means that some kinds of inferences, such as implicatures, cannot be believed as more than defaults no matter how strong the endorsements are on the assumptions.

3.3.2.3 Evaluating Coherence of Sets of Beliefs

We now have a way of comparing two beliefs and predicting which will be given up in revision.[9] However the idea is that sets of beliefs are evaluated for their coherence with one another, rather than individual beliefs.

Galliers addresses this problem by defining a way of evaluating alternate belief states. Belief sets are evaluated by comparing the coherence of the sets using an ordering of MORE-COHERENT, ie. belief state $\mathcal{A}$ is preferred to belief state $\mathcal{B}$ if $\mathcal{A}$ is MORE-COHERENT than $\mathcal{B}$ (Galliers, 1991a). Galliers defines the MORE-COHERENT relation with a three tiered system which depends on an additional construct of CORE BELIEFS, and a combination of deduction and endorsements on beliefs. Agents prefer belief states that support their core beliefs. The problem with Galliers' definition of the MORE COHERENT relation is that it depends on checking derivations for core beliefs over complete belief sets. This method is psychologically implausible given the AWM model of limited attention.

A solution is to assume that evaluation only operates on SALIENT beliefs so that the DISCOURSE INFERENCE CONSTRAINT constrains both revision and reasoning (Solomon, 1992). This means that at any one time belief sets consisting of only a small number of beliefs are compared, and thus belief sets are evaluated by a simple technique of counting how many beliefs are more strongly supported (Galliers, 1990). This formulation predicts that agents may allow their beliefs to be globally inconsistent, but once these inconsistencies are pointed out they will attempt to revise their beliefs so as to make them locally coherent.

Attitude IRUs provide evidence that default < linguistic and entailment < linguistic as an endorsement on mutuality. Affirmation IRUs show what types of beliefs can 'argue against' another belief.

[9] Along with other information about inconsistency such as that used by Gardenfors' postulates (Gardenfors, 1988).

Affirmation IRUs also show that agents take care to defeat conflicting defaults and that repetition may be used to demonstrate strength of belief or speaker commitment.

A problem for the treatment of Consequence IRUs is that both Galliers' and Gardenfors' models are based on the assumption that belief sets are closed under logical consequence. As discussed in chapter 1, theories that make this assumption must be augmented with additional distinctions to explain the function of utterances which make inferences explicit. This can be done by assuming that there is a difference between implicit and explicit beliefs (Levesque, 1984), and that agents are only aware of their explicit beliefs. Thus while epistemic states may be idealized to be closed under consequence, agents' reasoning and belief revision processes operate on explicit beliefs,

3.3.2.4 Belief Deliberation in AWM

The mechanisms proposed above of tagging beliefs with endorsements, and the proposed storage and retrieval mechanisms of AWM has two theoretical ramifications:

- PRINCIPLE OF POSITIVE UNDERMINING (Harman, 1986): It has been shown that beliefs persist when their supports are degraded (Harman, 1986; Tversky and Kahneman, 1982; Ross and Anderson, 1982; Galliers, 1990). By positing a loose association between beliefs and their supports, the AWM model predicts that beliefs can be retained even when their supports are degraded.
- PRINCIPLE OF ASSOCIATED SUPPORT: Agents can perceive that they believe a proposition with various strengths without having access to the reasons why they believe it. This is because the weakest link rule associates a belief with an endorsement at the time the belief is formed. This means that the endorsement on a belief is available even if the supporting beliefs are not explicitly available.

The principle of positive undermining was proposed by Harman (1986) to account for data presented in Tversky and Kahneman (1982) and Ross and Anderson (1982). The PRINCIPLE OF ASSOCIATED SUPPORT is proposed here as a desirable side effect of the proposed mechanism.

In addition, the AWM model makes certain predictions. The storage and retrieval operations for AWM means that some of the beliefs that were causal in forming a belief will be retrieved along with the belief because they were stored in memory at around the same time. In addition, sometimes agents can forget that they have changed their beliefs so that deliberation is affected by the frequency and recency of a belief. A belief stored in multiple locations is more likely to be retrieved, and thus available for deliberation. A belief that was stored recently is also more likely

to be retrieved and used in deliberation. This shows that the formation of beliefs and intentions is connected with the processes of manipulating attentional state via memory storage and retrieval.

The algorithm that agents use for deliberating about beliefs is:

1. Consider the set of salient beliefs currently in AWM. Each salient belief has an endorsement.
2. If the set of salient beliefs are inconsistent or incompatible, then generate alternate belief states as compatible sets of beliefs.
3. Evaluate alternate belief states by counting which state has the most strongly supported beliefs.
4. Store the results of deliberation at the current memory pointer locus.

3.4 Mutual Supposition

Sections 3.2 and 3.3 examined the cognitive processes of attention and belief deliberation internal to agents. This section examines how these internal processes affect how agents coordinate in dialogue and how they decide what is mutually believed. While it was convenient to talk about a single discourse model in section 1.2.1, in reality each conversant has his/her own discourse model. Certain conversational processes are aimed at keeping the two models coordinated. Thus it is clear that what has been called SHARED KNOWLEDGE, COMMON KNOWLEDGE, and MUTUAL BELIEF is more accurately described as an individual speaker's 'tacit assumptions' (Prince, 1978) or as 'mutual absence of doubt' (Nadathur and Joshi, 1983; Joshi, 1982). The MUTUAL SUPPOSITION account of mutual belief presented here models this 'absence of doubt' quality by representing the conversants' assumptions about mutuality as defeasible, depending on the evidence provided by the other conversants in dialogue. By building on the theory of belief revision discussed in section 3.3, beliefs about mutual beliefs can be revised as the discourse proceeds.

3.5 Shared Environment Model of Mutual Supposition

The common ground is a set of MUTUALLY SUPPOSED propositions, assumed to be shared between conversants in discourse based on a number of shared assumptions about conventions of language and shared background (Lewis, 1969; Stalnaker, 1978; Thomason, 1990). This is modeled by an explicit schema proposed by Lewis, called the SHARED ENVIRONMENT model of common knowledge (Lewis, 1969; Clark and Marshall, 1981; Barwise, 1988a). What Lewis called common knowledge

will be called mutual supposition here, reflecting the defeasible and uncertain nature of what is assumed to be mutual.

Because conversants don't have access to the mental states of other conversants, mutual supposition must be inferred, based on externalized behavior of various kinds. The inference of mutual supposition can be inferred using the MUTUAL SUPPOSITION INDUCTION SCHEMA, henceforth MSIS:

> **Shared Environment Mutual Supposition Induction Schema (MSIS)**
> It is mutually supposed in a population P that Ψ if and only if some situation $\mathcal{S}$ holds such that:
>
> 1. Everyone in P has reason to believe that $\mathcal{S}$ holds.
> 2. $\mathcal{S}$ indicates to everyone in P that everyone in P has reason to believe that $\mathcal{S}$ holds.
> 3. $\mathcal{S}$ indicates to everyone in P that Ψ.

The situation $\mathcal{S}$ of the MSIS is the discourse situation as defined in section 1.2. If each of the three conditions given in the MSIS is satisfied then the conversants are justified in inferring that a fact Ψ is mutually supposed. Condition (1) specifies that a public utterance event must be accessible to all of the discourse participants. Conditions (2) and (3) state that what is mutually supposed is derivable from the fact that all participants have access to this public event. In other words, what is believed to be mutually accepted is the set of mutual suppositions indicated by the occurrence of a sequence of utterance events in a discourse situation $\mathcal{S}$.

This account of mutual supposition is a weak model of what has previously been called mutual belief or common knowledge (Halpern and Moses, 1985; Lewis, 1969; Schiffer, 1972; Parikh, 1990; Barwise, 1988a; McCarthy et al., 1978; Fagin and Halpern, 1985; Vardi, 1989). The MUTUAL SUPPOSITION INDUCTION SCHEMA (MSIS) allows agents to infer mutual supposition in a finite amount of time using the finite decision procedure of checking three conditions rather than an infinite list of statements of the form (A believes (B believes (A believes....))) and (B believes (A believes (B believes ....))).

This finite decision procedure of checking the three conditions in the MSIS involves some 'risk' for the agent in the inference of mutual supposition, since what a discourse situation INDICATES can vary according to assumptions about background information and reasoning processes. What $\mathcal{S}$ INDICATES depends on the participants' interpretation of the utterance event, and the fact that this interpretation is mutual may be more or less endorsed (Nadathur and Joshi, 1983; Fox, 1987). One of the functions of IRUs that will be discussed below is reducing the 'risk' by the use of discourse strategies that (1) demonstrate what inferences are made; (2) provide evidence as to

what has been mutually accepted; and (3) make it easier for other agents to determine the relevant reasoning context and to make inferences that follow in that context. Section 6.2 will explain how the INDICATES relation is formalized using the belief model introduced in section 3.3.2.

3.6 Related Work on Belief and Attention Models

3.6.1 Discourse Structure and Attentional State

The model of attention/working memory, AWM, is an explicit model of limited attentional capacity, which models salience and non-salience by the simple measure that attentional capacity is limited and as new information comes in, old information must go out. In addition, frequency can have an effect on salience by making it easier to retrieve a proposition. However AWM is extremely simple; it is based solely on recency and frequency effects with no additional structure.

In contrast, most theories of discourse assume that discourse structure is hierarchical (Grosz, 1977; Sidner, 1979; Hobbs, 1979; Polanyi and Scha, 1984; Grosz and Sidner, 1986; Webber, 1986; Webber, 1988; Polanyi, 1987; Mann and Thompson, 1987). In these hierarchical models, sequences of utterances aggregate into discourse segments, with potential embedding among the segments. The embedding relation makes the discourse hierarchically structured.

Only Grosz and Sidner's (G&S) theory proposes an explicit relationship between discourse structure and a model of Attention. In G&S's theory, the hierarchical structure determines a stack-based model of attentional state. The point of this section is to compare this model of attentional state with AWM. I point out ways in which AWM may be too simple, as well as suggest ways that AWM might account for phenomena that hierarchical structuring has previously explained. I will also present examples from the financial advice corpus that suggest that G&S's model should take recency into account.

All of the discourse theories above define hierarchical discourse structure via embedding relations between discourse segments. Theories vary as to what determines the embedding relation, what counts as a discourse segment, and thus what the embedding relation is defined on. For example, most theories agree that there are two main types of discourse relations, coordinating and subordinating, but differ as to whether these are defined on text spans or on intentions underlying the discourse. In G&S's theory, the embedding is determined by the intentions of the conversants. The main intention for the whole discourse is called the discourse purpose (DP) and the intentions for the subsidiary segments are called discourse segment purposes (DSPs). Figure 3.4 shows how a sequence of utterances might be hierarchically structured, based on the utterance-level intentions

that they realize. DSPs may be realized more or less directly by an utterance, or left to be inferred; this is shown in figure 3.4 by the fact that DSP1 has no corresponding single uterance. In addition, multiple utterances may be required to realize a single intention (DSP); an example of this is the fact that DSP2 is based on the aggregation of $U_1 \ldots U_5$. Finally, utterances that are sequentially non-adjacent, but are treewise adjacent, may contribute to the same intention (DSP), such as U_1 and U_6.[10]

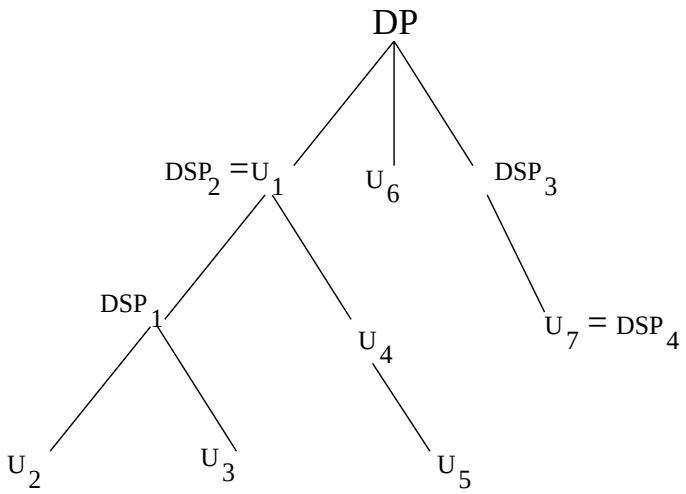


Figure 3.4: A Sequence of Utterances Structured into Discourse Segments

In G&S's theory, the hierarchical discourse structure is reflected in the attentional structure. Each intention has an associated focus space; the focus space contains those entities (objects, properties and relations) that are salient, either due to explicit mention or because they became salient in the process of producing or comprehending utterances in the segment. While G&S state that attentional state in their theory is a property of the discourse and not a property of the conversants, this view of how entities can become salient in a focus space suggests that aspects of G&S's model are modeling conversants' mental states, just as the AWM model is a simple model of conversants' mental states. However it is important to keep in mind that G&S did not intend their model to be a psychological one.

3.6.1.1 Stack Model of Attentional State

In G&S's model, the recognition of intentional structure is what determines the attentional structure. As a discourse is processed, the recognition of a subordinate intention results in PUSHING a focus space onto the attentional stack; each intention that contributes to some intention further up

[10]This means that there is an issue with determining whether an utterance alone with its utterance level intention counts as a DSP.

the tree (further down in the stack) results in first POPPING all the intervening focus spaces, and then pushing a focus space for the new intention on the stack (Sidner, 1979; Grosz, 1977).

Thus, when the decision is made to attach U_6 at the node just under DP, as shown in figure 3.4, the focus space associated with DSP_2 (U_1, U_4, U_5) is popped off the stack and no longer accessible. The focus space associated with the highest level DP is at the top of the stack, and the discourse entities realized by U_6 are part of this focus space. Although U_6 is adjacent to U_5, the referents and propositions realized by U_5 and stored in the focus space for DSP_2 are not accessible when processing U_6.

Similarly, when U_4 is attached under U_1, or more properly the intention recognized from U_4 is inferred to be dominated by that realized by U_1, the focus space on the top of the stack for DSP_1 (U_2, U_3) is popped from the stack. The focus space for DSP_2 (U_1) is at the top of the stack, and all the discourse entities associated with that focus space are more salient than those that were just talked about in U_3, by virtue of being on top of the stack.

As figure 3.5 shows, an utterance U_8 can be added anywhere along the right frontier of the tree-structured intentional structure. As the hearer interprets U_8 and attempts to infer the intention associated with it, the hearer also must determine where in the discourse structure this intention belongs. While it seems that hearers may not have to determine this at once, i.e. before they hear the next utterance, according to a hierarchical model of discourse structure, they must make this determination fairly quickly or the sequence of utterances won't make sense (but see (McKoon and Ratcliff, 1992)). In addition, because utterances may realize multiple intentions, it is possible that the same utterance fits in multiple locations in the intentional representation. Furthermore, since it is possible that Struct-$\mathcal{P}$ and Struct-$\mathcal{D}$ are not identical, an utterance may be in one relation to propositions in Struct-$\mathcal{P}$ and in another relation to referents in Struct-$\mathcal{D}$.

3.6.1.2 Approximating the Stack Model with AWM

Because of the strong effect of recency in AWM, the closest approximation to the stack model of attention is that certain utterances and the intentions inferred from them function as retrieval cues for retrieving entities stored in the discourse model to use in the current context. Thus a return to a prior context, such as that illustrated in figure 3.4 by the relation between U_4 and U_1, could be achieved if U_4 triggers a retrieval from memory that selects propositions and entities realized by U_1. This is encapsulated in the retrieval cue hypothesis given below:

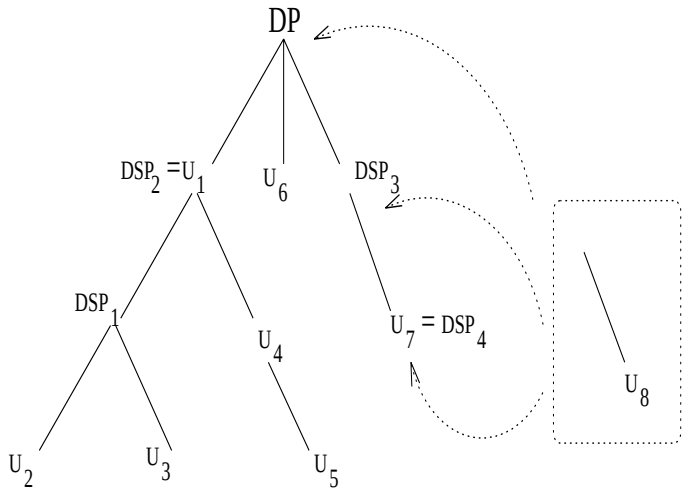


Figure 3.5: Adding a new utterance (utterance intention) to the Hierarchical Discourse Segment structure: 3 choices

(31) RETRIEVAL CUE HYPOTHESIS: the main function of an IRU is as a retrieval cue. It points back to a previous context and serves as the cue for the retrieval of a set of propositions from that context.

The RETRIEVAL CUE HYPOTHESIS is formulated for IRUs, but any type of utterance or the intention directly inferred from it could function as a retrieval cue. Formulating the POPPING operation as retrieval would mean that intentions can be strongly interrelated and directly affect attentional state without losing the effects of recency and frequency attested to by many psychology experiments on attention and working memory.

The AWM model as currently formulated provides no analog of a second aspect of the POPPING operation, i.e. that entities become no longer accessible. This might be emergent from the first aspect, that retrieval would displace other previously salient entities. In addition, a plausible analog might be that when conversing about a particular intention or topic, agents continually use certain propositions so these get continually restored in AWM. If it is clear that a proposition will no longer be needed, this 'rehearsal' process stops and the proposition quickly becomes no longer salient. An extension of the current AWM mechanism to model POPPING in this way, along with evaluating whether it would have the desired effect, is beyond the scope of this work.

3.6.1.3 The Stack Model doesn't Limit Attention

One advantage of AWM is that there are utterances that don't fit very well with a stack model of attentional state. For example, consider the two IRUs in 32-22 in the excerpt in 32. Here E has been telling H about how all her money is invested, and finally poses a question in 32-3:

(32) (3) E: .....
And I was wondering – should I continue on with the certificates or
(4) H: Well it's difficult to tell because we're so far away from any of them – but I would suggest this – if *all of these are 6 month certificates and I presume they are*
(5) E: *Yes*
(6) H: *Then I would like to see you start spreading some of that money around*
(7) E: uh hu
(8) H: Now in addition, how old are you?
.
(discussion and advice about starting an IRA)
.
(21) E: uh huh and
(22a) H: But as far as the certificates are concerned,
(22b) I'D LIKE THEM SPREAD OUT A LITTLE BIT -
(22c) THEY'RE ALL 6 MONTH CERTIFICATES
(23) E: Yes
(24) H: And I don't like putting all my eggs in one basket - and I would suspect that February 25 would be a good time to put it into something that runs for 2 and a half years. That first one that comes due. Call me on the others as they come due .....

Note that all of 22a,b,c are part of H's turn. Here, the relation of the IRUs in 32-22 to the structure of the previous discourse is similar to that of utterance U_4 in figure 3.5, which is related to the intention (topic) of U_1. In figure 3.5, the intervening material that was discussed was in U_2 and U_3; in excerpt 32 the intervening material started with 32-8 and extended to 32-22.

In G&S's stack mechanism, the utterance of *but as far as the certificates are concerned* could have the effect that intervening focus spaces would be popped so that the focus space representation of the previous segment from 32-3 to 8 would be on the top of the stack after 32-22a (Grosz, 1977; Sidner, 1979). It would seem that the propositions realized by the IRUs should be salient since they are on the top of the stack. But then it is difficult to see why they would be said again in the current segment.

However, if, as in AWM, the propositions from the previous segment must be retrieved, then the IRUs can function to save the hearer the retrieval time. Furthermore, if all the propositions from the previous segment aren't relevant in the current segment, then the IRUs select only what is relevant.

Example 32 shows that the stack model makes some nonintuitive predictions about how attentional state changes as a result of embedded segments. In addition, consider 33:

(33) (4) C: Ok harry, I'm have a problem that uh my - with today's economy my daughter is working.
(5) H: I missed your name.
(6) C: Hank.
(7) H: Go ahead hank
(8) C: as well as her uh husband
They have a child.
and they bring the child to us every day for babysitting.
.....

H interrupts C's narrative at 33-5, but in 33-8 C continues as though 33-4 had just been said. 33-8 contains both an anaphoric referent and an anaphoric property, and realizes the proposition *My daughter's husband is working as well.* The interpretation of the proforms in 33-8 depends on the recoverability of 33-4. In a hierarchical model of discourse this is explained by positing that 33-5 ... 33-7 is an embedded segment and presumably then the utterance of 33-7 closes the embedded segment and pops it off the stack. When 33-8 is interpreted, the focus space with the representation of 33-4 is at the top of the stack and supports the interpretation of the proforms.

We might expect then that continuations of this kind are possible over any embedded segment. However consider the following invented variation on 33:

(33') (4) C: Ok Harry, I have a problem that uh my - with today's economy my daughter is working.
(5) H: I missed your name.
(6) C: Hank.
H: Haven't you called me before?
C: Yes I called you last week.
H: Okay, that's what I thought.
(7') H: Well, Go ahead hank
(8) C: As well as her uh husband
They have a child.
and they bring the child to us every day for babysitting.
.....

In the variation given as 33', 33-8 is either very hard or impossible to interpret. Whether the hearer can access the utterance before the embedded segment may depend on the number of intervening utterances, the number of intervening topics, or some other factor. The invented variation in 33' is similar to the naturally occurring example in 34 which includes an IRU as 34-23:

(34) (13) E: Well however this is my question: I am a single woman, and retired. I have about 120 M, and I'll tell you how it's broken down, and perhaps you can advise me
(14) H: How old are you Elsa?
(15) E: I am seven six, alright?
(16) H: I got you.
(17) E: uh *70,000 in CDs*, 30,000 -
(18) H: When are they due?
(19) E: Pardon?
(20) H: When are they due?
(21) E: Well they're due right now, every month one is due until June
(22) H: They're due one a month?
(23) E: Yes sir.
NOW THAT'S 70.
Now another 30 in low-income CDs at 8% – the long term ones, you know?
(24) H: When are they due? .....

After an interruption by H, starting at 34-18, E, in 34-23, tries to continue her enumeration of how her money is invested. The IRU in 34-23 summarizes the extent of the investments that E had accounted for so far at the time of the interruption. The IRU is marked with a cue word *now* that indicates the beginning of a new segment in G&S's theory, and the following utterance also begins a new segment. Here E apparently did not feel as though she could continue her narrative without re-evoking propositions that had been said before the interruption.

The stack model makes no predictions about when we might expect propositions to get re-evoked. In contrast, AWM is too simple: it uses only a recency criterion. For the purpose of exploring IRUs, AWM produces the desired results by predicting that certain limits on attention capacity will have the effect that the speaker may want to produce an IRU to save the hearer retrieval, or in the worst case because otherwise the hearer may not retrieve the relevant proposition at all. The retrieval cue hypothesis as a way of formulating hierarchical structure in a model like AWM will be discussed further in chapter 7.

3.6.2 Propositional Relations and the Discourse Inference Constraint

In Hobbs', and Mann and Thompson's, theories of discourse structure, each utterance must be related to the prior utterance by a 'coherence relation' or 'rhetorical relation' (Hobbs, 1979; Mann and Thompson, 1987). These relations can hold between the content of utterances or between the utterances themselves, e.g. the speaker's right to make a particular assertion at this point in the discourse. In Grosz and Sidner's theory, each utterance must be related to the prior utterances on the intentional level by either the GENERATE relation that holds between actions or the SUPPORT relation that holds between beliefs.

None of these accounts have noted that IRUs are frequent in contexts where the speaker intends the hearer to infer that a particular relation holds between two propositions. However, much of this previous work has claimed that certain critical inferences in discourse rely on the adjacency of discourse segments (Hobbs, 1979; Cohen, 1987; Mann and Thompson, 1987; Polanyi, 1987). While I leave open the question of how these relations are inferred, the DISCOURSE INFERENCE CONSTRAINT is similar to the segment adjacency constraint.

It seems that it would be difficult to distinguish between these two accounts since what is adjacent is certainly salient. However, the accounts do make different predictions because adjacent discourse segments are defined on the hierarchical structure of a discourse. For example in figure 3.4, U_1 and U_6 are adjacent. This means that two utterances that are not be sequentially adjacent they may be adjacent at some level in the hierarchical structure, and the inference of 'coherence relations' could occur at this level. The AWM model used here would not allow these inferences to arise unless the propositions realized by the utterances themselves were salient at the same time. In addition, the inferences that arise from IRUs in the corpus are all cases of immediate juxtaposition and don't depend on hierarchical structure. Furthermore, recent experiments by McKoon and Ratcliff also suggest that inferences between distal parts of a text are not automatically derived (McKoon and Ratcliff, 1992).

3.7 Summary

Chapter 3 provides the theoretical framework used throughout the rest of the thesis. Section 3.2 introduces a model of limited Attention called AWM (Attention/Working Memory). The purpose of AWM is to model the posited resource bound of attention, and thus provide the functional basis for Attention IRUs.

Because AWM is so simple, the question arises as to whether it can model Attentional State in

discourse. Most discourse theories assume that discourse is hierarchically structured and Grosz and Sidner have proposed a model of Attentional State in discourse that is parasitic on this hierarchical structure. Section 3.6.1 compares the AWM model with G&S's stack model and suggests a simple reformulation of it in terms of AWM so that it is more psychologically plausible.

Section 3.3 briefly reviews accounts of belief deliberation and explains the theory underlying the Attitudes of ACCEPTANCE and REJECTION. The theory of belief deliberation is dependent on the model of limited attention because only salient, explicit beliefs are subject to belief revision and to reasoning in general. Following Galliers, I propose that the corrigibity of a belief is dependent on the ENDORSEMENTS on the belief and it supporting beliefs. I introduce a set of endorsement types needed to explain the types of beliefs that occur in the corpus and show how these can be used to predict preferred belief states.

These endorsement types are then used as part of the explanation of how agents infer whether another agent accepts a belief. The basis for the account of these inferences of acceptance is Lewis's theory of common knowledge, presented in section 3.4 as a theory of mutual supposition. The different endorsement types reflect the fact that inferences about what other agents accept are defeasible, and can be more or less supported by evidence in the discourse situation (Prince, 1978; Nadathur and Joshi, 1983). The account of mutual supposition will be used in chapter 6 as part of a theory of the function of Attitude IRUs.

Chapter 4

Empirical Method: Distributional Analysis

4.1 Introduction

In chapter 1, I proposed that IRUs have three general functions:

- Attitude: to provide evidence supporting beliefs about mutual understanding and acceptance
- Attention: to manipulate the locus of attention of the discourse participants by making a proposition salient
- Consequence: to augment the evidence supporting beliefs that certain inferences are licensed

Whether an IRU has one of these functions in a naturally occurring discourse is not apparent from simple observation supported by introspection; thus it seems clear that any theory about the function of IRUs must be supported by some empirical evidence. The theory proposed here relies on two methods: (1) corpus-based distributional analysis of IRUs, and (2) computational modeling to formalize and test the theory of the function of an IRU in a particular context. In this chapter, I will discuss the distributional analysis. Section 4.2 discusses the parameters of the distributional analysis and shows how each parameter can be used to argue for one IRU function or another.

Corpus-based distributional analysis is a method of the functional school of pragmatics (Kuno, 1987; Ward, 1985; Prince, 1986; Birner, 1992). A prototypical study using this method consists of (1) selecting a particular identifiable form; (2) collecting a large number of tokens of this form, along with the discourse context in which each token occurs; and (3) justifying, using multi-variate

analysis, which variables of the discourse context constrain the relation between the distribution and discourse function of the designated form.

The benefits of this method are that we get an overall picture of the distribution and discourse function of a form in many different discourse contexts. This is important to provide a general empirically grounded theory. In addition, the analysis provides a set of predictions about similar forms that we might expect to have the same or related discourse functions. Furthermore, the criteria for selecting the data set and the variables used in coding the discourse context are explicit and thus replicable. This means that the predictions and claims of the account provided can easily be tested and extended to other corpora representing other speech situations. This seems preferable to basing a theory on a few examples or one speaker's intuitions.

4.2 Distributional Analysis: The Function of IRUs

CLASS	SUBCLASS	CODING FACTOR VALUE
INFORMATION STATUS	SALIENCE	ADJACENT, SAME, LAST, REMOTE
	HEARER OLD	REPETITION, PARAPHRASE, ENTAILMENT, IMPLICATURE, PRESUPPOSITION , UNUSED
	SPEAKER	SELF , OTHER
PROSODIC REALIZATION	PHRASE FINAL INTONATION	LOW, MID, HIGH
DISCOURSE CORRELATES	CUE WORDS	SO , THEN, NOW, BUT , AND, OK , WELL
	SALIENT SET	YES/NO
	ARGUMENT OPPOSITION	YES/NO
	SEGMENT LOCATION	OPEN SEGMENT, CLOSE SEGMENT

Figure 4.1: Summary of Parameters used in Distributional Analysis

A distributional analysis uses distributional parameters to identify certain subclasses of IRUs. Then these distributional subclasses are claimed to realize particular discourse functions. The discourse functions are supported by the correlation of other distributional parameters that were not used to identify the class in the first place.

The data set for the distributional analysis consists mainly of a corpus of a radio talk show for

financial advice consisting of 55 dialogues collected over a period of a week of hourly shows.[1] 210 IRUs were analyzed in the distributional analysis.

The advantages of a talk show genre are that: (1) the conversants typically do not previously know one another, so the analyst has access to most of the information relevant for the discussion since it must be made explicit in the conversation; (2) the conversations are directed toward an explicit purpose which provides some constraints on what is relevant. An additional argument for this particular talk show is that its longevity indicates that the callers and audience have found it useful. A possible disadvantage is that because the conversation is not just between the talk show host and the caller, but includes the audience as well, there may be aspects of the conversation that are specific to this genre which are not being accounted for by an analysis of it as a two person conversation. Another source of tokens in this research are examples in the literature and opportunistically collected tokens.

The application of the functional method depends first on defining a way to select a form and then determining which contextual variables might possibly have a bearing on the distribution and function of this form. The definition of IRU repeated here from chapter 1 delimits the data set that the analysis is based on:[2]

> An utterance u_i is INFORMATIONALLY REDUNDANT in a discourse situation $\mathcal{S}$:
>
> 1. if u_i expresses a proposition p_i, and another utterance u_j that entails p_i has already been said in $\mathcal{S}$.
> 2. if u_i expresses a proposition p_i, and another utterance u_j that presupposes or implicates p_i has already been said in $\mathcal{S}$

Remember that, as discussed in chapter 1, IRUs only describe a subset of redundant propositions in discourse. Also that the utterance, u_j, which originally added the propositional content of the IRU to the discourse situation, is called the IRU's ANTECEDENT. A diagnostic of defeasibility can be used to test whether the proposition p_i expressed by an utterance u_i is already entailed. Section 2.2 discussed how presuppositions and implicatures are often linked to lexical items. In order to determine whether the proposition p_i is presupposed or implicated, the analyst must rely on standard accounts of presupposition and implicature such as those given in (Horn, 1972; Karttunen

[1] This corpus was originally taped from a live radio broadcast and transcribed by Julia Hirschberg and Martha Pollack. I am grateful to Julia Hirschberg for providing me with the tapes of the original broadcast. This allowed me to segment, pitchtrack and retranscribe the IRUs that this analysis is based on, using programs generously supplied by Mark Liberman.

[2] While it is possible that other information will be shared between a speaker and hearer, the definition below typically must be used in corpus analysis since observers cannot reliably determine when both speakers in a conversation have access to this other information.

and Peters, 1979; Gazdar, 1979; Hirschberg, 1985), where the process of how the implicature or presupposition is derived has been described in detail.

The second step of the distributional analysis is to select a number of contextual variables which can be reliably coded for and which support the analysis of the function of IRUs. There are three classes of distributional parameters: (1) information status, (2) utterance intention and (3) discourse correlates. Since one of the functions of IRUs is to manipulate ATTENTION by making a proposition salient that isn't currently salient, one goal is to find a way to determine the salience of the antecedent of the IRU. Since the function of CONSEQUENCE IRUs is to increase the evidence supporting mutual beliefs about what is inferrable, it is important to know the logical basis for the belief in the antecedent.

These factors have historical precedent in factors considered relevant to the determination of the function of other forms in discourse. It is well known that discourse entities, including both referents and propositions, vary in terms of the degree to which they are HEARER OLD (Prince, 1981b; Clark and Marshall, 1981; Horn, 1986; Moser, 1992), and this variation determines both syntactic form and prosodic realization.

One cue is to utterance intention is whether the IRU is prosodically realized with a phrase final Low, Mid or High tone. Prosodic realization also shows whether the speaker is treating a particular item of information as HEARER OLD.

A final set of parameters are the discourse correlates of IRUs. These discourse correlates are indicators of discourse structure such as cue words, the occurrence of salient sets, and conflicting default inferences, all of which can support arguments for both ATTENTION and CONSEQUENCE functions (Schiffrin, 1987; Horn, 1991; Prince, 1981a).

The overall structure of the distributional analysis and the way it is organized is shown in figure 4.1. The meaning of each coding factor will be discussed in detail in the remainder of this section. Factors related to INFORMATION STATUS and are discussed in sections 4.3 and utterance intention is discussed in section 4.4.1. Discourse correlates are discussed in section 4.5.

4.3 Information Status Parameters

Variation in information status can be classified along two independent dimensions: logical status and salience. These dimensions reflect Prince's distinction between HEARER OLD information and SALIENT information as discussed in section 1.2 (Prince, 1981b; Prince, 1992). First, I discuss types of HEARER OLD information, and then SALIENT information.

4.3.1 Hearer Old Information

On the HEARER OLD dimension, the first point to note is that all entailments are not created equal. There is a primary distinction between what has been said and what is inferrable from what has been said. This is reflected both in the distinction between EVOKED entities and INFERRABLE entities in Prince's taxonomy of HEARER OLD and in Levesque's distinction between EXPLICIT and IMPLICIT beliefs (Levesque, 1984). Here I make a number of finer distinctions relevant to propositions and their logical status in the discourse. There are five categories of HEARER OLD status for propositions based on the relation of the IRU to its ANTECEDENT:

- Repetition: the IRU is a partial or complete repetition of its antecedent, except for potential mappings of indexicals such as *I* to *you*.
- Paraphrase: the IRU is a syntactic or semantic paraphrase of its antecedent, resulting from the application of simple syntactic transformations such as topicalization or semantic mappings based on lexical semantics (McKeown, 1983; Melčuk, 1988).
- Entailment: the IRU follows from a logical inference rule such as modus ponens from a set of antecedents.[3]
- Implicature: the IRU is an implicature from a set of antecedents, based on lexical items or sets, as described in (Horn, 1972; Gazdar, 1979; Hirschberg, 1985).
- Presupposition: the IRU is presupposed by its antecedent, based on existential or factive presuppositional forms (Kiparsky and Kiparsky, 1970; Karttunen, 1973; Gazdar, 1979; Bridge, 1991).

The distinction between Repetitions and Paraphrases and other Entailments reflects the distinction between EXPLICIT and IMPLICIT beliefs and between EVOKED and INFERRABLE discourse referents (Prince, 1981b). Of course, both Repetitions and Paraphrases are entailed by their antecedents, but Repetitions are **explicitly** related to their antecedents. Paraphrases rely on semantic relations between propositions: these relations are more explicit than Entailments because they can usually be derived by substitution of lexical definitions or by simple syntactic transformations (Joshi, 1964; McKeown, 1983; Melčuk, 1988).

Like Entailments, Presuppositions and Implicatures are **implicit** in what has been said. Thus saying them can have a function of making them **explicit**, i.e. supported by a LINGUISTIC endorsement.

[3]While any set of propositions can be thought of as a single proposition by just conjoining the set members, I prefer to talk about sets of antecedents since they may not all be introduced into the discourse at the same time, and producing the conjunction may have processing consequences (Vardi, 1989).

A final category of HEARER OLD information is what was called UNUSED information in Prince (1981b). This is information which is old to the hearer but not old to the discourse. The examples of this type of IRU that I have are ones that I heard in conversation, where I could use my own knowledge of the conversants to determine that the information in the IRU was already mutually believed.

Examples of these different categories will be presented after the next section.

4.3.2 Salient Information

Discourse salience is coded with a textual basis, consisting of parameters based on the **location** of the IRU's ANTECEDENT. This is an approximation since salience is supposed to reflect what is in the hearer's consciousness (Chafe, 1976), but like all categories based on mindreading, approximations are necessary. There are four categories of SALIENCE relations between an IRU and its antecedent:

- Adjacent: the IRU sequentially follows its antecedent utterance, ie. the IRU is U_{n+1} and its antecedent is U_n. Speakers don't normally say the same thing twice in sequence so this usually only happens when the antecedent was just said by the Other speaker.
- Same: the IRU is in the same turn as its antecedent, and therefore said by the current speaker. However, it is normally not Adjacent.
- Last: the antecedent is in the last turn of the current speaker, i.e. there is one intervening turn by another speaker.
- Remote: the antecedent is remote from the IRU, said by either speaker in a turn that was prior to the last turn of the current speaker.

The final relevant information status parameter is the Speaker of the antecedent of an IRU. Speaker information is coded because it is possible that a speaker's own utterances are more available than what someone else said. In addition, an IRU whose antecedent was produced by another speaker may serve a different function than one whose antecedent was produced by the speaker of the IRU. The speaker parameter has two values:

- Other: the speaker of the antecedent of the IRU is different than the speaker of the IRU.
- Self: the speaker of the antecedent of the IRU is the same person as the speaker of the IRU.

The speaker parameters and the location salience parameters are not orthogonal in one instance: Same location entails Self as Speaker.

4.3.3 Examples of Information Status Parameters

Consider 35 and 36.

(35) H: That's right. as they come due, give me a call, about a week in advance. But the first one that's due the 25th, *let's put that into a two and a half year certificate*
E: PUT THAT IN A TWO AND A HALF YEAR. Would ...
H: Sure. we should get over 15 percent on that

(36) (15) H: Oh no. *I R A's were available as long as you are not a participant in an existing pension*
(16) J: Oh I see. Well I did work, *I do work for a company that has a pension*
(17) H: ahh. THEN YOU'RE NOT ELIGIBLE FOR EIGHTY ONE

In example 35, the HEARER OLD status is a Repetition: (e) produces a partial repetition of what H has just said. In example 36, the HEARER OLD status is an Entailment via the logical inference rule of MODUS TOLLENS. Since 36 is an Entailment, it can have a Consequence function of demonstrating that an inference was made or making sure that the other agent makes the inference. Because Repetitions are explicitly related to their antecedents, the function of the Repetition in 35 cannot be Consequence.

Because both 35 and 36 are Adjacent to their antecedents, the proposition they realize is already salient and they can't have the Attention function of making a proposition salient. The Speaker parameter for both 35 and 36 is Other: in 35 E uttered the IRU and H said the antecedent. Because they are both Adjacent and their antecedent was said by the Other speaker, they can be Attitude IRUs: they can demonstrate ACCEPTANCE of the previous speaker's assertion.

4.4 Utterance Intention Parameters

The distributional analysis relies on one prosodic cue to utterance intention: phrase final intonation.[4]

4.4.1 Phrase Final Prosodic Realization

IRUs occur with a phrase final Low, Mid, or High tone. Final High marks interrogative force, uncertainty, or new information(Pierrehumbert and Hirschberg, 1990; McLemore, 1992). Final

[4]See (Walker, 1993c) for a discussion of the other prosodic correlates of IRUs.

Mid marks continuation and HEARER OLD or predictable information(Liberman, 1975; Ladd, 1980). I will argue that final Mid is used to mark non-assertion. Final Low marks assertion or finality. Phrase final intonation indicates whether the speaker is treating the current utterance as a question, an assertion, or as HEARER OLD information.

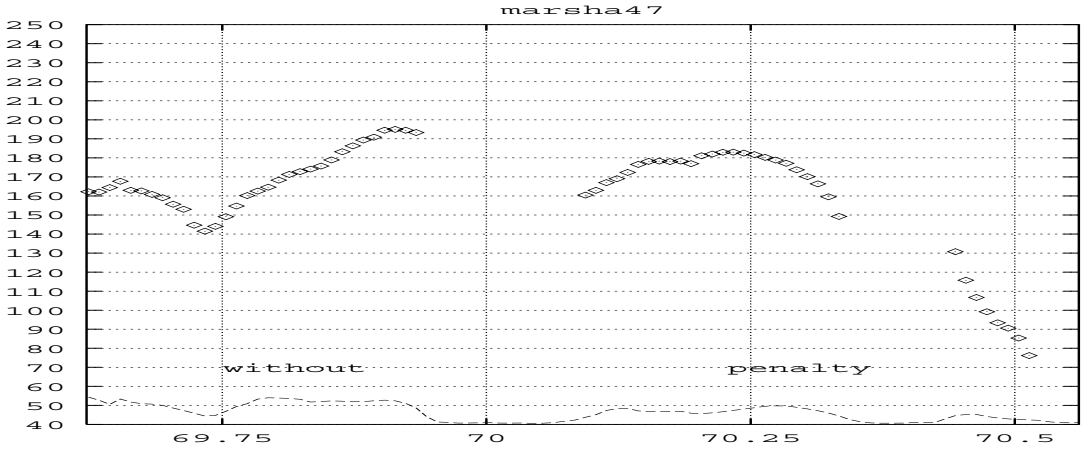


Figure 4.2: Fundamental frequency over time. Final tone is a phrase final Low

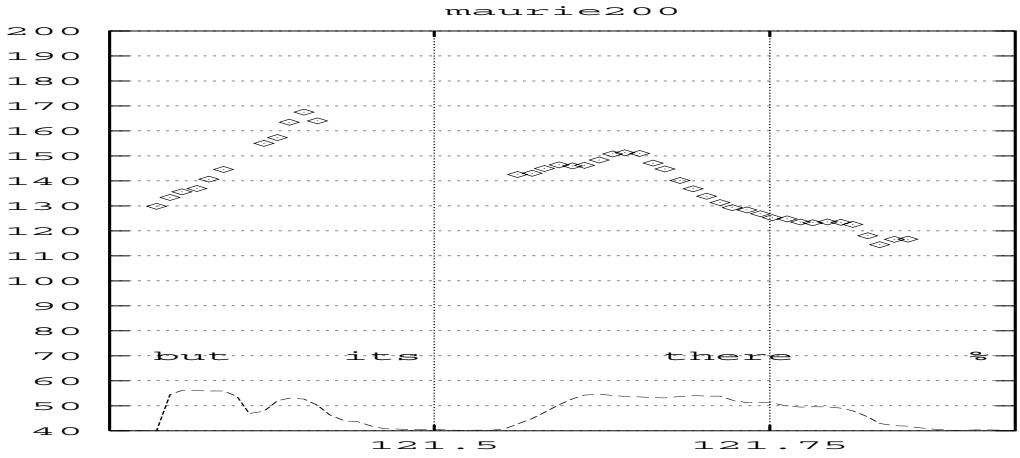


Figure 4.3: Fundamental frequency over time. Final tone is a phrase final Mid

An example of a phrase final low (fall) is given in Figure 4.2. Figure 4.3 shows a phrase final mid. Figure 4.4 shows a phrase final high. Low and Mid define two different falls, falls to Mid and falls to Low. Phrase final highs define rises.

Since all the tones are interpreted with relation to tones in context, there is no objective fundamental frequency (F0) which counts as a High, Mid or Low. Figure 4.5 shows however that if final tone is plotted as a function of the F0 of the previous High, we get three distinct distributions for the three

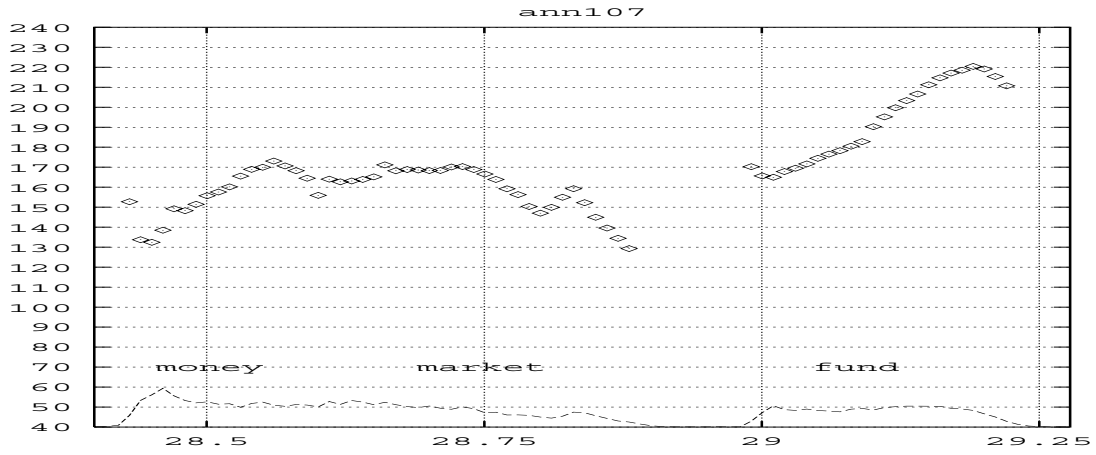


Figure 4.4: Fundamental frequency over time. Final tone is a phrase final High

tones. This argues that these tones can be reliably identified by conversants (and by analysts) (cf. (Liberman and McLemore, 1992; McLemore, 1992)).

4.4.2 Examples of Prosodic Realization Parameters

Consider again example 35 repeated here for convenience:

(37) H: That's right. as they come due, give me a call, about a week in advance. But the first one that's due the 25th, *let's put that into a two and a half year certificate*
E: PUT THAT IN A TWO AND A HALF YEAR. Would ...
H: Sure. we should get over 15 percent on that

Remember that the SALIENCE status of the IRU in 37 is Adjacent and its HEARER OLD status is a Repetition. As shown in figure 4.6, it was prosodically realized with a final Mid, which, as mentioned above, marks both continuation and HEARER OLD or predictable information. Since, as the transcript shows, E intended to continue before she was interrupted by H, there is no basis for concluding whether the Mid here is related to the fact that the proposition conveyed is both HEARER OLD and salient or E's intention to continue.

However, consider the difference between the phrase final realization of 4.2 and 4.7. The two contexts for the two utterances are below. Figure 4.2 is 38-26 and figure 4.7 is 38-27. The dialogue excerpt is given in 38:

(38) (24) H: that is correct, it could be moved around so that each of you have 2000
(25) M: I

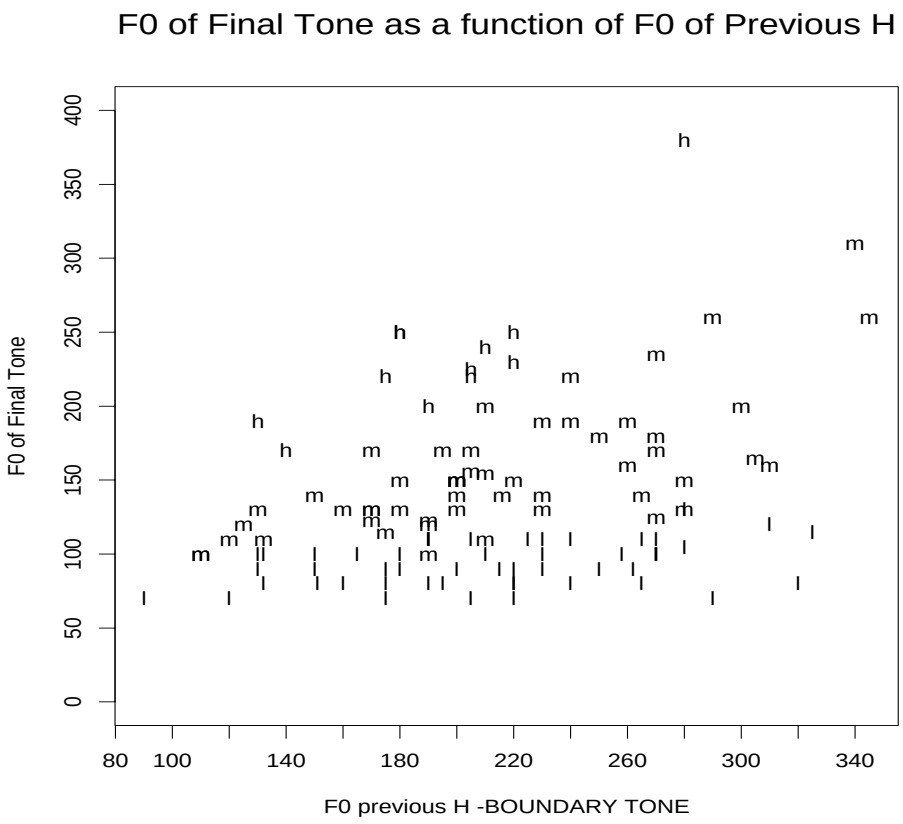


Figure 4.5: Distribution of Final Tones as a function of previous High

(26) H: *Without penalty*

(27) M: WITHOUT PENALTY

(28) H: Right

(29) M: And the fact that I have a an account of my own ...

Although it isn't clear from the excerpt above, both utterances are IRUs. 38-26 also had an antecedent in the dialogue shown below in 39-13 and 39-14.

(39) (13) M: Anyway what happens if in the future I should get a job where I could uh contribute to my own I R A? The local bank person indicated that that would mean *we*

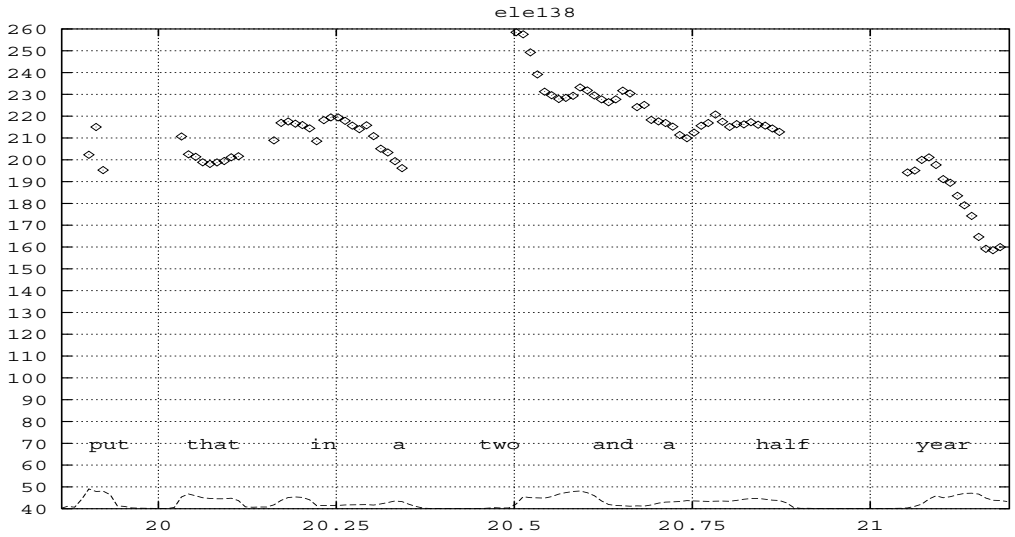


Figure 4.6: Eleanor 138. Phrase Final Low

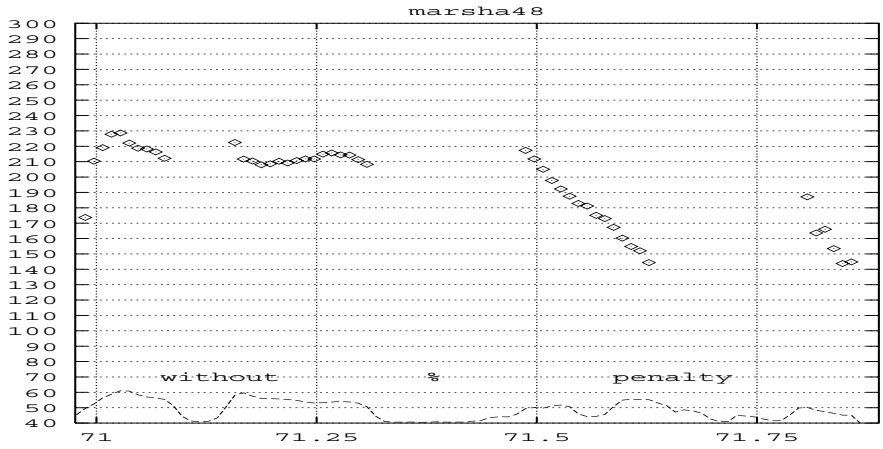


Figure 4.7: Marsha 27. Phrase Final Mid

would pay a penalty on anything that had ever been put in the spousal account.

(14) H: *nope not so ....*

As figure 4.2 shows, the phrase final tone for 38-26 is a Low. In contrast the phrase final tone for 38-27 is a Mid. In this case, it seems reasonable to argue that 27 is marked as HEARER OLD or non-assertive, while the speaker treats 26 as newly asserted information.

As a final case of final Mid, consider figure 4.3. The dialogue excerpt is given in 40:

(40) (25) M: So I just put that on the little part there on that new form there
(26) H: Well you'll have to list that as interest earned, and then you can knock that amount off, somewhere on schedule b,
there's a line for it, I don't remember what the line number is, but IT'S THERE

The phrase final Mid on the *its there* in 40 shows that H is perfectly cognizant of the fact that the proposition is currently salient, predictable information. Thus it would be hard to support an analysis of the IRU in 40 as an Attention IRU: a speaker cannot both intend to make a proposition salient and believe that it is currently salient. Note the contrast between HEARER OLD and SALIENCE here in that the IRU shown in 4.2 has a phrase final Low.

As a final example of the use of phrase final tones, consider the utterance shown in figure 4.4 which is 41-33 in the excerpt below.

(41) (32) A: I have uh some money accumulated, I really don't feel as though I'm using it to its best advantage. I have 80 thousand dollars in CDs, I have 20 thousand in an Allsavers, 10 thousand in a money market fund
(33) H: Money market fund?
(34) A: Yes.
(35) H: Right.
(36) A: And 5000 in stocks .....

The final rise of 41-33 can be used to argue that 41-33 has interrogative force. What H is apparently questioning here is whether he heard correctly.

4.5 Discourse Correlate Parameters

One discourse correlate is provided by discourse markers. Discourse markers are claimed to both indicate discourse structure and indicate relations between propositions in discourse (Polanyi and Scha, 1984; Reichman, 1985; Schiffrin, 1987; Grosz and Sidner, 1986; Hirschberg and Litman, 1987). Discourse markers such as *so* and *then* mark propositions as inferrable in the context, and thus provide supporting evidence for the function of CONSEQUENCE IRUs. Discourse markers such as *now, but* and *ok* mark the beginning of a new segment of discourse and thus may correlate with ATTENTION IRUs.

Another discourse correlate is intended to track when IRUs occur in 'contrastive' environments

(Ward, 1985; Prince, 1986; Horn, 1991; Ward, 1990; Walker, 1993b). The inference of rhetorical contrast is licensed in two situations: (1) a salient set is evoked in the dialogue (Prince, 1986); or (2) there is opposition in the argument structure. An example of (2) is provided in example 40. Salient Set and Argument Opposition are binary coding parameters with values of yes/no in order to track both whether an IRU occurs in an environment of potential contrast, and what the basis for the contrast is. Typically these 'contrastive' IRUs also cooccur with the discourse marker *but* and are analyzed here as a type of CONSEQUENCE IRU, as will be explained more fully in chapter 8.

The final discourse correlate is related to the intuition that the function of some IRUs is to manipulate attentional state or indicate discourse structure. Whittaker and Stention analyzed discourse structure in problem-solving dialogues and noted that IRUs tend to occur at the end of discourse segments. Based on the analysis of the financial advice dialogues, it appears that IRUs occur both as part of openings and as part of closings of discourse segments.

The problem with exploiting this intuition is that segmenting discourses is difficult and its operationalization is an open problem. In studies in which judges were asked to segment discourses, only a small number of the discourse segment boundaries were agreed upon by all judges (Whittaker and Stenton, 1988; Passonneau and Litman, 1993). The most successful examples are those in which the discourse structure closely parallels a well defined task structure (Grosz, 1977; Sibun, 1991). The difficulties include determining when one segment ends and another begins, defining the level of granularity of a segment, and with the perception that some utterances simultaneously belong to two disjoint segments.

I have not attempted to completely segment the dialogues in the corpus, but have constructed some operational criterion for judging whether an IRU is at the opening or closing of a discourse segment. Many of these critera are based on observations made in the literature, but they may not be fully general. An IRU U is an Open Segment IRU if:

- U is part of a turn that starts with a cue word use of *now*
- U is part of a turn that starts with a topic marking construction such as *as far as the Z*, *on the subject of Y*, presentational there sentences, etc.
- U and U_{-1} have no discourse entities that are related by coreference, inferential dependency, or poset relations.

An IRU U is a Close Segment IRU if:

- The dialogue ends after U, or any utterances after U and before the end of the dialogue are like *Okay, Thanks Harry, You're great* or the rest of the dialogue consists of IRUs and the above.
- U_{+1} starts with cue word use of *now*
- U_{+1} consists of an utterance such as *Yeah I figured* or *Okay, that was my question*, followed by an utterance that starts with a cue word use of *now*, or a topic marker like *as far as*
- U_{+1} is like *Okay, second question* or *as far as the certificates are concerned*, or *Another alternative would be to X* or a presentational-there sentence with a new discourse entity.
- U_{+1} represents a shift in initiative, ie. U ends a sequence of utterances (ie at least two or three turns), by one speaker and the next sequence of utterances is by another speaker. The cases which occurred in the corpus all correspond to the end of problem/data description by the caller (or client or apprentice) which is typically followed by an initiative shift in which the talk show host (expert) takes over initiative for the rest of the dialogue, as described in (Whittaker and Stenton, 1988).
- U ends an embedded narrative, e.g. a story/anecdote set in the past and temporally distinct from what is under discussion now.

The following section provides examples of IRUs coded with the discourse correlates.

4.5.1 Examples with Discourse Correlate Parameters

Consider the example given in 42:

(42) J: *I also have ten thousand in a C D*, six month C D
H: Oh, uh hang on, do you have anything else?
J: That that's it. that's it.
H: Well now we're talking about something slightly different.
YOU ALREADY HAVE SOME MONEY IN A C D.
Have you anyone dependent on you?

The arguments that this is an Attention IRU include the fact that SALIENCE status of the IRU is Remote. In addition, the discourse marker *now* at the beginning of the turn classifies the IRU as an Open Segment IRU. Furthermore, the IRU is prosodically realized with a phrase final Low, so it is not explicitly marked as HEARER OLD or predictable information.

As an example of Argument Opposition, consider again the example in 43 repeated for convenience from section 4.4.2.

(43) (25) M: So I just put that on the little part there on that new form there
(26) H: Well you'll have to list that as interest earned, and then you can knock that amount off, somewhere on schedule b,
there's a line for it, I don't remember what the line number is, but IT'S THERE

This example is a case of Argument Opposition because the assertion of *I don't remember what the line number is* could support the inference that there isn't a line number after all. The assertion that *it's there* argues that there is a line number. The discourse marker *but* prefaces the IRU and provides a further correlate for Argument Opposition (Horn, 1991).

As an example of Salient-Set consider the excerpt given in 44:

(44) (19) H: well, the medical and dental care you can deduct, provided you can establish that you have provided more than half support.
(20) R: uh huh
(21) H: BUT THE DEPENDENCY YOU CANNOT CLAIM
(22) R: um hm (breath) I see. ok. uhh, alright, the second question...

The complete example showing the antecedent for the IRU given in 44-21 will be given in chapter 7. Here the IRU would be coded as having a salient set because *the dependency* is being compared with *the medical and dental care*: both are ways to get tax deductions. Selection from salient sets provides another way to make contrasts, and as noted by Ward and Horn (Ward, 1985; Ward, 1990; Horn, 1991) and discussed in section 8.2, IRUs are prevalent in environments of contrast. This utterance also ends a discourse segment, as shown by the 44-22, where (r) goes on to a new question.

4.6 Summary of Corpus via Distributional Parameters

Figure 4.6 shows how the corpus is distributed by the HEARER OLD and SALIENCE coding factors. These two factors are used to define Attention IRUs (the row with Remote) and a subclass of Consequence IRUs called Inference-Explicit IRUs (the column with Infer). The Remote parameter defines Attention IRUs, based on the assumption that an IRU can be used to make a proposition salient if and only if that proposition is not currently salient.

	Repeat	Para	Infer	Presupp	Unused
Remote	6	39	8	0	4
Last	7	18	4	1	0
Same	0	5	3	5	0
Adjacent	54	20	17	9	0

Figure 4.8: All IRUs by Hearer Old Status and Salience Status

	Repeat	Para	Infer	Presupp	Unused
Remote	0	3	1	0	4
Last	0	0	0	0	0
Same	0	3	0	4	0
Adjacent-SELF	0	0	0	8	0

Figure 4.9: Contrastive IRUs: Argument Opposition

The HEARER OLD Inference factor is the combination of Entailment and Implicature. The Inference category defines a class of IRUs which make inferences explicit, which is one way of achieving the Consequence function of supporting beliefs that particular inferences are licensed.

The other subclass of Consequence IRUs, Affirmation IRUs are not definable by the HEARER OLD and SALIENCE coding factors. Affirmation IRUs are defined by the Salient-Set and Argument Opposition discourse correlates and by the use of the particle *but* to preface the IRU. The distribution of these with respect to the total number of IRUs shown in figure 4.6 is given in figure 4.6.

The class of Attitude IRUs is defined by a combination of SALIENCE and Speaker coding factors. Figure 4.6 shows how the Adjacent row in figure 4.6 is distributed between Adjacent/Other and Adjacent/Self. Chapter 6 will argue that IRUs that are Adjacent to their antecedents where the antecedent was said by the Other speaker are used to demonstrate Attitude.

These categorizations on the basis of purely distributional factors have two effects: (1) functional categories may be overlapping, and (2) some IRUs are left out of the categorization. An example of

	Repeat	Para	Infer	Presupp	Unused
Adjacent/Other	54	14	15	1	0
Adjacent/Self	0	6	2	8	0

Figure 4.10: Adjacent IRUs by Speaker and Hearer Old Status

overlapping categories can be seen in figure 4.6 by the fact that 8 IRUs are in both the Remote row and the Inference column. Whether these IRUs function as Attention IRUs or Inference-Explicit IRUs or both is discussed in chapter 8. Similarly figure 4.6 shows that there are 15 Attitude IRUs that are also Inference-Explicit IRUs. I argue that these IRUs function both for Attitude and Consequence in chapters 6 and 8.

The IRUs that are left out of the categorization are 25 IRUs in the Last SALIENCE category. Of these, 7 are repetitions and 18 are Paraphrases. These IRUs are reattempts by the speaker to achieve the intention of the antecedent of the IRU. Either the antecedent wasn't heard so the speaker repeats it, or the speaker paraphrases it because it wasn't understood or as a way to elaborate on what was said. The 8 Adjacent/Self IRUs shown in figure 4.6, which are not Affirmation IRUs, are also used to elaborate or summarize what the speaker was just saying. These are all cases where the speaker continues because they are not sure that they have communicated what s/he intended. Being unsure may be related to the lack of a response of an expected type from the hearer or due to some other reason. Thus these may not actually be redundant and are not discussed in the remainder of the thesis.

4.7 Summary: Distributional Analysis

This chapter discussed distributional analysis as one of two complementary empirical methods used in this thesis to support the theory of the function of IRUs. Distributional analysis is a good method for determining the function of IRUs by examining their occurrence in multiple contexts. Section 4.2 discusses the parameters used in the distributional analysis. Each parameter can be used to argue for a specific communicative function of IRUs as the following chapters will show.

While the distributional analysis of IRUs provides a primary empirical basis for the theory presented here, it should be obvious that there are certain claims that are difficult to support by a distributional analysis, namely those about the relationship between IRUs and limited processing. This is

because: (1) there are no performance measures for the dialogues in the financial advice corpus. It is also impossible to objectively assess the impact of the IRU on the communication, because we would need control dialogues for similar tasks in which IRUs were **not** presented. Furthermore, such comparisons are not possible because the participants in each financial advice dialogue were attempting to solve different financial problems;[5] (2) It is impossible to distinguish potential conventional uses of IRUs from ones that result from cognitive limitations such as limited attentional capacity or limited inferential capacity; (3) It is impossible to tell what inferences were made and how long participants retain discourse information in memory, because we have no access to the mental states of the participants as they are carrying out the dialogue.

Because of these limitations of distributional analysis, computational simulation is a complementary method. Chapter 5 presents the Design-World experimental simulation environment which will provide another source of empirical evidence about the function of IRUs.

[5]One possible performance measure would be time or length of the dialogue, but this measure ignores the quality of the solution, as well as the fact that different problems take different amounts of time to solve.

Chapter 5

Empirical Method: Design-World

5.1 Introduction

Design-World is an artificial domain which consists of the floor plan for a house and a number of pieces of furniture. The task involves two agents who must carry out a dialogue to come to an agreement about how to arrange some of the pieces of furniture in the rooms on the floor plan. Each agent starts out with a set of pieces of furniture, which must be selected from to accomplish the task. Each piece of furniture has an associated point value which contributes to a performance measure for the task, and the agents attempt to maximize the points achieved by their design. The task, while artificial, is based on cooperative design tasks used for experiments on distributed cooperative work (Suchman, 1985; Bly, 1988; Tang and Leifer, 1988; Whittaker, Geelhoed, and Robinson, 1993).

The following section sketches the goals of Design-World simulations. Section 5.3 introduces the domain and the task for the simulation. Section 5.4 presents the IRMA architecture for resource bounded agents, also used in the TileWorld simulation environment (Bratman, Israel, and Pollack, 1988; Pollack and Ringuette, 1990). Section 5.6 discusses the particular instantiation of the IRMA architecture used in the simulation. Section 5.7 discusses the discourse actions that the agents can engage in. Section 5.8 discusses the way the domain and the task limits the range of discourse strategies available and discusses in detail which task parameters and cognitive parameters can be varied in order to explore the interaction of a cognitive limitation with a particular discourse strategy. The way in which performance is measured in order to determine the effects of the strategies will be discussed briefly in section 5.3, and then more fully in section 5.10.

5.2 Design-World Goal: Testing the processing effects of IRUs

Design-World is a highly parametrizable simulation environment based on the cognitive architecture presented in chapter 3. The simulation parameters are:

- Search radius of AWM which determines the extent to which an agent is attention limited.
- Number of inferences agents can make. This is done with three distinct parameter settings NONE, HALF, ALL
- Number of memory retrieval operations
- Discourse strategies for communication. The Baseline strategy will be discussed here. Each communicative function has its own associated strategies.
- Task Difficulty: the task can be varied to increase inferential complexity by introducing extra goals to match colors of pieces, or to require that agents never make a mistake, or to require that agents agree on the reasons for their actions.

The idea is to parametrize agents with communication strategies that do or do not include IRUs, and then test for a cognitive benefit for strategies that include IRUs, in terms of improved performance or decreased amount of inference or retrieval. Thus, Design-World provides empirical support three different aspects of the theory: (1) it provides a testbed for the shared environment model of mutual beliefs; (2) it supports the exploration of the interaction of limited attention and inference with communication strategies; and (3) it allows claims about the effects of IRUs to be tested in terms of global performance measures such as efficiency of task execution.

5.3 Design-World Domain and Task

The basic Design-World task consists of a pair of agents achieving a design for a floor plan with two rooms. Figure 5.1 shows a potential initial state. Furniture items are of 5 types: couch, table, chair, lamp and rug. Each furniture item has a color and point value. A design for a room consists of any four pieces from these types. The points associated with a furniture item supports the calculation of utility of including that item in the design plan and provides the basis for an objective performance measure for the Design-World task.

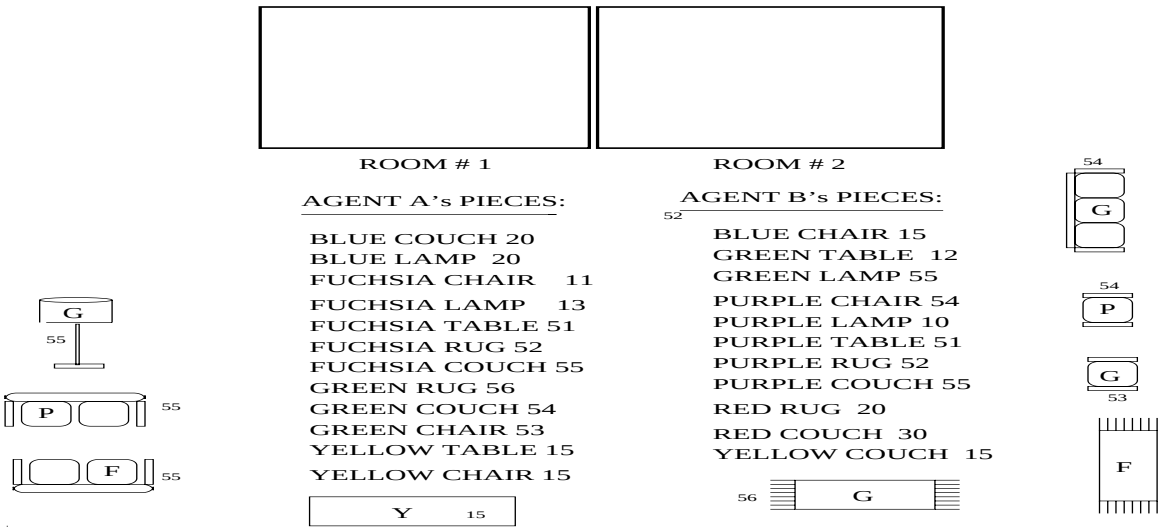


Figure 5.1: Initial State for the Design-World Task

The agents in Design-World communicate with an artificial language which will be described in more detail in the following sections. An example gloss of a dialogue excerpt for Design-World is shown in 45:[1]

(45) BILL: Let's put the green couch in the study
KIM: No, instead let's put the blue lamp in the study
BILL: Putting in the green couch is worth 17
BILL: No, instead let's put the green couch in the study
KIM: Putting in the blue lamp is worth 20
KIM: No, instead let's put the blue lamp in the study
BILL: Then, let's put the red couch in the study

The dialogues are mixed-initiative, since either agent can make a proposal and there is no enforced collaboration since both agents reason about whether to accept another agent's proposal. At the end of a dialogue, the agents have agreed on which four pieces of furniture should go in each room. This is called a COLLABORATIVE PLAN for achieving the intention Design-House. A COLLABORATIVE PLAN will be defined more precisely in section 6.5. Figure 5.2 shows a potential final state.

The main advantage of an artificial world like Design-World is that it is simple and conceivably simulable by robots; there is a limited amount of domain knowledge and only a limited number of actions to reason about. Yet despite this, it provides a natural set of inferences in terms of

[1] This dialogue is produced from the artificial language by adhoc methods.

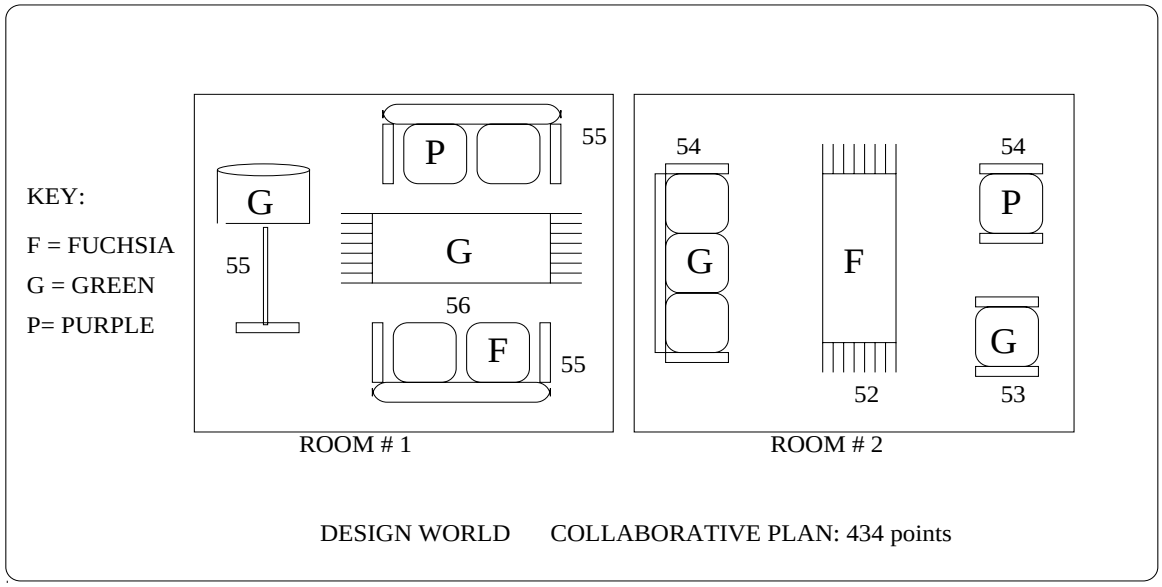


Figure 5.2: One Final State for Design-World Standard Task: Represents the Collaborative Plan Achieved by the Dialogue, 434 points

calculations to maximize utility and the subgoals remaining to complete the design for a room. Another advantage over a real domain is the ease with which the complexity of the domain and the task are varied. [2]

5.4 IRMA Architecture for Resource-Bounded Agents

The agent architecture used in the Design-World simulation environment is based on the IRMA architecture for resource-bounded agents, shown in figure 5.3 (Bratman, Israel, and Pollack, 1988; Pollack and Ringuette, 1990). The IRMA architecture has not previously been used to model the behavior of agents in dialogue. For the purpose of exploring the effects of resource-bounds on attention, this architecture has been extended with the model of limited attention that was discussed in section 3.2 and the model of belief deliberation discussed in section 3.3. The basic components of the version of the IRMA architecture developed in this thesis are:

[2] There are a number of variations that could be constructed in this domain. For instance, the complexity of the task is easily varied by manipulating constraints or restrictions on the sequencing of actions, e.g. the couch must go in the room first. Extra constraints or conflicts could be added by giving agents preferences for putting a particular piece of furniture in a particular room. If the simulation included both a planning and an execution phase, these could be interleaved or done in two phases, and it would be possible to explore which parts of the collaborative plan were best determined at planning time and which at execution time.

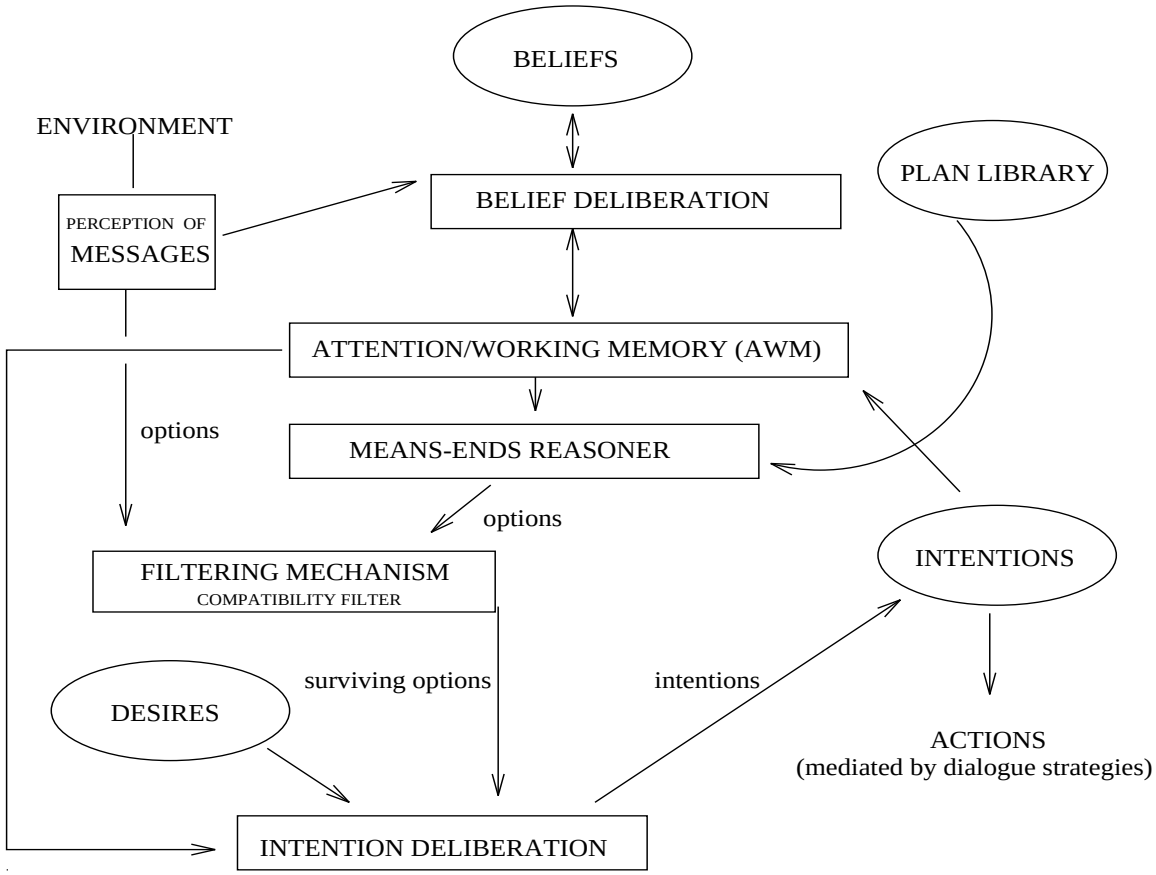


Figure 5.3: Design-World version of the IRMA Agent Architecture for Resource-Bounded Agents with Limited Attention (AWM)

- Attention/Working memory (AWM): the limited attention module constrains working memory and the retrieval of current beliefs and intentions that are used by the means-end reasoner. It provides a level of discourse structure that supports modeling the effects of discourse salience.
- Beliefs: a database of an agent's beliefs. This includes beliefs that an agent supposes are shared. These are stored in an agent's AWM along with intentions.
- Intentions: a database of an agent's intentions. This includes intentions that an agent supposes are shared. These are stored in an agent's AWM along with beliefs.
- Plan Library: what an agent knows about plans as recipes to achieve goals. As a matter of convenience the plan library is not constrained by Attention/Working Memory so that agents always know what the next step in their plan is.

- Means-end reasoner: reasons about how to fill in existing partial plans, proposing options that serve as subplans for the plans an agent has in mind. Means end reasoning will be discussed further in section 5.5.
- Filtering Mechanism: checks options for compatibility with the agent's existing plans. Options deemed compatible are passed along to the deliberation process.[3]
- Intention Deliberation: decides which of a set of options to pursue. Intention Deliberation will be discussed further in section 5.5.
- Belief Deliberation: when there are conflicts in beliefs, decides what to believe. This is based on the belief revision model discussed in section 3.3.2.

Note that the beliefs and intentions used by the means-end reasoner must be in Attention/Working Memory. Thus what is used in reasoning and deliberation is a subset of agents' beliefs and intentions that are currently salient. Also in the Design-World simulation environment, the action types that are reasoned about include both dialogue actions and domain actions. This will be explained in more detail in section 5.1.

5.5 Means-End Reasoning and Intention Deliberation

Means-end reasoning is the process by which an agent considers how it can best achieve a given intention. For example, if the intention is (C) to make coffee, an agent might means-end reason that it can do so by option (A) of boiling water and using instant coffee or by option (B) of using the percolator.

In the IRMA architecture shown in figure 5.3, means-end reasoning is **limited** by attentional capacity. Only the options identifiable from premises currently salient are possible means.[4] Options can also be generated by communication and correspond to proposals that other agents make.

Once the options for how to achieve a goal, e.g. make coffee, are generated by either means-end reasoning or by communication, they must be evaluated. This is the process of intention deliberation. As discussed in section 3.3.1, deliberation is often difficult in real life because there are multiple incompatible evaluation functions that can be applied to competing intentions, because

[3]The filtering mechanism presented in (Bratman, Israel, and Pollack, 1988) and used in Tileworld is more complex than that presented here because BIP were interested in how current intentions act as a filter on intention formation and when current intentions get over-ridden. This use of intentions as a filter is not part of what is to be explained here.

[4]The only limit imposed in Design-World is this limit on the availability of premises for means-end reasoning. Another way to limit means-end reasoning is by limiting the time for reasoning so that an agent can act in real time. Thus an agent might stop means-end reasoning after figuring out methods A and B because it doesn't have any more time to reason about other methods for achieving C.

the probability of particular outcomes is not clear, and because agents may not be able to calculate all possible courses of action.

An adequate account of intention deliberation that takes agents' cognitive capabilities into account awaits future research, so Design-World has a simple evaluation function associated with the points associated with a plan. Agents can evaluate each proposal by first retrieving the score proposition associated with the piece of furniture in the content of the proposal and then checking for additional optional goals that the action might contribute to. Agents evaluate each higher level goal such as Design-Room by simply summing the value of the actions that contribute to the Design-Room goal.

5.6 Design-World Agents' Initial Beliefs, Intentions and Plans

The dimensions of the 3-dimensional space for Attention/Working Memory is set to be 16x16x16. Each agent's memory is initialized with private beliefs as to what pieces of furniture it has and what colors these pieces are. Both agents know what pieces of furniture exist and how many points they are worth. The domain state predicates that form the content of beliefs are:

- (Has-Available ?Agent ?Furniture)
- (Points ?Furniture ?Value)

These belief predicates are annotated as to their positive (POS) or negative (NEG) polarity in the belief database. They are also annotated with their endorsements. So the actual structure of what is represented fits one of the schemas below:

- (?Polarity (Has-Available ?Agent ?Furniture ?Time))
- (?Polarity (Points ?Furniture ?Value))

In other words, what is shown in figure 5.1 as Agent A's piece of furniture, such as a red couch worth 30 points, would be represented by two beliefs in A's belief database: (POS (Has-Available A Red-Couch)), (POS (Points Red-Couch 30)).

In Design-World, there are two types of intentions: utterance intentions and domain intentions. Utterance intentions will be discussed in section 5.7. The domain intentions are intended acts such as putting a particular furniture piece into a particular room in the schema below:[5]

[5]The IRMA architecture allows for intentions to be over-ridden subject to the filter-override. In Design-World, intentions that are already agreed to are never over-ridden. Thus there is no need for an inverse action to Put, such as Remove.

- (Put ?Agent ?Furniture ?Room ?Time)

The only plan in Design-World is DESIGN-HOUSE which has two subgoals DESIGN-ROOM-1 and DESIGN-ROOM-2. Each of these subgoals consists of 4 PUT-ACTS. For each put-act, the agent who has the furniture instantiates the agent variable, ?Agent. The simulations are not concerned with actual execution; if execution were to be added, agents would want to reason about whether two agents are needed for execution, in which case a pair of agents would instantiate the ?Agent variable. The ?Time variable for each action and state can be either NOW or LATER, but since no execution is performed, is instantiated as LATER and is not changed throughout the course of the simulation.

Some versions of the task include additional optional MATCHED-PAIR goals. A matched pair consists of two pieces of furniture of the same color. Matched pair is set up as a task parameter in order to vary inferential complexity. The Matched-Pair version of the task is described in more detail in section 5.9.3.

5.7 Discourse Actions and Discourse Structure

There are four aspects of communicative actions: (1) the utterance level intentions that an agent can achieve, (2) the propositional content of these utterance level intentions, (3) the discourse level acts that these utterance intentions contribute to, and (4) the way in which these intentions are reasoned about. I will discuss the first two of these in the remainder of this section. The way in which these intentions are reasoned about is encapsulated in agents' communicative strategies. These strategies will be briefly discussed in section 5.10 and in more detail in the following chapters.

In Design-World the overall structure of the discourse is completely determined by the task structure (Power, 1974; Grosz, 1977; Sibun, 1991). Agents in Design-World communicate via 7 utterance intentions:[6] Open, Close, Propose, Accept, Reject, Ask and Say (See also (Carletta, 1992; Sidner, 1992)).[7]

Agents take turns sending messages, but each turn may consist of more than one utterance intention. The schema of discourse actions shown in figure 5.4 controls how the utterance intentions are composed. [8] For example, an OPENING may consist of an OPEN utterance, but may include

[6]This is implemented with agents in the same process with a controller to switch control between them. This is a common way of simulating individual processes and has been used in the Tileworld simulation, in Power's robot world and in Carletta's JAM system (Pollack and Ringuette, 1990; Power, 1974; Carletta, 1992).

[7]It would be possible in this domain to leave these intentions implicit and only communicate propositional content. Then hearers would need to infer whether an utterance counts as one of these actions. If this were done, additional inference procedures would be needed to identify the utterance intention (Allen, 1983; Sidner, 1985; Litman and Allen, 1990), and in particular, it would be important to distinguish implicit acceptance from rejection.

[8]This schema is probably not adequate to describe all discourse action transitions in every type of dialogue

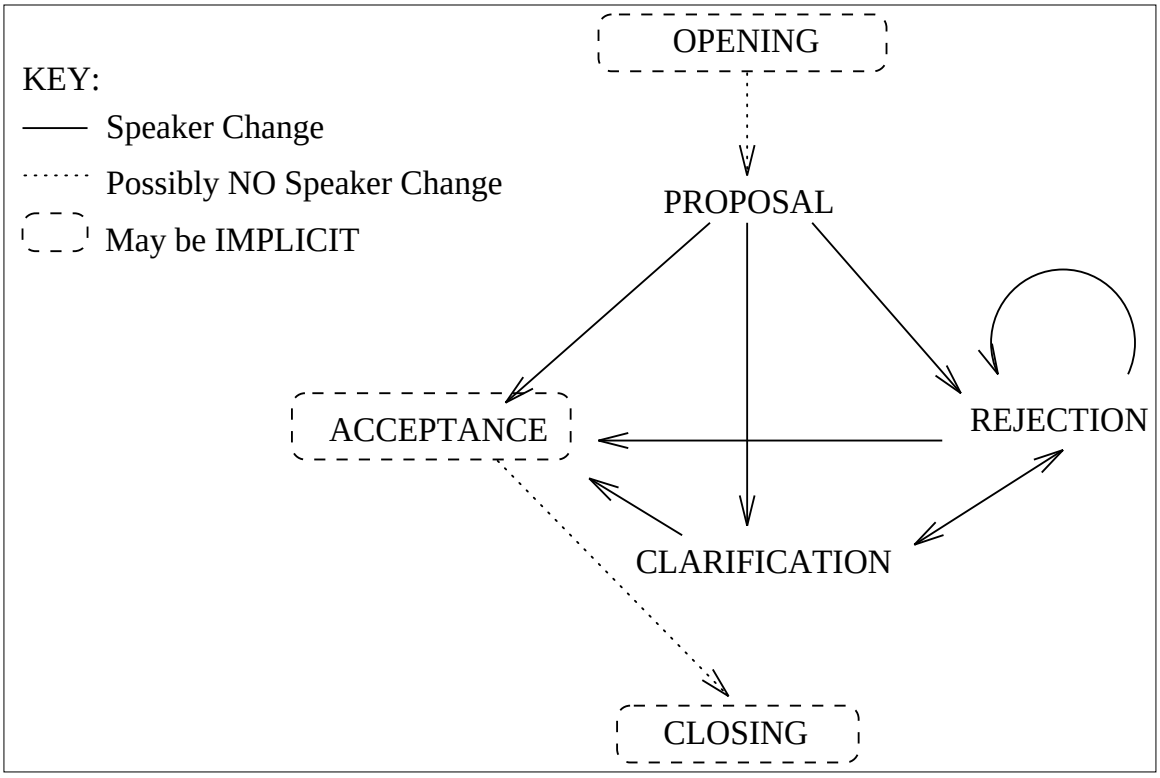


Figure 5.4: Discourse Actions for the Design-World Task

assertions that set the context of the discourse. These variations are determined by an agent's strategy for OPENING.

The form of the openings, closings, proposals, and acceptances shown in figure 5.4 depend on the dialogue strategies. Variations in strategies include whether or not an action is explicitly realized. For example, as the figure shows, opening and closing discourse acts are not always explicitly realized. A proposal can be followed by an acceptance, a clarification or a rejection. Rejections always include counter-proposals. Other variations in openings, acceptances, proposals and closings will be discussed further when I discuss the range of possible strategies in section 5.8.

Utterance intentions are all communicated explicitly, and each communicative act of an agent fits the schema below:

(?Utterance-Intention ?Agent1 ?Agent2 (?Pred ?Vars))*

Possible predicates to instantiate the ?Pred of the content of the utterance are: (1) the domain

(Levinson, 1979; Levinson, 1981; Schegloff, 1987), but is a useful abstraction for the simulation. This abstraction facilitates the investigation of the effects of resource-bounds, which operates independently of the range of utterance level intentions and other strategies available to an agent.

intention predicate of Put which corresponds to proposals that the speaker is making; (2) the state predicates Points, Has. For instance agent A might say:

(Propose A B (Put A&B Red-Couch Room-1 Later))

Agent B can either accept, reject or ask for more information about this proposal. For example, B could reply with an acceptance in the form of a repetition:

(Accept B A (Put A&B Red-Couch Room-1 Later))

B can also clarify the proposal by asking for more information.

(Ask B A (Points Red-Couch ?))

Remember that each agent starts out with beliefs about the points associated with each furniture item, but the model of limited attention/working memory means that they can *forget* what these values are.

B can also reject the proposal:

(Reject B A (Put A&B Blue-Couch Room-1 Later))

The predicate clause in a rejection is a counter-proposal that encodes the reason that B rejects the proposal that A made. B compared A's proposal with other options that B knows about, and in this case, B believes that putting the Blue-Couch in Room-1 is a better option because it is of higher utility.[9]

The utterance intention of Say is used to communicate beliefs about the current state, and to reply to questions. For instance if B asks A:

(Ask B A (Points Red-Couch ?))

A can reply with:

(Say A B (Points Red-Couch 30))

In addition, each of these utterance acts has an effect on the beliefs, intentions or state of the agents when they are processed.

[9]The production of a counter-proposal is a common, but not the only way that rejections are done in human-human dialogues (Whittaker and Stenton, 1988; Walker and Whittaker, 1990).

- Implicit Acceptance:
  - Store (MS (Intended A B ACT)) as a default
  - Store (Act-Effects Act)) as an default (weakest link)
- Explicit Acceptance:
  - Store (MS (Intended A B ACT)) as linguistic
  - Store (Act-Effects Act)) as an entailment
- Implicit Open: Find a discourse segment in task structure that matches the current proposal and mark its status as Open
- Explicit Open: Open the discourse segment that matches the explicit Open
- Implicit Close: Close the current discourse segment
- Explicit Close: Close the current discourse segment
- Say: Store the content of the Say in AWM as linguistic.
- Proposal:
  - Check whether Proposal is compatible with current beliefs, e.g. that no current beliefs contradict its preconditions.
  - Infer that (Has Agent Piece) and store in memory as entailed.
  - Means-End Reason (ME-Reason) about Intention the proposal contributes to.
  - Deliberate by evaluating the proposal against other options generated by Means-End reasoning.
  - Decide whether to Accept or Reject.
- Rejection:
  - Evaluate the proposal that constitutes the content of the rejection.
  - Accept the rejection proposal if it is better than your current proposal.
  - If you decide to reject the rejection proposal then reject with reason for rejection.

When I discuss particular dialogue strategies below I will use the effects of each utterance act as defined above to illustrate the relationship between the strategy and underlying cognitive processes.

The dialogue actions in combination with agents' means-end reasoning and deliberation, lets agents achieve a COLLABORATIVE-PLAN. In chapter 6 I will discuss making plans in more detail. In Design World, a sequence of a proposal followed by an acceptance means that agents agree on that step

of the plan. For example, the plan for (Design Room-1) requires that the agents both know that a room design consists of any 4 pieces of furniture. The agents have to agree that: (1) they intend to put each piece of furniture into a room; (2) these actions of putting furniture into a room contributes to achieving (Design Room-1); and (3) the actions they have chosen have the maximum utility as far as they know at the moment they agree. The dialogue actions shown in figure 5.4 represent a cycle in which:

1. individual agents perform means-end reasoning about options in the domain;
2. individual agents deliberate about which options are preferable;
3. then agents make PROPOSALS to other agents, based on the options identified in a reasoning cycle, about actions that contribute to the satisfaction of their goals;
4. then these proposals are ACCEPTED or REJECTED by the other agent based on calculating whether they maximize utility

For example, Ann conducts means-end reasoning for a set period of time, deliberates on the options generated by her reasoning, selects putting the red couch in room-1 as the best option to pursue, and then proposes to Bob that they put the red couch in room-1. Deliberation requires calculating the point value associated with an option. Bob deliberates about the proposal based on the calculated utility and compares it with other options he knows of. Then Bob may accept the proposal, reject the proposal, or leave the proposal on the table by requesting additional information.

Since, COLLABORATIVE PLANS are established through dialogue, agents' discourse strategies when carrying out the dialogue can affect how efficiently a collaborative plan is established, how good the plan is, and how robust it is to changes in the environment or to changes in agents' intentions. The remaining section briefly discusses discourse strategies that Design-World agents can use that affect how good their collaborative plans are and describes how the performance of the agents in Design-World is evaluated.

5.8 Possible Strategies

Remember that the main goal of Design-World is to demonstrate that IRUs are motivated by agents' cognitive limitations. Demonstrating this hypothesis involves varying the **way** that agents open and close dialogue segments, make proposals and communicate acceptance. There are a limited

number of possible strategies: (1) structure for a discourse segment related to a particular Design-World intention is limited, and (2) the domain limits what kinds of inferences can be made and what facts are relevant to performing the task.

5.8.1 The All-Implicit (Baseline) Strategy

The simplest strategy is when only proposals and rejections are said explicitly and all the other actions for a segment are inferred to have occurred. This strategy is All-Implicit, and is the BASELINE against which more complex strategies are compared.

Figure 5.5 shows the discourse action protocol for All-Implicit agents.

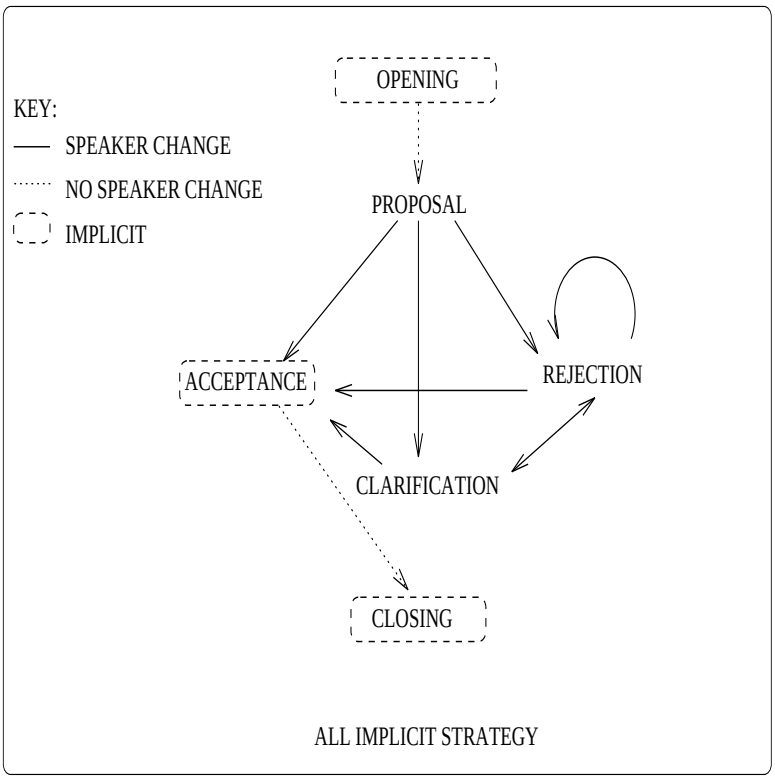


Figure 5.5: Discourse Action protocol for agents using the All-Implicit strategy

The All-Implicit strategy is exemplified by both agents Bill and Kim in the complete dialogue given in 46.

(46) BILL: *Then, let's put the green rug in the study.*
1:(propose agent-bill agent-kim option-10: put-act (agent-bill green rug room-1))
KIM: *Then, let's put the green lamp in the study.*
2:(propose agent-kim agent-bill option-33: put-act (agent-kim green lamp room-1))
BILL: *Then, let's put the green couch in the study.*

Agent A	Agent B
PROPOSAL 1	
	PROPOSAL 2
PROPOSAL 3 REJECTION 3,4	
	PROPOSAL 5
...	...
	PROPOSAL N

Figure 5.6: Sequence of Discourse Acts for Two All Implicit Agents in Dialogue 46

3:(propose agent-bill agent-kim option-45: put-act (agent-bill green couch room-1))
KIM: *No, instead let's put the purple couch in the study.*
4:(reject agent-kim agent-bill option-56: put-act (agent-kim purple couch room-1))
BILL: *Then, let's put the green couch in the study.*
5:(propose agent-bill agent-kim option-61: put-act (agent-bill green couch room-1))
KIM: *Then, let's put the purple chair in the living room.*
6:(propose agent-kim agent-bill option-81: put-act (agent-kim purple chair room-2))
BILL: *Then, let's put the fuchsia couch in the living room.*
7:(propose agent-bill agent-kim option-95: put-act (agent-bill fuchsia couch room-2))
KIM: *Then, let's put the purple rug in the living room.*
8:(propose agent-kim agent-bill option-108: put-act (agent-kim purple rug room-2))
BILL: *No, instead let's put the green chair in the living room.*
9:(reject agent-bill agent-kim option-116: put-act (agent-bill green chair room-2))
KIM: *Then, let's put the purple rug in the living room.*
10:(propose agent-kim agent-bill option-125: put-act (agent-kim purple rug room-2))

Note that acceptance is never communicated explicitly in 46; agents infer that another agent accepted its proposal from the fact that nothing was said to indicate rejection. In addition, opening and closing statements are left implicit, and no IRUs that might function to reduce inference or retrieval are included. Figure 5.6 abstracts from the dialogue in 46 to the discourse acts and figure 5.7 shows the cognitive processes that correspond to each utterance.

5.8.2 Communicating Rejection

Some dialogue behaviors are used by all of the agents for particular situations. For example, every agent has a way to resolve conflicts that involves IRUs. If an agent makes a proposal and another agent rejects its proposal, and if the first agent wants to reject the counter proposal, then she makes her reasoning explicit as to why she is committed to her original proposal. Remember that the

Discourse Act	Utterance Act	A PROCESS	B PROCESS
PROPOSAL	2:(propose A B option-10)	ME-REASON Put-1 Deliberate Propose, All Imp	Infer Open Put-1 ME-REASON Put-1 Check Preconds 10 Deliberate Decide to Accept Store (MS intend 10) Store (Act-Effects 10) ME-REASON Put-2 Deliberate Propose, All-Imp
PROPOSAL	3:(propose B A option-33)	Infer Acceptance 10 Store (MS intend 10) Store (Act-Effects 10) Infer Close Put-1 Infer Open Put-2 ME-REASON Put-2 Check Preconds 33 Deliberate Decide to Accept Store (MS intend 33) Store (Act-Effects 33) ME-REASON Put-3 Deliberate Propose	
	... .		

Figure 5.7: Cognitive Processes for Sequence of Discourse Acts in Dialogue 46: Agent A and B are All Implicit Agents

point values of furniture items are known to both agents at the start of the dialogue, but agents can forget due to limits on AWM. The conflict resolution strategy is to produce an IRU reminding the other agent of the value of the piece in the current proposal. This is shown in example 47 below, where Bill in 6 and 7 presents his reason and repeats his proposal.

(47) BILL: *Then, let's put the green couch in the study.*

4:(propose agent-bill Agent-Kim option-45: put-act (agent-bill green couch room-1))

Kim: *No, instead let's put the green lamp in the study.*

5:(reject agent-kim agent-bill option-49: put-act (agent-kim green lamp room-1))

BILL: PUTTING IN THE GREEN COUCH IS WORTH 54.

6:(say agent-bill agent-kim bel-70: score (option-45: put-act (agent-bill green couch room-1) 54))

BILL: *No, instead let's put the green couch in the study.*

7:(reject agent-bill agent-kim option-45: put-act (agent-bill green couch room-1))

Another strategy that all the agents use is to reject a proposal when, according to their own beliefs, the preconditions for the proposal are not met. In other words, the rejecting agent believes that the furniture item is not available. An example of a rejection like this is shown in 48.

(48) BILL: *Then, let's put the green chair in the living room.*

16:(propose agent-bill agent-kim option-77: put-act (agent-bill green chair room-2))

Kim:*Agent-Bill doesn't have the green chair.*

17:(say agent-kim agent-bill bel-1006: has n't (agent-bill green chair))

Kim: *No, instead let's put the purple chair in the living room.*

18:(reject agent-kim agent-bill option-87: put-act (agent-kim purple chair room-2))

5.8.3 Range of Strategies that include IRUs

For the purpose of showing the range of strategies and explaining the general strategies, I will assume that there are no costs for communication and processing. In sections 5.9.4 and 5.10 I will discuss ways of evaluating performance that take both task definition and processing costs into account. I will use the term RAW SCORE to refer to the number of total points the agents achieve on the task without any consideration of costs to achieve those points. Thus in the remainder of this section, the agents' score on a task is simply their RAW SCORE, i.e. the sum of the pieces used in the plan. For dialogue 46, Bill and Kim's AWM was set to 11 and their RAW SCORE is 434 points. Bill and Kim got the highest possible RAW SCORE even without being able to search all of memory.

To see what the range of possible strategies are, consider the dialogue actions that can make up a single discourse segment as shown in figure 5.4 and schematized here in figure 5.8. The way that REJECTION and CLARIFICATION are communicated is the same for all agents, so the places where variations in strategy can occur is in the OPENING, PROPOSAL, ACCEPTANCE and CLOSING discourse acts.

The types of IRUs in general are the types of HEARER OLD information, such as REPETITION, PARAPHRASE, ENTAILMENTS, IMPLICATURES and PRESUPPOSITIONS. The only types of IRUs that are used in the simulation are REPETITION and ENTAILMENT, ie. repeating something that is already known or making an inference explicit.

Thus considering the places where variation can occur and the types of IRUs in the simulation, it is possible for each of OPENING, PROPOSAL, ACCEPTANCE and CLOSING to include one or more repetitions or entailments. Attention strategies are defined as those in which a repetition occurs in either an opening, proposal or closing. In addition the Attention strategies vary as to whether only one fact is repeated or multiple facts are repeated. Consequence strategies are those where an inference is made explicit in either an acceptance or a closing. Attitude strategies are those where a repetition occurs in an acceptance. These possible variations are shown in figure 5.8, and their instantiations will be discussed in the following chapters.

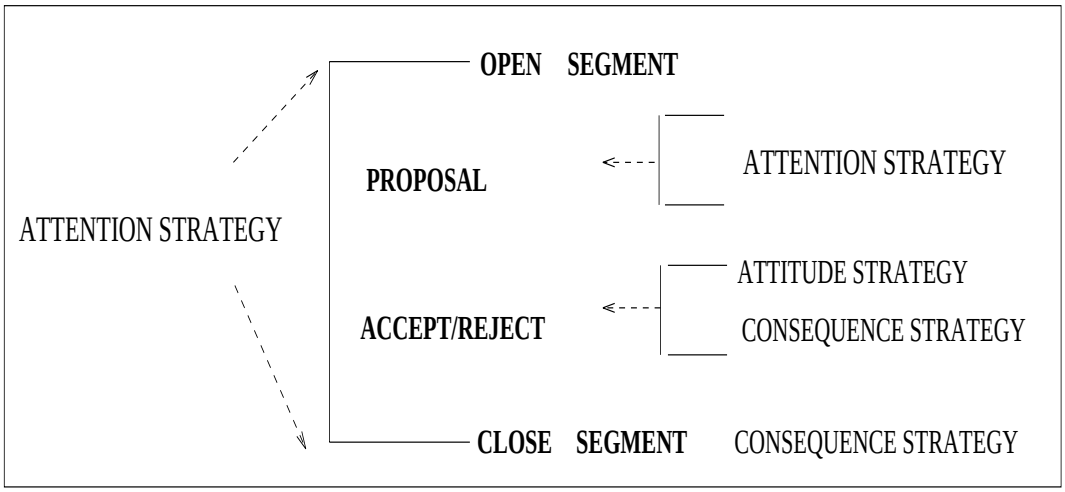


Figure 5.8: Schema for the Primary Effect of Dialogue Strategies on Dialogue Structure

5.9 Parameters for Cognitive Capabilities and Tasks

This section discusses the range of parameters for agents' cognitive capabilities related to limits on attention and inference. I also introduce a variation in the basic task that increases the inferential complexity of the task. In addition to the cognitive and task parameter variations discussed here, another way to effectively vary the task is by varying the way that performance is evaluated. These task variations will be discussed in section 5.10. These parameters will be used to show that the benefits of discourse strategies are situation specific.

5.9.1 Limited Attention Effects

.

For every strategy type discussed above, the parametrization of AWM provides a major limitation on how agents perform. For example, consider the dialogue given in 49 where Bill and Kim's AWM is set to 4.

(49) BILL: *Then, let's put the green couch in the study.*
1:(propose agent-bill agent-kim option-5: put-act (agent-bill green couch room-1))
KIM: *Then, let's skip the next one.*
2:(propose agent-kim agent-bill option-9: put-act (agent-nobody))
BILL: *Then, let's skip the next one.*
3:(propose agent-bill agent-kim option-10: put-act (agent-nobody))
KIM: *Then, let's put the red rug in the study.*
4:(propose agent-kim agent-bill option-12: put-act (agent-kim red rug room-1))
BILL: *Then, let's skip the next one.*
5:(propose agent-bill agent-kim option-13: put-act (agent-nobody))
KIM: *Then, let's skip the next one.*
6:(propose agent-kim agent-bill option-14: put-act (agent-nobody))
BILL: *Then, let's skip the next one.*
7:(propose agent-bill agent-kim option-15: put-act (agent-nobody))
KIM: *Then, let's put the purple couch in the living room.*
8:(propose agent-kim agent-bill option-17: put-act (agent-kim purple couch room- 2))

In 49, because AWM is set to 4, Bill and Kim cannot perform as well. When they cannot come up with an option for a plan step, they propose that they skip that step. As with other proposals, this proposal can be accepted or rejected by the other agent. With limited AWM, the agents cannot remember which pieces they have and what their values are. The raw score for the task is 129 points.

As discussed above, limited attention is modeled in Design-World by manipulating the memory radius parameter. The value of this parameter can range from 1 to 16, so that agents have limited attention at lower values. This provides an operationalization of the effect of salience.

Figure 5.9 shows how the parametrization of limited attention affects performance on the Design-World task. As shown in figure 5.9, scores increase as attentional capacity (memory search radius) increases.

The highest raw score achievable for the task is 434 points. Note also that the raw score is optimal at a certain point depending on the number of facts stored in memory and the complexity of the task. For the particular version of the task executed here, figure 5.9 shows that the agents achieve

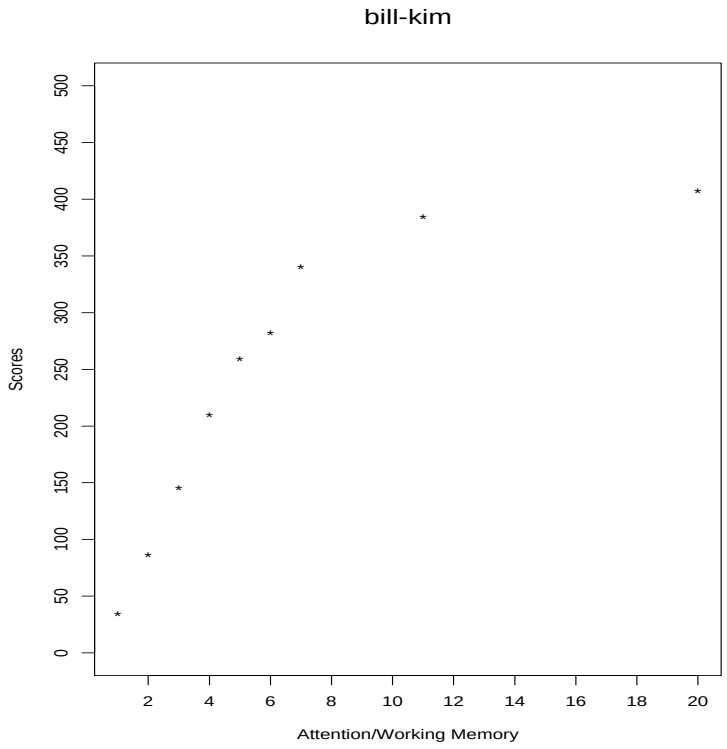


Figure 5.9: Baseline, Two Attention Implicit Agents

the maximal raw score when the search radius is set to 16, but are near optimal by AWM of 11. This means that if a particular level of performance is required, agents' resource parameters can be set so that that level of performance is guaranteed. However, if action is required at low levels of resources, reasonable behavior can be produced (Dean and Boddy, 1988).[10]

5.9.2 Varying Inferential Capacity

Another parameter that can be varied in Design-World is whether agents are inference limited. While it is widely agreed that human agents are not logically omniscient, there is currently no adequate theory of which inferences are made and what factors are involved in determining that these inferences, and not others, are the ones that are derived (Levinson, 1985; Konolige, 1986). Design-World implements limited inference in a straightforward way because there are only a limited number of possible inferences.

One source of inference limitations is provided by the fact that inferences can only be carried out on beliefs that are currently salient. This is the DISCOURSE INFERENCE CONSTRAINT. The primary

[10]Note that what is reasonable behavior depends on the definition of the task. If the task definition requires that agents put 4 pieces of furniture in every room in order to complete the task, then clearly, as shown in figure 5.9 they could not accomplish the task at low attention values.

effect of this is to limit which premises can be using in reasoning about how to achieve goals, i.e. means-end reasoning about put-acts and about achieving matched-pairs (to be discussed below). All agents are inference limited via this 'data limit' on what can be used as premises in inferences by the parameter setting for AWM, as discussed in section 5.9.1.

A second set of inference limitations is to limit inferences based on actions that agents agree on in the domain. These are primarily act-effect inferences of the form: If A has just agreed to use the red couch in room-1, then A no longer has the red couch available. If A cannot make this inference, then A might believe that the red couch could be used again to satisfy another intention. Design-World supports parametrizing agents as to whether they consistently draw these act-effect inferences or not. While this is a trivial inference, it provides a reasonable test-bed of the effect of not making critical inferences.

Agents can be ALL-INFERENCE agents, i.e. they are logically omniscient and always make all the act-effect inferences. Agents can also be parametrized to be NO-INFERENCE agents, i.e. they draw no inferences about the effects of proposals that they make or that they agree to. Finally, agents can be HALF-INFERENCE agents, which means that half the time they draw the act-effect inferences.

Figure 5.10 shows the scores for dialogues between two NO-INFERENCE agents. As one would expect they do very poorly on the task, because they don't infer that once they have decided to use a piece of furniture, then it is no longer available to use.

Figure 5.11 shows the scores for dialogues between two agents who make half of the valid inferences. These agents perform much better than the NO-INFERENCE agents.

5.9.3 Varying Inferential Complexity: Matched Pairs

Another set of parameters vary inferential complexity. For example, another version of the task includes optional goals to try to get matched pairs in each room at the same time as agreeing on the design for the house. A matched pair is worth the points for each furniture item used in the matched pair, plus an 50 point bonus for each matched pair. In this version of the task, Bill and Kim evaluate each proposal to see whether it can also achieve a matched pair. Compare 46 with 50. Again Bill and Kim both have AWM at 11.

(50) BILL: *Then, let's put the green rug in the study.*
1:(propose agent-bill agent-kim option-10: put-act (agent-bill green rug room-1))
KIM: *Then, let's put the green lamp in the study.*
2:(propose agent-kim agent-bill option-35: put-act (agent-kim green lamp room-1))
BILL: *Then, let's put the green couch in the study.*
3:(propose agent-bill agent-kim option-50: put-act (agent-bill green couch room-1))

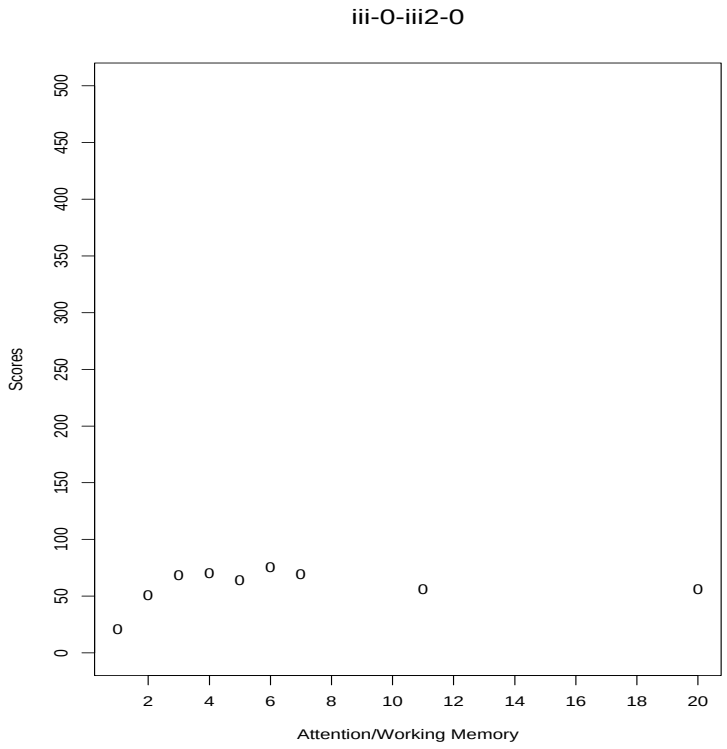


Figure 5.10: Scores for dialogues between Two No-Inference Agents: Evaluation Function = COMPOSITE-COST, commcost = 0, infcost = 0, retcost = 0

KIM: *No, instead let's put the purple couch in the study.*

4:(reject agent-kim agent-bill option-61: put-act (agent-kim purple couch room-1))

BILL: *Then, let's put the green couch in the study.*

5:(propose agent-bill agent-kim option-66: put-act (agent-bill green couch room-1))

KIM: *No, instead let's put the purple chair in the study.*

6:(reject agent-kim agent-bill option-81: put-act (agent-kim purple chair room-1))

BILL: *Then, let's put the green couch in the living room.*

7:(propose agent-bill agent-kim option-87: put-act (agent-bill green couch room-2))

KIM: *Then, let's put the green table in the living room.*

8:(propose agent-kim agent-bill option-106: put-act (agent-kim green table room-2))

BILL: *No, instead let's put the green chair in the living room.*

9:(reject agent-bill agent-kim option-115: put-act (agent-bill green chair room-2))

KIM: *Then, let's put the purple rug in the living room.*

10:(propose agent-kim agent-bill option-126: put-act (agent-kim purple rug room-2))

BILL: *Then, let's put the fuchsia rug in the living room.*

11:(propose agent-bill agent-kim option-141: put-act (agent-bill fuchsia rug room-2))

KIM: *No, instead let's put the purple table in the living room.*

12:(reject agent-kim agent-bill option-151: put-act (agent-kim purple table room-2))

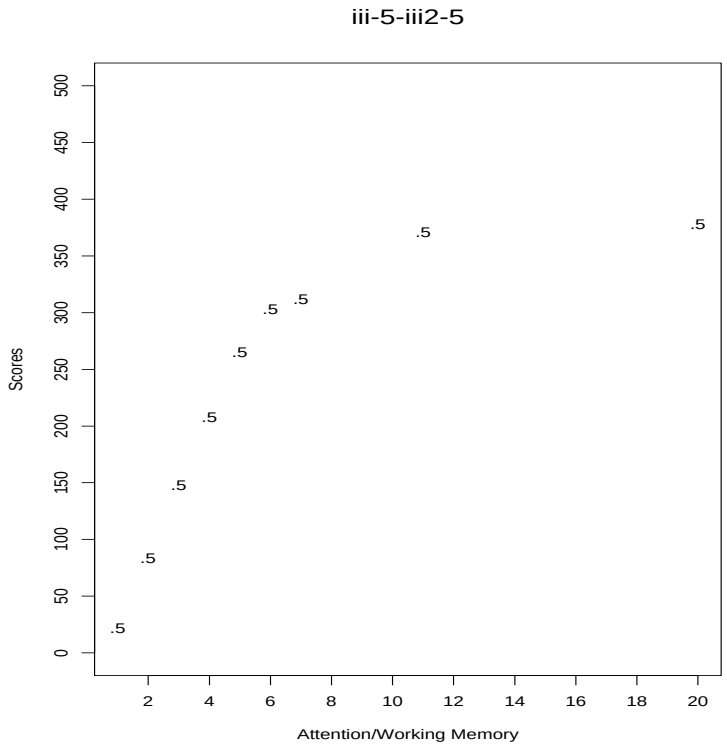


Figure 5.11: Scores for dialogues between Two Half-Inference Agents: Evaluation Function = COMPOSITE-COST, commcost = 0, infcost = 0, retcost = 0

In 50, Bill and Kim get a raw score of 430 points for the Design-Room task, but get additional points for the Matched-Pair optional task. Because the point of MATCHED-PAIR inferences is to test the value of agents coordinating on what inferences they make, a matched pair is only achieved if both agents make the inference that two intended actions will constitute a matched pair. If one agent makes the inference and the other doesn't, they do not get any additional matched-pair points. In the dialogue shown in 50, Bill and Kim achieve 4 matched pairs:

1. (intended-35: put-act (agent-kim green lamp room-1)), (intended-9: put-act (agent-bill green rug room-1));
2. (intended-81: put-act (agent-kim purple chair room-1)), (intended-61: put-act (agent-kim purple couch room-1));
3. (intended-108: put-act (agent-bill green chair room-2)); (intended-84: put-act (agent-bill green couch room-2));
4. (intended-151: put-act (agent-kim purple table room-2)), (intended-126: put-act (agent-kim purple rug room-2))

Since these matched pairs use all actions that the agents agreed on, the matched pair total is 430 plus 200 incentive points for a total of 630. The incentive points allow the performance measure to clearly show whether a strategy makes it easier to make matched pair inferences.

For each strategy it must be possible to evaluate the benefit of that strategy. The benefits of a strategy depend on how the task is defined and what the costs are to achieve the raw scores. Here I assumed that the costs involved in achieving raw scores could be ignored, but in general communication, inference and retrieval will cost something. The range of evaluation functions used to judge the efficacy of a strategy will be discussed in section 5.10. First in section 5.9.4, I will define ways in which the basic Design-World task can be varied.

5.9.4 Varying the Definition of Task Success

The Design-World task is simple and straightforward but can be parametrized to make it more difficult to perform, or so that different aspects of the task weigh more heavily in performance evaluation. There are 5 variations in the task definition.

The Standard task is defined so that the RAW SCORE that agents achieve for a particular dialogue consists of the sum of all the furniture pieces for each valid step in their plan. The point values for invalid steps in the plan are simply subtracted from the score, ie. agents are not heavily penalized for making mistakes due to trying to use the same piece twice.

Additionally, there are two other factors related to the quality of a solution, the degree of MATCHING BELIEFS required, and the effect of mistakes. The reason to explore the effect of the degree of matching beliefs required is that different collaborative tasks require different levels of agreement. The reason to explore the effect of mistakes, is that this effect varies depending on the interdependency of different subparts of the problem solution. It is simple to vary the task definition in the artificial Design-World task to reflect these very real differences between tasks in the real world. Below I will discuss two variation on the Standard task: (1) Zero-Invalids in which agents get no points for their plan if it has any invalid steps in it, and (2) Zero-Nonmatching-Beliefs: in which agents get no points for their plan if the have different reasons underlying each plan

Another task variation is whether agents are supposed to get MATCHED-PAIRS as discussed in section 5.9.3. Another issue is how the achievement of matched pairs is defined to contribute to the score. I will discuss two variations of this: (1) Matched-Pair (MP): RAW SCORE consists only of the points that agents get for achieving matched-pairs and the points they get for the standard task are ignored; (2) Matched-Pair-All (MPALL): RAW SCORE consists of both the points that agents get for achieving matched-pairs and the points for the standard task.

5.9.4.1 Zero Invalids Task Definition

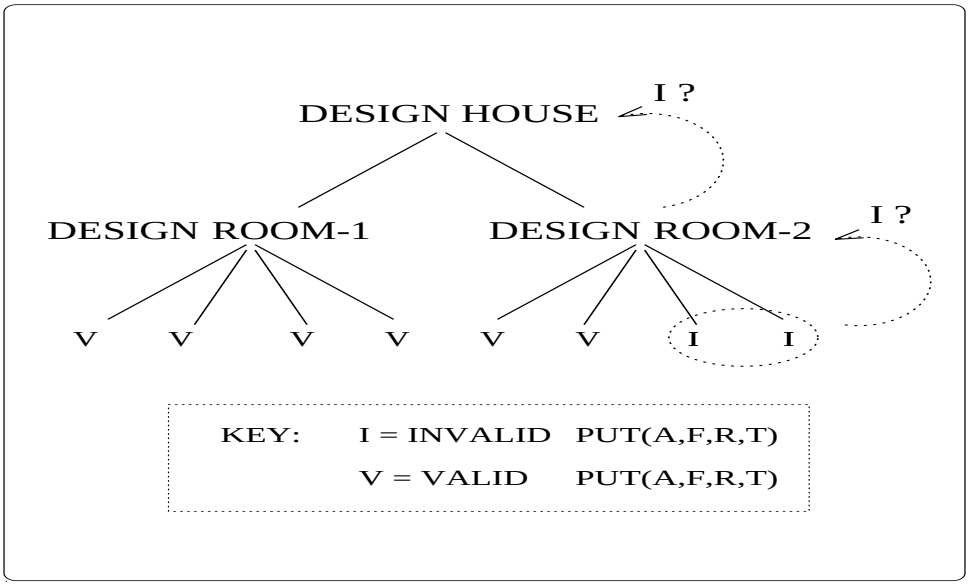


Figure 5.12: Evaluating Task Invalids: for some tasks invalid steps invalidate the whole plan.

One of the task definition parameters is related to the the effect of mistakes. A mistake results in an invalid step in a plan where an invalid step is one that cannot actually be executed. The effect of making a mistake in a plan depends on how interdependent different subparts of the problem solution are. For some tasks a mistake can invalidate the whole solution, for other tasks, partial solutions without the invalid step may be adequate. For example in a task like furnishing a room it may be desirable to have both a couch and a chair, but if the agents don't manage to agree on which chair to put in the room, or assume they can use a chair that will end up in a different room, the room is still partially furnished and usable. On the other hand, in a task such as building a tower, each step depends on the successful execution of the previous step and the whole plan may be invalid if a step to put down a foundation block cannot be executed.

In Design-World it is simple to define the level to which the substeps of the task are interdependent. Figure 5.12 shows what the choices are for the effect of invalid steps for the Design-World task. The score for invalid steps (mistakes) can just be subtracted out; this is how the Standard task is defined. Alternately, invalid steps can propagate up so that a room that isn't completely furnished doesn't contribute to the final score. In other words if the plan for a room has an invalid step in it, then the plan for the whole room is invalid. Finally, mistakes can completely propagate to the top level of the plan so that the whole plan is invalid if one step is invalid.

The Zero-Invalids task definition reflects task situations in which invalid steps invalidate the whole plan.

5.9.4.2 Matching Beliefs Task Definition

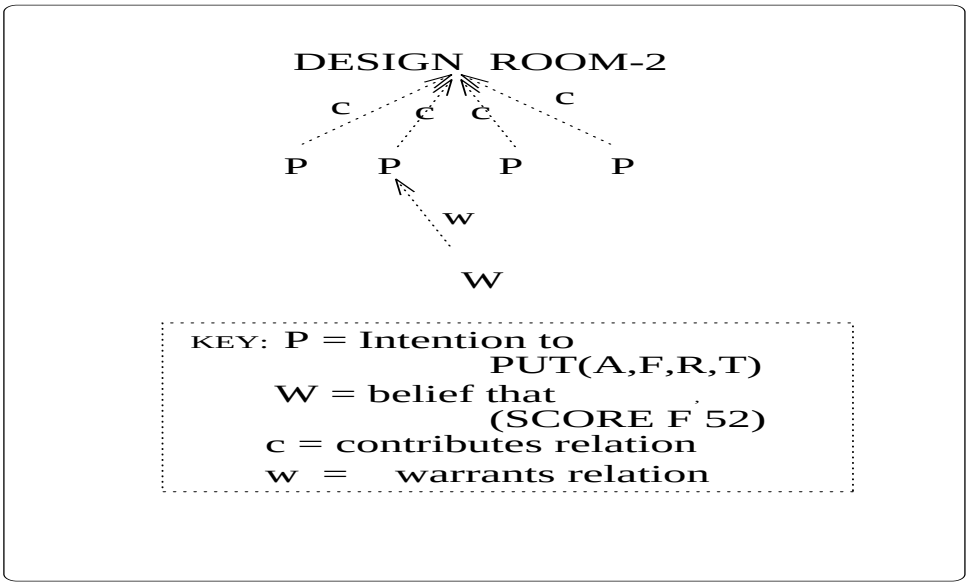


Figure 5.13: Tasks can differ as to the level of mutual belief required. Some tasks require that W, a reason for doing P, is mutually believed and others don't.

The Zero-Nonmatching-Beliefs task definition is designed to investigate the effect of different task requirements on the level of agreement that agents must achieve. It is easy to see that there are different degrees of agreeing in a collaborative task. It may be necessary for agents to agree on every aspect of the task, but it is not **always** necessary for agents to agree on e.g. the **reasons** for carrying out a particular action.

For example, in the negotiation between the union and the management of a company, any agreement that is reached is agreed to by each party for different reasons. An agreement for a shorter work week is supported by the union because more overtime pay is possible for those who want to work more and is supported by the management because the company's insurance premiums will be lower.

If two agents agree on a plan, but have different reasons for doing so, they may change their beliefs and their intentions under different conditions. The most stable, long-term, collaborative plans will be those in which agents agree on both the actions to be performed, as well as the reasons for doing those actions. Under these conditions the agents will be more likely to change their goals in a compatible way.

Figure 5.9.4.2 shows the structure of beliefs about intentions and warrants in Design-World. Depending on how the task is defined, the level of agreement required varies. As shown in figure

5.9.4.2, the Design-World task could be defined so that W, a reason for doing P, must also be mutually supposed. The variation in task definition for level of agreement is formulated as to whether agents have to agree on the WARRANT for an action under discussion.

In Design-World, a warrant W is a belief about the points proposition associated with a piece of furniture. In terms of the definition of a collaborative plan, to be given in section 6.5, agents A and B can mutually suppose that they have maximized utility without necessarily agreeing on what that utility is. Evaluating performance in Design-World can reflect these different task situations simply by changing the way that RAW SCORE is tabulated. If the task definition requires W to be mutually believed, then the function ZERO-NONMATCHING-BELS is applied to the RAW SCORE. The new RAW SCORE is then modified further by taking into account the costs of achieving that score. This will be discussed in section 5.10.

5.9.4.3 Matched Pair and Matched Pair All Task Definition

Finally, the task definition can be modified specifically for the Matched-Pair tasks discussed above that increase inferential complexity. The Matched-Pair (MP) task tallies only the points achieved from matched-pair goals and ignores the goals for the standard task. This is used to reflect situations in which it is desirable to see whether a strategy helps make a particular type of inference, possibly at the expense of performance overall.

In the Matched-Pair-All (MPALL) task, the raw score is the sum of the points achieved from matched pair goals and the points achieved in the Standard task. This task definition is used to model situations in which an inference is only valuable as long as it doesn't disrupt any other aspect of the agents' performance.

5.9.5 Summary

In sum, the variations in the task definitions are:

- Standard: the raw score for a collaborative plan is simply the sum of the furniture pieces in the plan with the values of the invalid steps subtracted out.
- Zero-Invalid: give a zero score to a collaborative plan with any invalid steps in it, reflecting a binary division between task situations where the substeps of the task are independent from those where they are not. See figure 5.12.
- Zero-Nonmatching-Beliefs: give a zero score to a collaborative plan when the agents disagree on the reasons for having adopted a particular intention. These reasons are the WARRANTS

for the intention. The way this difference is reflected in Design-World is to check whether the agents agree on the value of each piece of furniture in the final plan. See figure 5.9.4.2.

- Matched-Pair (MP): tabulate only the scores for the matched-pair optional goals of the Matched-Pair and Matched-Pair-Two-Room task.
- Matched-Pair-All (MPALL): tabulate the scores for both the standard task and the matched-pair optional goals of the Matched-Pair and Matched-Pair-Two-Room task.

5.10 Evaluating Performance

The term RAW SCORE refers to the number of total points the agents achieve on the task without any consideration of costs to achieve those points. The term COSTS refer to retrieval, inference and sending and receiving messages, and PERFORMANCE is used to describe raw score with costs subtracted.

The Design-World task provides an objective measure of benefits that is based on summing the values of the pieces of furniture used in the final plan, or summing these along with the points achieved for matched-pair inferences. In addition, every solution has certain costs associated with it. The costs that must be weighed along with the objective measure of the quality of the solution are: (1) costs of communication, based on number of messages; (2) costs of inference, based on tracking the number of inferences required to complete the task; and (3) costs of retrieval, based on tracking the number of steps required to retrieve needed facts from memory. These costs are highly relevant when agents are resource-bounded; if agents have limited resources then a solution that is too expensive cannot be achieved.

5.10.1 Composite Cost Evaluation Function

The goal of the COMPOSITE-COST evaluation function is to vary evaluation of the resulting design under different assumptions about the relation between (1) COMMCOST: cost of sending a message; (2) INFCOST: cost of inference; and (3) RETCOST: cost of retrieval from memory. PERFORMANCE is calculated as:

PERFORMANCE = Task Defined RAW SCORE
- (COMMCOST × number of messages)
- (INFCOST × number of inferences for both agents)
- (RETCOST × number of memory retrieval steps for both agents)

The way PERFORMANCE is defined reflects the fact that agents are meant to collaborate on the task. The RAW SCORE for each individual agent is simply the RAW SCORE that they achieve together. The costs that are deducted from this RAW SCORE are the costs for both agents' processing. The total amount of communication is the sum of both agents' messages. Total inferences is the sum of both agents' inferences. Total retrieval is the sum of both agents' retrieval. Thus the COMPOSITE COST function gives a measure of COLLABORATIVE EFFORT and reflects the assumption that agents want to achieve the LEAST COLLABORATIVE EFFORT (Clark and Wilkes-Gibbs, 1986; Clark and Schaefer, 1989; Brennan, 1990).

Varying costs reflects assumptions about different kinds of processors or different kinds of communication situations. In some situations, communication may be very inexpensive or free, for example picking up the telephone and making a local call. In other situations, say if messages can only be sent via a keyboard or an agent can only type with his toes, communication may be costly. Inference costs may vary from one agent to another. Retrieval costs could be varied to demonstrate differences between human and artificial agents. Most humans are limited as to the amount of retrieval they can do, but fast parallel machines can retrieve considerably more at less cost. Thus, there is no 'right' answer as to what these costs might be.

The approach here is to make these costs parameters of the model and explore what the effects are of different assumptions about costs. All of the simulations in the following chapters have been evaluated in each task situation for a range of cost parameters.

There are two limitations of the current definition. One is that all inferences are counted the same, and thus is dependent on the domain model and how much information each inference contains. It is possible to limit the amount of difference between different types of inferences by setting INFCOST very low in order to reduce the effect of inference overall. However, it would be necessary to add other parameters to the evaluation function to reflect the fact that some inferences may be more costly than others.

A second limitation is that it might be desirable to demonstrate different assumptions about memory access. In this implementation, retrieval is linearly related to overall performance by some factor determined by retcost. In order to simulate parallelism, one would want to define total retrieval cost as, e.g. $log_3(retrievalsteps)$. This would mean that each memory radius is accessed serially, but all the loci at each memory radius value are accessed in parallel. Using a retrieval cost function like this would be closer to Landauer's original formulation of the model (Landauer, 1975).

Independent of particular assumptions that are used to calculate performance, the goal is to be able to compare different strategies and to test whether the differences in performance are significant. How significance is determined will be discussed in section 5.10.2.

5.10.2 Evaluating Statistical Significance

This section discusses issues with evaluating the statistical significance of differences in performance in the simulations. First, the number of samples required to approximate the theoretical distribution was determined. Second, the Kolmogorov-Smirnov two sample test was identified as an appropriate test for comparing the distributions of the two samples.

5.10.2.1 Raw Score approximates Beta Distributions

Sample distributions from runs of two All-Implicit agents for each AWM setting are shown in figure 5.14. The distributions in figure 5.14 demonstrate the increase in RAW SCORE that we would expect with increases in AWM. These distributions approximate Beta distributions, and this approximation was used to determine the number of runs necessary for stable results.[11]

The Beta distribution with the largest variance, for parameters R and S greater than or equal to 1, is the uniform distribution. This largest variance distribution would require approximately 133 samples. Since none of the variances were that large (see figure 5.14), sample size was set at 100. An empirical evaluation of the adequacy of this sample size for three different strategies was tested to see if any differences showed up in alternate runs of 100; no differences were found.[12]

5.10.2.2 Kolmogorov-Smirnov Two-Sample Test

Generally, we are interested in how the performance of one pair of agents compares to the performance of another pair of agents on a particular version of the task. The point of the dialogue strategies is to see whether performance can be improved across several AWM settings simply by varying the dialogue strategy, while holding everything else constant. Thus particular strategies will be shown to beneficial for inference-limited agents, others for attention-limited agents and others dependent on the task situation.

[11] I am indebted to Max Mintz for consultancy on the best way to determine the number of runs required for stable distributions and for guidance on the development of a program to calculate differences in distributions using the Kolmogorov-Smirnov two-sample test.

[12] We will see below in section 5.10 that changing the definition of the task transforms these distributions so that they are no longer Beta distributions. However since these transforms are applied **after** the initial sample collection, I assume that a sample size of 100 is adequate.

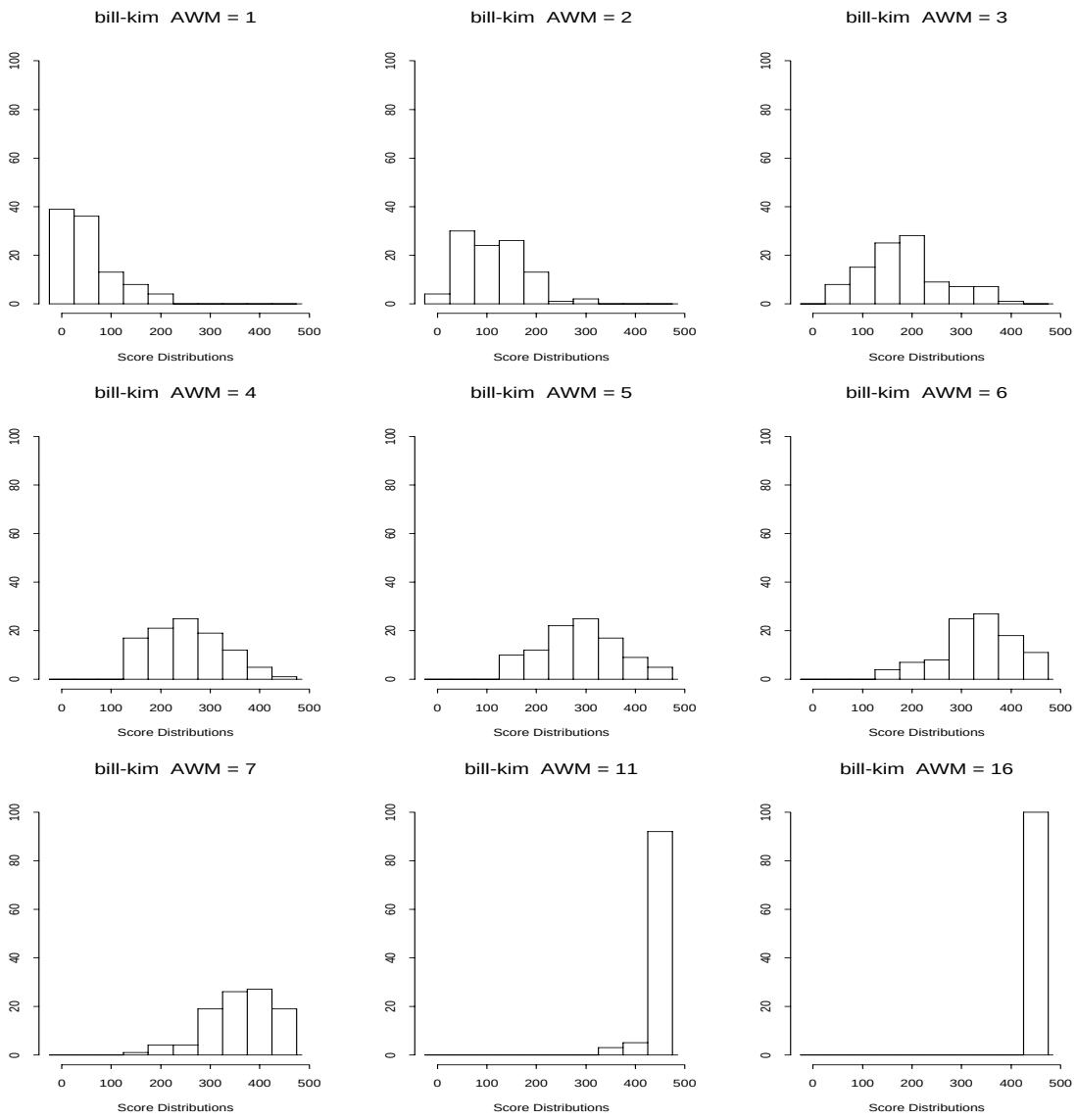


Figure 5.14: Histograms of Score Distributions for Dialogues between two All-Implicit Agents for all AWM settings

Because the approximation to Beta distributions is not exact, and because various transforms change the distributions in non-predictable ways, a non-parametric test was required. The differences in performance are evaluated using the Kolmogorov-Smirnov two-sample test (henceforth KS) (Siegel, 1956; Wilks, 1962). The KS two-sample test is a test of whether two independent samples have been drawn from the same population or from populations with the same distribution. The results below use the two-tailed version of the test. The two-tailed test is sensitive to any kind of difference in the distributions from which the two samples were drawn, e.g. differences in location (central tendency), in dispersion, in skewness etc.

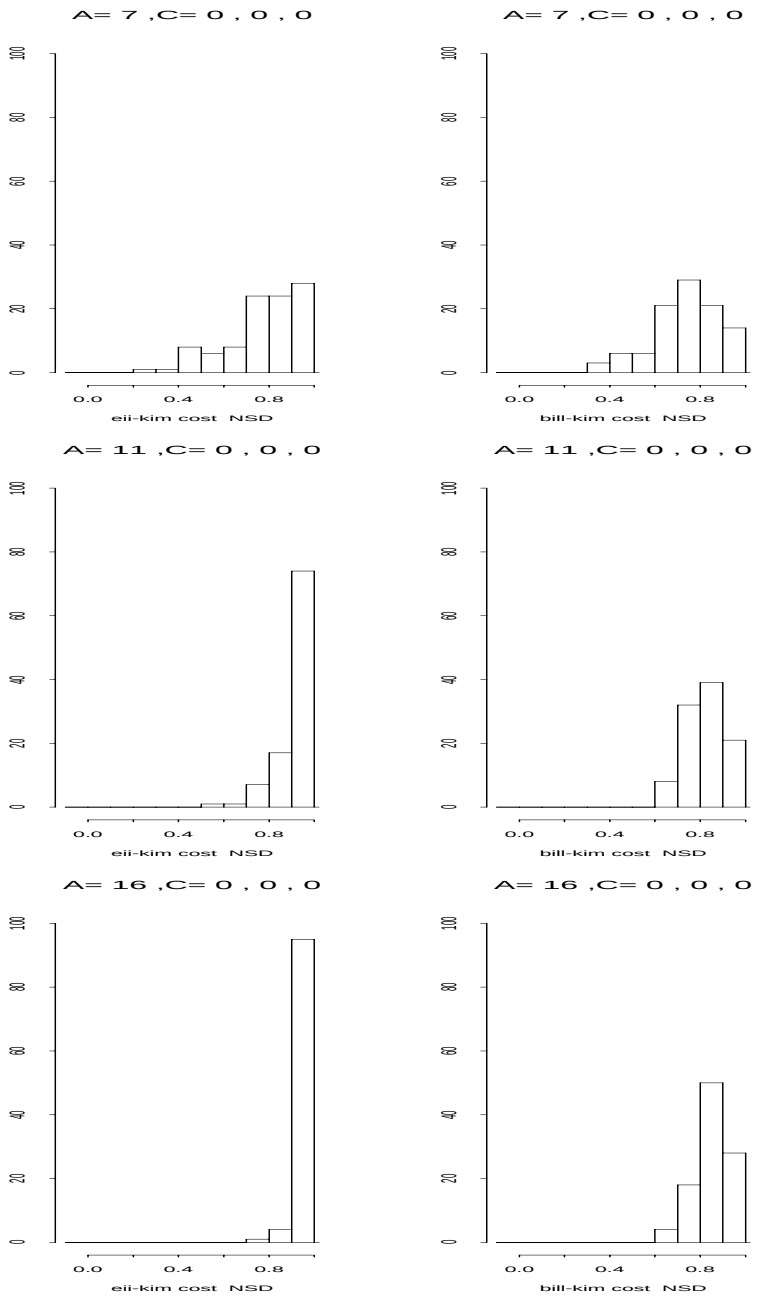


Figure 5.15: Comparing the Score Distributions for Dialogues between an Explicit-Acceptance Agent and an All-Implicit agent (EII-KIM) and two All-Implicit Agents (BILL-KIM), for AWM of 7, 11, and 16, for Standard Task, for commcost = 0, infcost = 0, retcost = 0

The two-sample test is concerned with the amount of agreement between two cumulative distributions. If the two samples have in fact been drawn from the same population distribution, then the cumulative distributions of both samples are expected to be close to each other. If the two sample cumulative distributions are too far apart at any point, this suggests that the samples come

from different populations. Thus a large enough deviation between the two sample cumulative distributions is evidence that they are not from the same population.

In terms of Design-World simulations, this means that if the distributions between the cumulative distributions of two samples are different 'enough' for a particular AWM setting, task definition, cost setting and strategy, then we can assume that agents using different strategies perform differently under those parameters. For example, consider a strategy to be introduced in chapter 6 of always explicitly indicating both ACCEPTANCE and REJECTION of a proposal. This is the Explicit-Acceptance strategy.

Figure 5.15 shows the distributions for dialogues in which one agent uses the Explicit-Acceptance strategy as compared with dialogues in which both agents are All-Implicit. Histograms are given for AWM of 7, 11 and 16 under the assumption that all processing costs are free. The KS test will show that the differences in the distributions at AWM of 7 are significant at the $p < .05$ level, and at AWM of 11 and 16 at the $p < .01$ level.

To apply the KS test we normalize performance measures to values between 0 and 1 and make a cumulative frequency distribution for each sample of observations, using the same intervals for both distributions. For each interval then, we subtract one step function from the other. The test focuses on the largest of these observed deviations.

Let $S_{n_1}(X)$ = the observed cumulative step function of one of the samples, that is $S_{n_1}(X) = K/n_1$ where K = the number of scores equal to or less than X. Let $S_{n_2}(X)$ = the observed cumulative step function of the other sample, that is $S_{n_2}(X) = K/n_2$. The KS test focuses on the maximum absolute value of the differences:

$$D = \text{maximum} \mid S_{n_1}(X) - S_{n_2}(X) \mid$$

Critical values for D can be calculated from the sample size and the level of significance desired. (see Table M in Siegel (1956).) Here sample size is 100 for both samples. If $D > .19$ then the difference is significant at the $p < .05$ level and if $D > .23$ then the difference is significant at the $p < .01$ level.

Now we are in a position to specify when a strategy is BENEFICIAL:

> A strategy A is BENEFICIAL as compared to a strategy B, in the same task situation, with the same cost settings, if the difference in distributions using the Kolmogorov-Smirnov two sample test is significant at $p < .05$, in the positive direction, for two or more AWM settings.

The converse of BENEFICIAL is DETRIMENTAL:

> A strategy A is DETRIMENTAL as compared to a strategy B, in the same task situation, with the same cost settings, if the difference in distributions using the Kolmogorov-Smirnov two sample test is significant at $p < .05$, in the negative direction, for two or more AWM settings.

Strategies need not be either BENEFICIAL or DETRIMENTAL, there may be no difference between two strategies. Also with the definition given above a strategy may be both BENEFICIAL and DETRIMENTAL depending on the range of AWM that the two settings are selected from. This is because the KS test compares distributions for each AWM setting at a time, but whether a strategy is BENEFICIAL is defined over all the AWM settings.

The next section shows how different assumptions about costs and the task definition determines whether or not a strategy is beneficial.

5.10.3 Effects of Task Definition and Changes in Costs

Task	Costs	AWM 3	AWM 7	AWM 11
Standard	1,1,.0001	107	277	340
Standard	10,1,0	4	170	255
Zero Invalids	0,0,0	110	306	424
Zero NonMatching Beliefs	0,0,0	3	8	108

Figure 5.16: Sample means of 100 runs for each evaluation function for All-Implicit agents at 3 different AWM settings

One way of seeing how the evaluation function costs and task definitions change the scores for a particular strategy, is to examine the performance means. In figure 5.16, performance means are shown for different AWM settings and different assumptions about costs. In the first row, COMPOSITE-COST is run with COMMCOST and INFCOST set to 1 and RETCOST to .0001, while in the second row COMMCOST is set at 10, INFCOST at 1, and RETCOST at 0. In the nonstandard task situations, we assume that all processing is free.

Each task definition and cost combination produce different results for different strategies and allow us to explore which situational parameters determine whether or not a communication strategy is BENEFICIAL. Remember that when using the Standard task, invalid steps in the plan are simply substracted out the RAW SCORE, whereas in the Zero-Invalids task, agents that make mistakes get

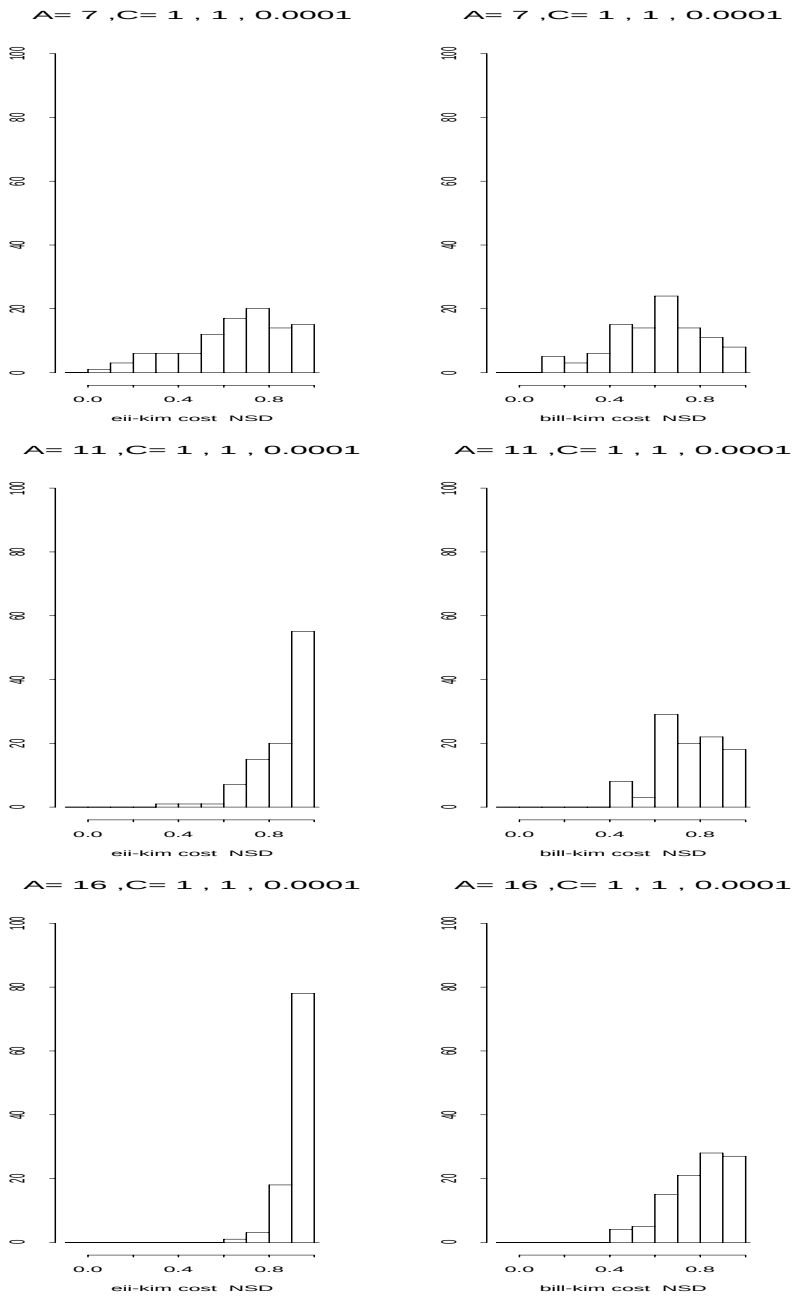


Figure 5.17: Comparing the Performance Distributions for Dialogues between an Explicit-Acceptance Agent and an All-Implicit agent (EII-KIM) and two All-Implicit Agents (BILL-KIM), for AWM of 7, 11, and 16, for Standard Task, for commcost = 1, infcost = 1, retcost = .0001, Differences are significant at AWM of 11 and 16

a RAW SCORE of 0. In the Zero-Nonmatching-Beliefs task, agents must agree both on the actions to be performed as well as the reasons for performing these actions.

As shown in figure 5.16, the performance means give a good indication of how agents perform with

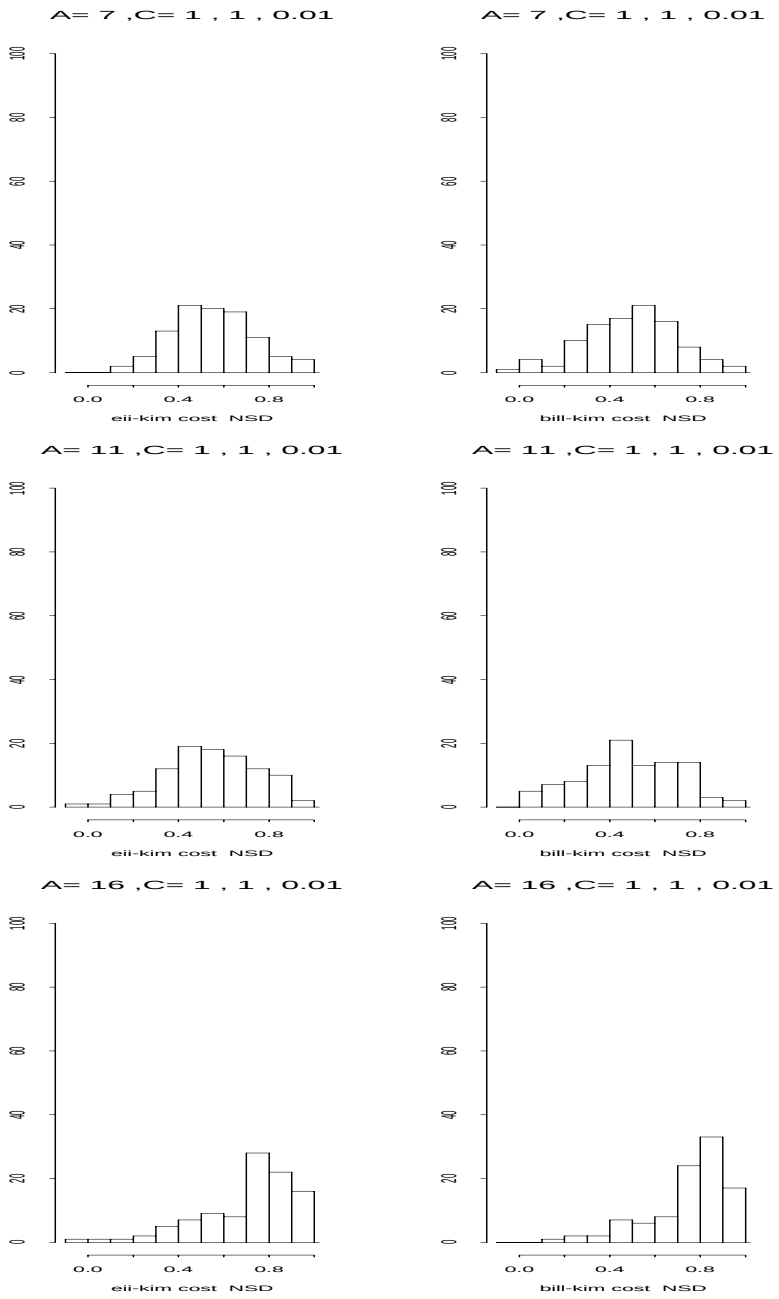


Figure 5.18: Comparing the Performance Distributions for Dialogues between an Explicit-Acceptance Agent and an All-Implicit agent (EII-KIM) and two All-Implicit Agents (BILL-KIM), for AWM of 7, 11, and 16, for Standard Task, for commcost = 1, infcost = 1, retcost = .01, No Signficant Differences

a particular strategy. They show that performance is highly situation specific. For example, the All-Implicit agents' strategy does well at avoiding invalid steps (row 3), does reasonably well when

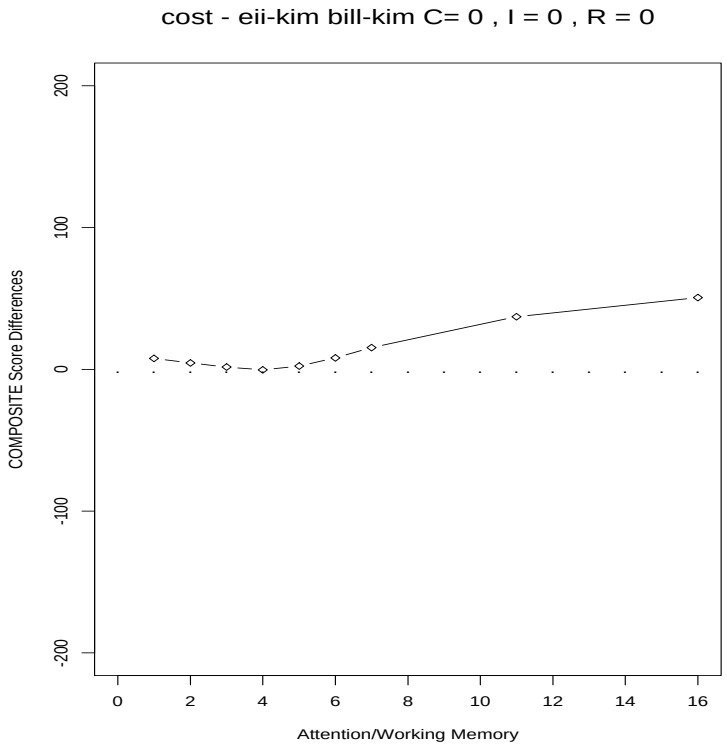


Figure 5.19: Sample Difference Plots: Differences using KS at 7,11,16 are significant

retrieval is not too expensive (row 1), and does rather less well when matching beliefs for warrants must be achieved (row 4).

Another way of depicting differences in performance is the performance distribution plots such as those presented in figure 5.15. Let's examine how communication costs alone effect changes in distributions. Keeping strategy differences constant, note that figure 5.15 showed performance distributions when all processing costs were assumed to be free. Figure 5.17 shows the performance distributions for the situation in which commcost is 1, infcost is 1 and retcost is .0001. Figure 5.18 shows the performance distributions for the situation in which commcost is 1, infcost is 1 and retcost is .01. The differences in the distributions are significant for AWM of 7,11, and 16 when processing is free (figure 5.15). When retcost is at .0001, the distributions are flattened slightly but we still get significant differences at AWM of 11 and 16. As we increase the cost of retrieval, the differences in distributions is flattened so that when retcost is .01, there are no significant differences in performance for any of the AWM values shown (figure 5.18).

5.10.4 Difference Plots: Comparing Performance

We have seen two ways of examining performance: looking at means for particular parameter/task settings and looking at performance distributions. In what follows, I will use a third way of visualizing performance. Differences in performance will be shown with **difference plots**. A difference plot compares two strategies: strategy 1 and strategy 2. A sample difference plot is shown in figure 5.19. The idea of difference plots is to give an overview of the differences between two strategies at all AWM settings without depicting histograms. If strategy 1 is better than strategy 2, the curve plotted will be above the 0 x-axis line in the plot. If strategy 2 is better than strategy, then the curve is below the plotted line. The actual distributions that correspond to the difference plot in figure 5.19 are also shown in the Performance distributions in figure 5.15 for AWM of 7, 11, and 16.

The motivation for difference plots is that, while presenting histograms of distributions doesn't lose any information, presenting these distributions in the remainder of the thesis may be overly complex. For example, presenting performance distributions requires visual comparisons between two sets of histograms such as those given in figures 5.15, 5.17, and 5.18. In other words, 18 performance distributions are needed to present the same information encapsulated in a difference plot such as that shown in figure 5.19.

The problem is that there is no guaranteed correspondence between the differences in means presented in difference plots and the differences in distributions used by the KS test to calculate statistical significance for performance differences. However, in practice here, since the evaluation of performance is always over a bounded interval, in most cases when the distributions are significantly different, the means are also look significantly different. However the converse is not always true. For example, the differences shown in figure 5.20 for AWM of 7,11, and 16 are **not significant**.

Despite these issues, it seems that presenting difference plots does not seriously misrepresent the results. In each case, KS is used to calculate significance, for each AWM setting, for each strategy and the significance values will be presented with each plot.

5.11 Related Work

Design-World was inspired by the TileWorld simulation environment: a rapidly changing robot world in which an artificial agent attempts to optimize reasoning and and planning (Bratman, Israel, and Pollack, 1988; Pollack and Ringuette, 1990). TileWorld was designed to test a theory of resource-bounded reasoning and deliberation. Unlike Design-World, TileWorld is a single agent

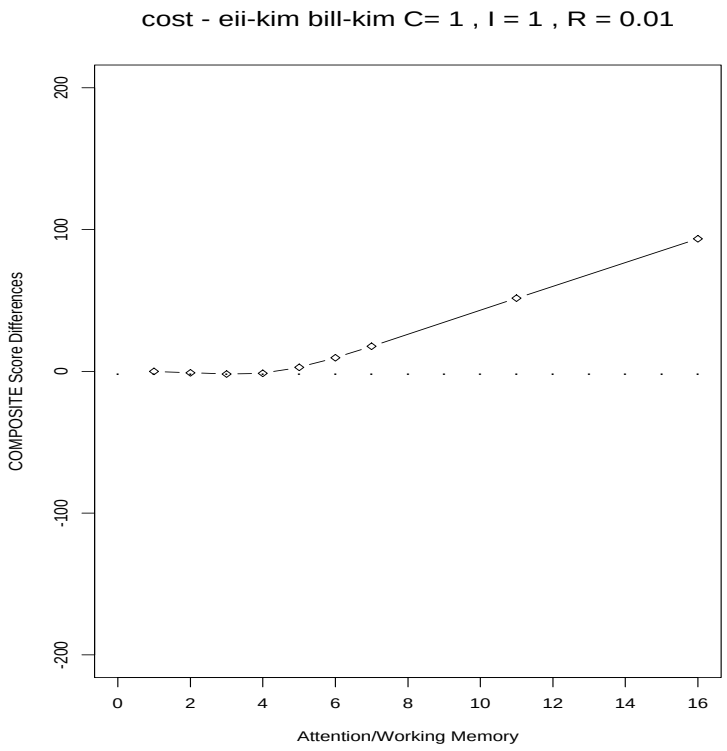


Figure 5.20: Visual Differences in Means are not always significant: Differences using KS at 7,11,16 here are NOT SIGNIFICANT

world in which the agent interacts with its environment, rather than with another agent. Design-World uses similar methods to test a theory of the effect of resource limits on communicative behavior between two agetns.

Design-World is also similar to the method used in the Edinburgh Dialogue Programs (Power, 1974; Houghton and Isard, 1985; Carletta, 1992). This method is distinct from other computationally-oriented research on discourse by attempting to provide a theory and a testbed for the roles of both the speaker and the hearer in a dialogic interaction rather than for one role alone. Each of these Dialogue Programs tests out different aspects of a theory of dialogue. For example, Power developed a computational simulation environment for dialogues between two artificial agents, conceptualized as robots, who must converse in order to be able to carry out actions to achieve their goals in an artificial world. Carletta's work focuses on variations in dialogue strategies between agents who may be 'risk-takers', and methods for agents to recover from the errors engendered by taking such risks. Carletta's methods are the closest to those presented here and will be discussed in more detail below.

The benefits of the Edinburgh method are that a number of hypotheses can be tested by a simulation that would be difficult to test in other ways. The method also forces the operationalization of 'goodness of fit' between the discourse strategies used by the speaker and the interpretation strategies used by the hearer in an extended dialogue. Furthermore extensions of the Edinburgh method developed in this thesis provide empirical support for claims about processing which are extremely difficult to judge either through introspection or by distributional analysis.[13]

Design-World is most similar to Carletta's JAM simulation for the Edinburgh Map-Task (Carletta, 1992). Carletta's simulation is based on the Map-Task Dialogue corpus, where the goal of the task is for the planning agent, the instructor, to instruct the reactive agent, the instructee, how to get from one place to another on the map. The instructor has a map with a number of landmarks and a route on it. The instructee must draw the route on his version of the map but his map's landmarks may not exactly match that of the instructor. This defines one difference between Design-World and the JAM system. Design-World simulates mixed-initiative dialogues, while JAM simulates an expert-apprentice type dialogue and all of the Map-Task dialogues are single initiative except for the clarification requests of instructees.

Carletta's simulation parametrizes agents according to their communication strategies and recovery strategies. The simulation explores many issues that are relevant for this thesis, including 'high risk' dialogue strategies, and methods of recovering from these strategies when failure occurs. Agents can take risks with respect to whether or not they make the context explicit, what Carletta calls 'context articulation'. This is similar to the explicit Attention strategies to be discussed in chapter 7. However Carletta notes that, in the Map-Task, context articulation doesn't serve any function because the context is always obvious.

A more critical 'high risk' strategy involves definite references to locations that may not appear on the instructee's map, e.g. *the big river*. The 'low risk' version of achieving the same goal might involve asking the instructee whether or not they have *a big river* on their map, before producing an instruction that refers to it. High risk strategies produce failures in communication that must be recovered from. The instructor then initiates one of several recovery strategies, using a notion of utility to select between them. The recovery strategy selected is based on calculations of the expected cost of repairing the current plan or replanning from scratch.

Based on her simulation Carletta argues that 'high risk' strategies are more efficient, where efficiency is a measure of the number of utterances in the dialogue. Here I have argued that the number of

[13]Processing claims can of course be partially judged by psycholinguistic experiments, which have the advantage of studying real humans, but the parameters that are investigated here cannot be varied as easily with real humans. Also, reaction time is a blunt instrument when it comes to providing support for a theory about the tradeoffs between inference, retrieval from memory, and producing and processing messages.

utterances is just one parameter for evaluating performance. In addition, if we wanted to have a more precise performance measure, there are problems with defining a performance measure for the Map-Task. The human subjects who carried out the task were not told that their performance would be measured by any particular criteria, and some of the subject pairs produced routes that barely approximated the one they were supposed to draw. Thus, the only measure of efficiency possible is the length of the dialogue or time to solution, but these measures ignore the quality of the solution. This means that the Map-Task dialogue corpus cannot be used to explore the efficiency trade-offs of different dialogue strategies.

In contrast, Design-World is set up so that there is a clear notion of maximizing the utility of the COLLABORATIVE PLAN. The fact that the value of an accepted plan is easily calculable means that Design-World provides an objective measure of the success or failure of various strategies.

5.12 Design-World Summary

This chapter discussed Design-World as one of two complementary empirical methods used in this thesis to support the theory of the function of IRUs. Design-World provides a method of testing the function of IRUs and claims about processing which are impossible to test with a distributional analysis. Design-World provides an objective performance measure for dialogue strategies, and ways to measure various cognitive costs. This allows us distinguish a strategy that makes it possible for an agent to do something as compared with one that makes it easier for an agent to do something.

In the following chapters, results from Design-World will be used along with the distributional analysis to argue for each communicative function of IRUs. Each communicative function has a number of associated strategies implemented in Design-World. Sections 5.3 and 5.6 described the Design-World domain and the task. Section 5.7 presented the discourse actions that agents use to communicate; these actions are the primitives out of which the communicative strategies are created.

Evaluating the performance of agents using various strategies depends on other parameters of the communication situation. Different task definitions were considered in section 5.9.4 and section 5.10 defined a performance measure COMPOSITE-COST that is used to evaluate the effectiveness of these strategies depending on the different costs as defined in the communication situation. Design-World results for Attitude strategies will be presented in chapter 6, for Consequence strategies in chapter 8 and for Attention strategies in chapter 7.

Chapter 6

Attitude

6.1 Introduction

The communicative function of ATTITUDE IRUs is to demonstrate the hearer's attitude to an utterance just contributed by a speaker to the discourse situation $\mathcal{S}$. Previous work has noted that whether a proposition is added to the discourse model does not follow automatically just from the fact that a speaker made an assertion; the hearer has the opportunity to reject the assertion (Stalnaker, 1978; Gazdar, 1979; Clark and Schaefer, 1989). The hearer's demonstration of attitude constrains whether and how the propositional content of the utterance and any inferences that follow from this content are added to the discourse model.

While speakers can display various attitudes to a proposition conveyed, such as understanding, acceptance, surprise, dismay or rejection, this chapter focuses on the less emotive attitudes central to problem-solving dialogues: UNDERSTANDING, ACCEPTANCE and REJECTION. Acceptance is not automatic, even when agents are cooperating. This is reflected in the ATTITUDE ASSUMPTION:

- ATTITUDE ASSUMPTION: Agents deliberate whether to ACCEPT or REJECT an assertion or proposal made by another agent in discourse.

The ATTITUDE ASSUMPTION means that agents have to work at achieving mutuality. This is represented by COORDINATION ASSUMPTION 1:

> COORDINATION ASSUMPTION 1:
> Achieving mutuality of beliefs and intentions is a coordination problem for conversants.

Furthermore, whether an utterance counts as an acceptance or rejection must often be inferred

because speakers don't make their propositional attitudes explicit by saying *I agree* or *I don't believe you*. Often the way listeners interpret which attitude is conveyed depends on the prosodic realization of an utterance. An example of an Attitude IRU, prosodically realized with a phrase final fall, is given in 51-27, where M repeats what H has said in 51-26.

(51) (24) H: That is correct. It could be moved around so that each of you have two thousand.
(25) M: I see.
(26) H: *Without penalty.*
(27) M: WITHOUT PENALTY.
(28) H: Right.
(29) M: And the fact that I have a, an account of my own from a couple of years ago, when I was working, doesn't affect this at all.

The repetition in 51-27 conveys the fact that M heard and UNDERSTOOD the verbatim content of what H said in 51-26 (Clark and Brennan, 1990). In addition, M's utterance conveys that she ACCEPTS H's assertion, both because it is realized with a final fall and because M passes up an opportunity to say anything more about what H has asserted (Schegloff, 1982; Whittaker and Stenton, 1988). Thus M has demonstrated her ATTITUDE to H's assertion.

Section 3.4 presented the framework for representing what is mutual called MUTUAL SUPPOSITION. The mutual supposition of UNDERSTANDING is treated separately from ACCEPTANCE and REJECTION. Section 6.2 examines the intention to achieve UNDERSTANDING and then section 6.3 extends the treatment of understanding to support the inference of ACCEPTANCE. Previous work has assumed that rejection must be conveyed by denial or logical contradiction. In section 6.4 I will argue that neither denial or contradiction are **necessary** to communicate rejection and that certain types of IRUs may convey rejection. Section 6.5 builds on the analyses in the earlier sections to explain how agents can achieve COLLABORATIVE PLANS.

The theory presented here is supported by two empirical bases: a distributional analysis of Attitude IRUs and Design-World experiments that vary how Attitude is demonstrated. Aspects of the distributional analysis that support the theory are presented throughout. The Design-World simulations, presented in section 6.6, show that Attitude IRUs make interaction more robust and help agents avoid making mistakes even when communication is not uncertain.

6.2 Mutual Understanding

MUTUAL UNDERSTANDING is the mutual supposition that an utterance has been **understood as**

intended and is usually inferred rather than communicated directly. Since the minimal purpose of any dialogue is that an utterance be understood, this goal is a prerequisite to achieving other goals such as another agent believing the proposition conveyed, or committing to a future action described by that proposition.

According to the SHARED ENVIRONMENT model an utterance event U in a discourse situation $\mathcal{S}$ licenses the inference of certain mutual suppositions, depending on what U INDICATES in $\mathcal{S}$. To formalize the INDICATES relation, it is useful to augment the representation of utterance events with two additional constructs: ASSUMPTIONS and ENDORSEMENTS. The endorsement types were discussed in section 3.3. Assumptions are beliefs that support a MUTUAL SUPPOSITION.

Let Δ be a function on utterances that represents the set of DEFEASIBLE ASSUMPTIONS associated with each utterance event. The INDICATES function will be represented as $leadsto_{\Delta}$ where Δ gives the set of associated assumptions. For utterance events, Δ always includes a set of assumptions about UNDERSTANDING (Clark and Schaefer, 1989; Walker, 1992b):

1. ATTENTION: the addressee B is attending to U;
2. COMPLETE HEARING: the addressee hears U correctly;
3. REALIZE: the addressee believes that U, said in discourse situation $\mathcal{S}$, realizes a proposition Φ, and that the speaker intended to convey Φ .[1]
4. LICENSE: the addressee believes that U, said in discourse situation $\mathcal{S}$, realizes a proposition Φ and that the speaker intended the addressee to infer Ψ, which follows from Φ as an entailment or by non-logical inference.

The REALIZE and LICENSE assumptions reflect the interpretation process of conversants as they attempt to determine what the utterance U INDICATES in a discourse situation $\mathcal{S}$. For convenience I will define two functions associated with these assumptions. The function $\mathcal{R}$: ($\mathcal{S}$,U) $\rightarrow \mathcal{P}$, returns the proposition that the speaker of an utterance U in discourse situation $\mathcal{S}$ intends to realize, and which the addressee must identify. The function $\mathcal{L}$: $(\mathcal{S}, \mathrm{U}) \rightarrow \mathcal{P}$, returns the proposition licensed as an inference by an utterance U in a discourse situation $\mathcal{S}$, which again the addressee must identify. The realize function $\mathcal{R}$ and license function $\mathcal{L}$ are relativized to both the conversants through the utterance event U, and to the discourse situation $\mathcal{S}$. This captures the fact that what an utterance means can depend on shared assumptions between two conversants as well as properties of the discourse situation.

[1]Of course the addressee may believe that U in $\mathcal{S}$ realizes some other proposition besides the one that the speaker intended to convey. The realization assumption represents the fact that Φ is not conveyed directly (Reddy, 1979; Schegloff, 1990; Brennan, 1990).

The defeasibility of the assumptions in Δ(U) is represented by associating an ENDORSEMENT with each assumption from the set of endorsement types discussed in section 3.3. All of the assumptions start out with an endorsement of HYPOTHESIS. One of the roles of Attitude IRUs, discussed in more detail in the next section, is that of upgrading the endorsement for a mutual supposition. To illustrate here how Attitude IRUs can affect endorsements, consider example 52.

(52) A: The number is 427 899.
B: 427 899.

B's repetition in 52b means that the COMPLETE HEARING assumption now has an endorsement type of LINGUISTIC. After B's utterance, B would be inconsistent to assert later in the dialogue that *I didn't hear what the number was.* Without B's utterance, a later assertion of this type would be acceptable; it would simply defeat the default assumption that B did hear as long as no other beliefs contradicted this later assertion.

The UNDERSTANDING INFERENCE RULE, UIR, represents the fact that inferences about what $\mathcal{S}$ indicates depend upon the assumptions above. This rule represents how the mutual suppositions are updated based on agents' behavior in discourse. Let A and B represent arbitrary members of a population of conversants P, i an arbitrary element of sequential indices I, σ an arbitrary member of the set of sentences Σ.

UNDERSTANDING INFERENCE RULE:
An utterance event U = (A, B, i, σ) $\sim\!\!>_{\Delta}$ MS(P, understand(B, $\mathcal{R}(\mathcal{S}$,U))

This rule says that given an utterance event U it is possible to derive a mutual supposition that the addressee understood the content of the utterance. The rule is based on the fact that an utterance U demonstrates a public hypothesis that it will achieve its purpose. The fact that things do not always go as planned is represented by the assumptions Δ(U) defined for each utterance event. Each assumption $\delta \in \Delta$(U), associated with U, underlying the inference of mutual supposition, is initially endorsed as a HYPOTHESIS as shown below:[2]

[2]This lets all members of P start supporting the mutual supposition simultaneously (Halpern and Moses, 1985), but this mutual supposition is very defeasible.

Δ(U)	Endorsement
attend (B, U)	hypothesis
hear (B, U)	hypothesis
realize(B, U, $\mathcal{R}(\mathcal{S}$, U))	hypothesis

If the speaker intends Ψ to be inferred, where Ψ is an inference from Φ, the content of U, the license assumption is included in Δ(U).

Δ(U)	Endorsement
license(B, U, $\mathcal{L}(\mathcal{S}$, U))	entailment $\vee$ default

Since the addressee is never in a position to know for sure what the speaker's intent is, s/he must determine what the likely intended inferences are.

Utterance U_{i+1} operates directly on the assumptions $\Delta(U_i)$ underlying utterance U_i. For example, since the assumptions associated with the UIR are very defeasible when they are only hypotheses, subsequent events can easily defeat these assumptions, and if an assumption is defeated the mutual supposition is defeated. Subsequent events can also **upgrade** the strength of these assumptions. The effect of U_{i+1} on the assumptions underlying U_i depends on showing that conventionally the next utterance position is the ATTITUDE LOCUS. Evidence that this is so will be provided in section 6.5.3.

Both the content and the form of U_{i+1} determine whether the assumptions $\Delta(U_i)$ are defeated or upgraded. For example, the complete hearing assumption of U_i can be defeated via conventional means by a $\Sigma(U_{i+1})$ of *What?*. This adds ($\neg$ hear (B, U_i)) to the discourse model, with an endorsement of linguistic, which defeats the hypothesis endorsement on the complete hearing assumption, and thus defeats the mutual supposition of understanding.

If a mutual supposition is defeated, its negation becomes mutually supposed, and the endorsement on the negated belief is based on whatever kind of evidence defeated the positive supposition. For example, each of not attending, not hearing, not knowing what the utterance realizes, or not knowing what the utterance licenses, supports the mutual supposition of not understanding.[3]

[3]The distributional analysis shows that the repair strategy of the speaker in response to the defeat of a mutual

Prosodic realization of an Attitude IRU is critical to interpretation: a phrase final rise on an Attitude IRU, ('question intonation'), can defeat the REALIZE and/or the LICENSE assumption, depending on whether the IRU is a repetition, a paraphrase, or makes an inference explicit.[4] Attitude IRUs with phrase final falls serve a different function: that of **upgrading** the endorsements on the assumptions underlying the inference of mutual understanding(Walker, 1993c). I will call this subset of Attitude IRUs, which communicate understanding and which may license the inference of acceptance, Attitude Acceptance IRUs. These will be discussed in the next section.

6.2.1 Attitude Acceptance IRUs

This section discusses the effects of U_{i+1} on the attention, hearing, realize and license assumptions underlying the inference of mutual understanding of the content of U_i. Each type of Attitude Acceptance IRU, the assumption addressed and the endorsement provided is shown in figure 6.1. The inferences about what is mutually supposed, licensed by the addressee's following utterance, are CONVERSATIONAL DEFAULTS (Joshi, Webber, and Weischedel, 1986). They are implicatures that arise from norms of interaction and have the diagnostic properties of implicatures of being both reinforceable and defeasible (Sadock, 1978).

NEXT Utterance Type	ASSUMPTION ADDRESSED	ENDORSEMENT TYPE
REPETITION	attention, hearing	linguistic
PARAPHRASE	attention, hearing, realize	linguistic
ENTAILMENT	attention, hearing, realize, license	linguistic
IMPLICATURE	attention, hearing, realize, license	linguistic
ANY Next Utterance	attention, hearing, realize, license	default

Figure 6.1: How the Addressee's Following utterance upgrades the endorsement for assumptions underlying the inference of mutual understanding

supposition is different in each case. For the attend and hear assumptions, speakers repeat their utterance. For the realize assumption, speakers paraphrase their utterance, and for the license assumption, speakers paraphrase or make inferences explicit. Thus speakers' repair strategies are directed toward failed assumptions as would be expected.

[4]This function of phrase final rises is probably specific both to the sequential position of the utterance, the fact that it is an IRU, and possibly other aspects of the discourse situation codified by the HG corpus. Other research has demonstrated a more general function for rises, that of soliciting a response of some kind from the hearer (McLemore, 1991; McLemore, 1992).

As detailed in figure 6.1, each HEARER OLD type, e.g. repetitions, paraphrases and making entailments explicit, provides a linguistic endorsement on the ATTENTION assumption because some response was generated by the next speaker. Each type also has specific distinct properties in the way it operates on the discourse model. The specific contribution of each of these types will be discussed in the remainder of this section.

Note here that **any** next utterance U_{i+1} by the addressee upgrades the endorsements on each of $\Delta(U_i)$ to default. See Figure 6.1. The basis for these conversational defaults will be discussed in section 6.3.

6.2.1.1 Repetition

A repetition demonstrates complete hearing of the verbatim content of what was said (Clark and Brennan, 1990). Consider the example below:

(53) (6) r. ....does that income from the certificate of deposit rule her out as a dependent?
(7) H: Yes *it does*
(8) r. IT DOES.
(9) H: Yup, that knocks her out. .....

The caller (r), in 53-8, repeats H's assertion from 53-7 with a phrase final fall. This repetition upgrades the endorsement on the HEARING and ATTENTION assumptions associated with 53-7 from hypothesis to linguistic. The REALIZATION assumption is upgraded to DEFAULT since (r) repeated what he heard, but chose not to provide evidence that he actually understood what H meant. The effect on $\Delta(U_7)$ is as follows:

$\Delta(U_7)$	ENDORSEMENT
attend(r, U_7)	linguistic
hear(r, U_7)	linguistic
realize(r, U_7, $\mathcal{R}(\mathcal{S}, U_7)$)	default

Because of the WEAKEST LINK rule, the mutual supposition licensed by UIR is a default.

MS(P, understand(r, $\mathcal{R}(\mathcal{S}, U_7)$)) default

However, the attention and complete hearing assumptions are no longer defeasible by linguistic evidence. Thus it would be contradictory for (r), in subsequent conversation, to say *Oh I didn't hear you say that*, or *I thought you said that it doesn't rule her out as a dependent.*

6.2.1.2 Paraphrase

As shown in figure 6.1, paraphrases and making entailments explicit also upgrade the complete hearing assumption. A paraphrase demonstrates complete hearing by showing that the verbatim content has been semantically incorporated into the addressee's memory. Making an entailment explicit demonstrates complete hearing by showing that the verbatim content has been incorporated into the addressee's memory and that at least one inference has been performed on this content.

In addition, a paraphrase and making an entailment explicit provide a linguistic endorsement for the realize assumption, of what proposition the paraphraser believes the previous utterance realizes. Paraphrases do this by demonstrating that the content has been semantically incorporated into memory. Consider example 54:

(54) (18) H: I see. *Are there any other children beside your wife?*
(19) d. *No*
(20) H: YOUR WIFE IS AN ONLY CHILD.
(21) d. right. and uh wants to give her some security .....

Utterance 54-20 is said with a phrase final fall and modifies the assumptions associated with 54-19 as follows:

$\Delta(U_{19})$	ENDORSEMENT
attend(h, U_{19})	linguistic
hear(h,U_{19})	linguistic
realize(h, U_{19}, $\mathcal{R}(\mathcal{S}, U_{20})$)	linguistic

In other words, h has demonstrated that as far as he is concerned 54-19 in $\mathcal{S}$ realizes the same proposition as 54-20. The effect of this demonstration is to upgrade the realize assumption associated with 54-19. Because of the WEAKEST LINK rule, the UIR now licenses a mutual supposition of understanding with a linguistic endorsement since all of the underlying assumptions are endorsed

as linguistic. Thus a paraphrase provides excellent evidence that an agent actually understood what another agent meant.

MS(understand(h, $\mathcal{R}(\mathcal{S}$, U_{19}))) linguistic

6.2.1.3 Making Inferences Explicit

IRUs that make inferences explicit, in addition to their ability to address the assumptions that paraphrases do, provide evidence of what inferences the speaker of the IRU believes are licensed in $\mathcal{S}$. Explicit entailment IRUs in the ATTITUDE LOCUS provide evidence about what is licensed by the prior utterance. The LICENSE assumption represents both inferences such as entailments and inferences based on non-logical inference rules such as implicatures. Consider example 55, where h makes an inference explicit in 55-(17). This inference follows from modus tollens, the content of 54-15 and 54-16, and the content of other utterances in $\mathcal{S}$ that make it clear that the tax year under discussion here is 1981.

(55) (15) H: oh no. *IRA's were available as long as you are not a participant in an existing pension*
(16) j. Oh I see. well *I did work I do work for a company that has a pension*
(17) H: ahh. THEN YOU'RE NOT ELIGIBLE FOR EIGHTY ONE

The fact that the proposition realized by 55-17 was inferrable from $\mathcal{R}(\mathcal{S}$, $U_{16})$ is represented by the license assumption, which after 55-16 is endorsed as an ENTAILMENT as shown below:

license(h, U_{16}, $\mathcal{R}(\mathcal{S}$, U_{17})) entailment

However, 55-17 upgrades the endorsement on this assumption to LINGUISTIC.

license(h, U_{16}, $\mathcal{R}(\mathcal{S}$, U_{17})) linguistic

Making an inference explicit also upgrades the endorsements on the attention, complete hearing, and realize assumptions to linguistic. These assumptions are upgraded because making an inference from an utterance presupposes attending to it, hearing it, and understanding its propositional content.

6.2.2 Distribution of Types of Attitude IRUs

I presented a model above of how different types of IRUs defined by the HEARER OLD distributional parameters have different effects on the assumptions underlying the mutual supposition of understanding. All of these types of IRUs occur in the HG corpus with similar frequency as shown in figure 6.2.

	Repetitions	Paraphrases	Inferences
Attitude	54	15	24
Not Attitude	6	43	32

Figure 6.2: Distribution of Attitude vs. Not Attitude IRUs by Hearer Old parameters; Attitude = Adjacent and Other

However, if we compare Attitude IRUs to other IRUs, figure 6.2 shows that Attitude IRUs are more likely to be repetitions than paraphrases ($\chi^2 = 49.96, p < .001$, df = 1). They are also more likely to be repetitions than inferences ($\chi^2 = 29.25, p < .001$, df = 1). These repetitions require less effort than either paraphrases or making inferences explicit, but only provide evidence of hearing the verbatim content, rather than any deeper understanding.

While it would seem that speakers would repeat when they think only evidence of verbatim hearing is required, I have not discovered any distributional correlates of repetitions that predict when a speaker might choose to repeat rather than paraphrase or make an inference explicit.

It is possible that in the HG corpus, much of the information given is fairly straightforward and hearing is all that is required. On the other hand there are very few misunderstandings in the corpus which seem to be due to mis-hearing. As argued earlier, an alternate motivation would be that Repetition Attitude IRUs are selected in these situations as a very minimal response requiring little effort. They are presented as HEARER OLD information prosodically as well (Walker, 1993c), which should mean that are allocated a reduced amount of processing time due to their prosodic realization alone (Cutler, 1976; Cutler and Foss, 1977; Nooteboom and Kruyt, 1987; Terken and Nooteboom, 1987).

6.2.3 Summary: Understanding

This section has argued that Attitude IRUs in general provide evidence as to the addressee's attitude toward a proposition conveyed by a speaker in a discourse situation $\mathcal{S}$. I focused here on the inference of mutual understanding which is a prerequisite to other goals in discourse and which

can be assumed to always be a shared intention for any conversation. Attitude IRUs with phrase final falls can upgrade the assumptions Δ(U) associated with each utterance event U, making the inference of mutual understanding less defeasible in the context. Section 6.2.1 illustrated for each HEARER OLD category of IRU the assumptions it addresses and the effects on the discourse model.

In each case, the IRU addresses one or more assumptions that have to be made in order to infer that mutual understanding has actually been achieved. The assumption, rather than endorsed as hypothesis or default, is upgraded to an endorsement type of linguistic as a result of the IRU. The fact that different IRUs address different assumptions leads to the perception that some IRUs are better evidence for understanding than others, e.g. a paraphrase is stronger evidence of understanding than a repetition (Clark and Schaefer, 1989). Section 6.3 will discuss the extension of the account given here to the inference of mutual acceptance. Section 6.5.3 will present distributional correlates for attitude IRUs.

Attitude Acceptance IRUs may simultaneously achieve other effects in addition to the upgrade function focused on here. For example, Attitude IRUs can be a rehearsal strategy for the speaker. Consider a situation in which the caller repeats the phone number that the information operator just gave him. This IRU is an effective way to remember the phone number.

Attitude IRUs can also maintain a proposition as salient, i.e. simultaneously function as an Attention IRU. This intention is made more plausible by the fact that in this discourse situation Repetition Attitude IRUs are common, but complete hearing doesn't seem to be a problem. It is plausible that Attitude IRUs are used conventionally or that they are intended to merely reflect the speaker's attentional state, since it is surprising that speakers would provide evidence of complete hearing when there is little chance that complete hearing wasn't achieved.

Finally, in previous work, we argued that a speaker who produces an Attitude IRU conveys that s/he has no reason to want to take control of the dialogue and the current speaker should continue (Whittaker and Stenton, 1988; Walker and Whittaker, 1990). This is a paraphrase of aspects of the theory presented here: one reason a conversant might want to take control is because they don't understand the content of the utterance. Other reasons will be discussed in the following section.

For particular instances of Attitude IRUs it is not clear which, if any, intention is primary, and attempting to make this distinction in corpus analysis is problematic. The Design-World experiments discussed in section 6.6 will however provide some insight on these multiple effects. One effect that will be tested is whether Attitude IRUs can have a rehearsal benefit, even in a situation in which there is no uncertainty as to whether an utterance is heard, understood or accepted.

6.3 Acceptance

In section 6.1, I distinguished the attitudes of understanding, acceptance and rejection. Section 6.2 presented an analysis of how the mutual supposition of understanding is inferred. This section will extend that analysis to the inference of acceptance.

In computational models of discourse, the inference of acceptance is usually accomplished via 'helpful' or 'cooperative' axioms. These axioms are based on the simplifying assumption that helpful agents will adopt other agents' beliefs and intentions (Cohen and Levesque, 1990; Litman and Allen, 1990). However a problem is that in many situations these assumptions may not be warranted.

Another version of the 'helpful' assumption is Grosz and Sidner's conversational default rule CDR2 in which the inference of acceptance depends on whether or not the addressee previously believed ¬P (Grosz and Sidner, 1990; Perrault, 1990). It isn't clear whether CDR2 is used by both agents in planning utterances, but if so, it is hard to see how it helps an agent engaged in a dialogue. If an agent thinks that another agent previously believed that P, then there is no point in making an assertion that P.[5] On the other hand, if an agent knows that another agent previously believed that ¬P, then there can be no point in making the assertion, since by this assumption the other agent will not change his/her beliefs. This assumption is of course useful in precisely those situations where an agent can be sure that another agent has no previous beliefs about P, but it is not clear how often agents are actually in this position. Examination of naturally-occurring dialogues in this work and elsewhere shows the even in situations where the addressee has asked a question as to whether P, s/he often has beliefs that would support P or ¬P (Pollack, Hirschberg, and Webber, 1982; Whittaker and Stenton, 1988; Walker and Whittaker, 1990).

The account presented here departs from previous work in dropping these 'helpful' and 'cooperative' simplifying assumptions and providing another way to make inferences about acceptance and rejection. It cannot be assumed that the speaker A always knows enough about B's mental state to be able to predict whether or not B will accept A's assertion that P.[6] The key observation that supports dropping the 'cooperative' and 'helpful' assumptions is that agents' behavior provides observable evidence for whether or not the intended effect of an utterance has been achieved, and that speakers deliberately provide such evidence to indicate their level of understanding and acceptance. There are two separate issues here: (1) whether or not B accepts A's proposal; and (2) how acceptance and rejection is displayed or negotiated.

[5] Modulo motivations of Attention and Consequence that are proposed here and that will be discussed in chapters 8 and 7.

[6] This is distinct from that fact that A may want plan his/her utterances strategically to improve the chances of having them accepted. This will be discussed in more detail in chapter 8.

6.3.1 Extending the Understanding Inference Rule

In order to extend the treatment of understanding to acceptance, there needs to be an extension of the UIR for inferring the mutual supposition of acceptance of a proposal or assertion in dialogue. Let A and B represent arbitrary members of a population of conversants P, i an arbitrary element of I, σ an arbitrary member of Σ.

ACCEPTANCE INFERENCE RULE (AIR)
An utterance event U = (A, B, i, σ) $\sim_{\Delta}$ MS(P, accept(B, $\mathcal{R}(\mathcal{S}$, U)))

The difference between the UIR and the AIR are the underlying assumptions Δ(U). For the AIR, acceptance is one of the assumptions $\delta \in \Delta$(U) associated with an utterance event U as shown below:

Δ(U)	ENDORSEMENT
attend (B, U)	hypothesis
hear (B, U)	hypothesis
realize(B, U, $\mathcal{R}(\mathcal{S}$, U))	hypothesis
license(B, U, $\mathcal{L}(\mathcal{S}$, U))	entailment $\vee$ default
accept (B, U, $\mathcal{R}(\mathcal{S}$, U))	hypothesis

The inference of acceptance is licensed by the response of the hearer B, in a way that is similar to the inference of mutual understanding. Whether the conversants mutually suppose that the AIR is in effect depends on the communication situation. Mutual awareness or agreement about the communiation situation is critical because acceptance is rarely conventionally conveyed by an utterance such as *I believe you* or *I agree*. To complicate matters, a form such as *uh huh*, that strictly provides only evidence of attention may support the inference of acceptance. I will argue that the conditions under which the AIR is used is specific to certain dialogue situations where agents are attempting to construct a COLLABORATIVE PLAN. Before considering these arguments, I will first examine evidence from the distributional analysis that shows that Attitude IRUs are prosodically marked. Then I will briefly discuss how rejection is conveyed before returning to the inference of acceptance.

6.3.2 Attitude Acceptance IRUs are neither Assertions nor Questions

I've argued that the primary role of Attitude IRUs is to display evidence of understanding and acceptance. Further evidence for this account is that Attitude IRUs are prosodically distinct from assertive utterances. We briefly discussed the fact that Attitude IRUs which indicate a lack of understanding are realized with Rises. Here I will focus on the fact that Attitude IRUs realized with falls do not sound like assertions.

First, note that in figure 6.5, Falls to Mid are significantly more likely to occur on Attitude IRUs than on IRUs that are not in the ATTITUDE LOCUS, ($\chi^2 = 5.695, p < .02, df = 1$). The frequency of phrase final Mids shown in figure 6.3 supports the view that Attitude IRUs are marked as hearer old and salient information(Prince, 1981b; Prince, 1992). Mids mark turn-medial utterances, non-completion, and hearer old or predictable information. Since Attitude IRUs are repetitions, paraphrases or inferences from propositions already asserted and currently salient, they unarguably consist of old and predictable information.

The ramifications of this for processing is that when a conversant hears an Attitude IRU, s/he can distinguish it from an assertive utterance. In addition, the fact that the utterance is marked as hearer old information means that the utterance requires less processing effort (Cutler and Foss, 1977).

This use of Mid has theoretical ramifications for a theory of intonational meaning. Final Mids are a distinguishing characteristic of the Warning/Calling contour (Pike, 1945; Liberman, 1975; Ladd, 1980; McLemore, 1992). However, warnings are often repetitions of old information. Calling someone isn't normally informative either; often merely the sound of the speaker's voice carries all the information. Furthermore, both noncompletion and nonfinality can be treated as special cases of non-assertion. The prevalence of Mid in all these environments supports the hypothesis that Mids mark non-assertion.

Boundary T	Rise	Mid	Low
Attitude (93)	24	28	15
Not Attitude (84)	3	22	32

Figure 6.3: Attitude IRUs have final Mid more frequently than Other IRUs

It is not clear however what determines whether an Attitude IRU is realized with a final Low or a final Mid. One possibility is that the final value reflects discourse structure: lower phrase final tones will be more common at the 'perceived' end of a discourse segment, whereas Mids, indicating

continuation(McLemore, 1991), will occur throughout the rest of a discourse segment. Testing this hypothesis must be left to future work.

6.4 Rejection

Previous work has assumed that rejection must be conveyed via denial or logical contradiction (Allwood, 1992; Gazdar, 1979; Hamblin, 1971). However, an Attitude IRU realized with a phrase-final rise (question intonation) can also indicate non-acceptance. This is encapsulated in the FINAL RISE HYPOTHESIS:[7]

> FINAL RISE HYPOTHESIS: Phrase final rises on Attitude IRUs defeat the realize and license assumptions, and thereby defeat the mutual supposition of understanding.

The FINAL RISE HYPOTHESIS corresponds to earlier characterizations of Attitude IRUs with phrase final Rises as querying 'the whole or some part of the previous utterance of another speaker, often with a note of incredulity' (Cruttenden, 1986).[8] An example of an Attitude IRU, realized with a phrase final rise, which seems to 'query part of a previous speaker's assertion' is shown in 56:

(56) (38) H: And I'd like 15 thousand in a 2 and a half year certificate
(39) R: The full 15 in a 2 and a half?
(40) H: That's correct
(41) R: Gee, not at my age.

In 56-39, R queries whether she actually heard and understood what H said. Formally, 56-39 defeats the REALIZE and LICENSE assumptions for 38, and 40 is a reassertion. The ATTITUDE LOCUS for 40 is 41, and R explicitly rejects H's proposal in 56-41.

In addition, even if we restrict our attention to assertive utterances, there is an additional class of rejection IRUs that shows that logical inconsistency is not **necessary** for REJECTION. Consider 57 from (Levinson, 1979).

(57) U_1: There's a man in the garage.

U_2: There's something in the garage.

[7] This final rise hypothesis is specific to IRUs in the ATTITUDE LOCUS. See McLemore (1991) for an analysis of phrase final rises on utterances in general as 'connecting speakers' or 'connecting turns'.

[8] Attitude IRUs realized with a Rise Fall are characterized as 'echo exclamations'.

In 57, U_2 is an entailment of U_1 via existential generalization, yet U_2 REJECTS U_1. That the IRU in U_2 can reject U_1 is surprising. How can a logically consistent assertion function to reject another assertion? I will argue that the basis for this type of rejection is a QUANTITY implicature (Grice, 1967; Horn, 1972; Gazdar, 1979), and that the implicature depends on the FOCUS/OPEN PROPOSITION structure of U_1 and U_2 (Prince, 1986).

Consider the quantity implicature in 59, which arises from 58:

(58)U_1: Is the new student brilliant and imaginative?

U_2: He's imaginative.

(59) He's not brilliant.

In 58, U_1 introduces a question as to whether *a*, *the new student*, is both *brilliant* and *imaginative*. U_2 is 'less informative' than it could have been and so implicates the denial of *brilliant*(*a*) shown in 59. I argue that 57 can be analyzed similarly. First, note that the implicature shown in 59 still arises in the context of the assertion in 60:

(60)U_1: The new student is brilliant and imaginative.

U_2: He's imaginative.

Thus the implicature is not dependent on the question context given in 58. Since implicatures only arise when they are consistent with the context and 59 is not consistent with 60:U_1, 60:U_1 cannot have been added to the context as an assertion before the utterance of 60:U_2. This supports the analysis of mutual acceptance presented above in which U_1 is added to the context as an HYPOTHESIS until after U_2. This means that it can be **cancelled** by an implicature in U_2. The implicature in U_2 has an endorsement of DEFAULT which defeats an hypothesis.

However one issue remains. As we saw in section 6.2, IRUs in the ATTITUDE LOCUS often indicate acceptance rather than rejection. Clearly a less informative U_2 following an assertion U_1 need not REJECT U_1. The difference between communicating acceptance and rejection with an IRU in the attitude locus depends on the EXCLUSION OF FOCUS CONDITION in 61:

(61) EXCLUSION OF FOCUS CONDITION: If an utterance U_2 by a speaker B asserts an alternate instantiation of the salient open proposition contributed to the context by an utterance U_1 as uttered by a speaker A, and U_2 excludes the focus of U_1, then U_2 REJECTS U_1.

The salient open proposition in 61 is the p-skeleton of U_1 (Prince, 1986; Rooth, 1985). In 57:U_1 the focus is *a man* whereas in 57:U_2, the focus is *something*. In 58:U_1 and 60:U_1 the focus includes *brilliant and imaginative*, whereas in 58:U_2 and 60:U_2, the focus is only *imaginative*. In each case, U_2 REJECTS U_1 because it meets the EXCLUSION OF FOCUS CONDITION. As further support for the EXCLUSION OF FOCUS CONDITION consider the difference in foci of the naturally occurring example 62 and its alternate 62':

(62)U_1: We bought these pajamas in New Orleans for me.

(62')U_1: We bought me these pajamas in New Orleans.

(63)U_2: We bought these pajamas in New Orleans.

The focus was *for me* in 62 whereas in 62' focus is most natural on *New Orleans*. 63 can reject 62 by excluding its focus but is infelicitous as a rejection of 62'. Felicitous rejections meet the EXCLUSION OF FOCUS CONDITION.

In sum, IRUs can indicate rejection showing that rejection need not be conveyed by denial or contradiction. The basis for this rejection is a quantity implicature dependent on focal structure. The fact that this class of rejection exists provides support for the account of mutual understanding and acceptance presented above.

6.5 Making Plans: Mutual Suppositions about Beliefs and Intentions

6.5.1 Collaborative Planning Principles

A major constraint on the inference of acceptance is that the conversants must believe that acceptance is **required** to achieve the overall purpose of the dialogue, i.e. the mutual intentions as understood at the initiation of the discourse constrain whether any utterance licenses the inference of mutual acceptance. The mutual supposition that acceptance is necessary then licenses the inference of acceptance via the operation of a simple principle of cooperative dialogue:

> COLLABORATIVE PRINCIPLE: Conversants must provide evidence of a detected discrepancy in belief as soon as possible.

The COLLABORATIVE PRINCIPLE is a simplification of the COLLABORATIVE PLANNING PRINCIPLES of Whittaker and Stenton (1988) and Walker and Whittaker (1990). It is an instance of a general

rule incumbent on a cooperative conversant to prevent his/her conversational partner from making false inferences (Joshi, 1982; Joshi, Webber, and Weischedel, 1986).

- Collaborative Planning Principles
  - Information quality: The listener must believe that the information that the speaker has provided is true, unambiguous and relevant to the mutual goal. This corresponds to the two rules: (A1) Truth: If the listener believes a fact P and believes that fact to be relevant and either believes that the speaker believes not P or that the speaker does not know P then interrupt; (A2)Ambiguity: If the listener believes that the speaker's assertion is relevant but ambiguous then interrupt.
  - Plan quality: The listener must believe that the action proposed by the speaker is a part of an adequate plan to achieve the mutual goal and the action must also be comprehensible to the listener. The two rules to express this are: (B1)Effectiveness: If the listener believes P and either believes that P presents an obstacle to the proposed plan or believes that P is part of the proposed plan that has already been satisfied, then interrupt; (B2) Ambiguity: If the listener believes that an assertion about the proposed plan is ambiguous, then interrupt.

The collaborative planning principles distinguish utterances about beliefs, assertions, from utterance about intentions, proposals, and describe under what conditions a conversant should demonstrate non-acceptance in order to keep default inferences about acceptance or understanding from going through. An interrupt can be realized by a question, which indicates non-acceptance, or by a rejection. Thus these principles provide specific guidelines to agents that determines their language behavior. While they do not say what an agent should say in an interruption, the specified reasons for interrupting can be used to provide the content of an interruption. For example the plan quality effectiveness clause suggests that the content of the listener's interruption should be (reject P), where P provides the reason for rejection.

The generalization presented in the collaborative principle is that evidence of any discrepancy in belief should be made apparent. Furthermore, the convention in the dialogues in the financial advice corpus is that this evidence is given in the attitude locus. I will discuss this convention further in section 6.5.3.

Figure 6.5.1 shows how the collaborative principle licenses the infernce of acceptance. The inference of acceptance is a conversational implicature, i.e. a default. The fact that it is a conversational implicature can be seen from the standard diagnostics; it is reinforceable, cancelable, nondetachable and calculable (Sadock, 1978; Horn, 1989; Hirschberg, 1985). The inferences shown

in figure 6.5.1 follow if (1) it is assumed that the collaborative principle is in operation in this discourse situation; and (2) if the hearer provides no evidence of rejection. Once the addressee has had an opportunity to reject an assertion or proposal, and has passed up this opportunity, the mutual supposition of acceptance of $\mathcal{R}(\mathcal{S}, \mathrm{U})$ is licensed as a default.

NEXT Utterance Type	ASSUMPTION ADDRESSED	ENDORSEMENT TYPE
ANY Non-Rejection	acceptance	default
Next Utterance	attention, hearing, realize, license	default

Figure 6.4: Inferences from the Collaborative Principle

6.5.2 Collaborative Plans

If we focus on proposals we can use the COLLABORATIVE PRINCIPLE to show how agents construct a COLLABORATIVE PLAN, i.e. the mutual suppositions about intentions that agents must achieve to have the beliefs necessary to carry out a collaborative task (Schelling, 1960; Lewis, 1969; Clark and Carlson, 1982; Power, 1984; Grosz and Sidner, 1990).

Agents can be motivated to achieve a COLLABORATIVE PLAN whenever they, for whatever reason, have the same goals and think that it is in their own interest to collaborate. One way agents can formulate collaborative plans is through dialogue; agents assert facts to other agents and if the other agents ACCEPT these facts then they can be part of a COLLABORATIVE PLAN. Agents make proposals to other agents about intentions, and if other agents ACCEPT these proposals, then they can be part of a COLLABORATIVE PLAN:

COLLABORATIVE-PLAN A&B Intention $\Longleftrightarrow$

1. MS A&B (Intend A $\vee$ B ($\alpha_1 \wedge \ldots \alpha_n$))
2. $\forall \alpha_i$ (MS A&B (Contribute α_i Intention)
3. MS A&B (Max-Utility ($\alpha_1 \wedge \ldots \alpha_n$) Intention)

The definition of COLLABORATIVE PLAN depends on the account of MUTUAL SUPPOSITION (MS), presented in chapter 3. An action α_i CONTRIBUTES to an intention if it is either a substep in a plan to achieve the intention or it enables achieving a substep (Pollack, 1986; Grosz and Sidner, 1986; Di Eugenio, 1993). Just as Pollack characterized plans as a set of beliefs and intentions (a complex

mental attitude) (Pollack, 1990), a COLLABORATIVE PLAN is simply a set of **mutual** suppositions about intentions and beliefs that are achieved through dialogue.

The Max-Utility constraint means that typically agents must have some mutual suppositions about what beliefs are serving as WARRANTS for their collaborative plan. Utility must be maximized over the combination of all the actions involved in the proposed plan rather than for each individual action. Chapter 5 discussed how agents achieve a collaborative plan in a simple world like Design World.

6.5.3 Evidence for the Attitude Locus

The COLLABORATIVE PRINCIPLE licenses the inference of acceptance by assuming that U_{i+1} is the ATTITUDE LOCUS for U_i. This means that U_{i+1} is the addressee's opportunity for demonstrating understanding, acceptance or rejection. We saw that the COLLABORATIVE PRINCIPLE embodies the assumption of the attitude locus, and thus will license default inferences about understanding and acceptance when no evidence of rejection is provided. This section uses facts from the distributional analysis to argue for the attitude locus.

The ATTITUDE LOCUS can be characterized by the information status parameters of Adjacent and Other. Remember that Adjacent and Other are SALIENCE parameters that mean that the IRU is adjacent to its antecedent and that the antecedent was said by the Other speaker. Any utterance U_{i+1}, which occurs adjacent to another utterance U_i, where U_i was said by the Other speaker is in the ATTITUDE LOCUS.

The argument that U_{i+1} is the ATTITUDE LOCUS for U_i depends on the final rise hypothesis presented in section 6.4. If the FINAL RISE HYPOTHESIS is true, IRUs with phrase final rises are an unambiguous demonstration of Attitude. Thus the location of phrase final rises can be used as an argument about where Attitude is conventionally demonstrated. An examination of the distribution of phrase final rises on IRUs provides support for the ATTITUDE LOCUS as shown in figure 6.5.

Phrase Final Intonation	Rise	Fall
Attitude	24	43
Not Attitude	3	54

Figure 6.5: Distribution of Boundary Tones on IRUs, Attitude vs. Not Attitude; Attitude = Adjacent and Other, Not Attitude = Not (Adjacent + Other), Fall = Low + Mid

Phrase final rises on IRUs are much more likely to occur in the hypothesized ATTITUDE LOCUS, characterized by the distributional parameters of Adjacent and Other ($\chi^2 = 16.88, p < .001, df = 1$). This provides evidence for the ATTITUDE LOCUS because the vast majority of utterances with this function occur in this locus.

Figure 6.5 also shows that there are three cases of IRUs realized with phrase final rises, which are not in the attitude locus. These exceptions provide further support for the ATTITUDE LOCUS because each of them is explicitly marked as being 'out of sequence'. For example, the utterance of *Oh hang in there my friend* marks two of them in 64:

(64) M: .... First of all, you know, ah there's an outfit here in Philadelphia that you know, that I put money in at a certain interest, and I can I can borrow on it at one percent more. And *it's a 30 month deal*, so I have not received any interest on it. You know I don't have to show any interest on it *because they have not given me you know any 1099's or anything*, but..
H: Oh hang in there my friend
M: I'm hanging in there.
H: YOU HAVE A 30 MONTH CERTIFICATE?
M: Right.
H: AND THEY HAVE NOT SENT YOU A 1099?

One question associated with the ATTITUDE LOCUS is what happens when the next utterance is an interruption. For example, in 65, H interrupts C at 64-5 to ask for his name:

(65) (4) C: Ok harry, I'm have a problem that uh my - with today's economy my daughter is working.
(5) H: I missed your name.
(6) C: Hank.
(7) H: Go ahead hank
(8) C: as well as her uh husband
They have a child.
and they bring the child to us every day for babysitting.
This is while she works.
(9) H: um hm
(10) C: Now we're wondering how can we handle this to help them get a tax deduction.

At 65-8, C returns to his narrative, apparently picking up exactly where he left off. He co-specifies

his daughter with the possessive pronoun *her* in the phrase *her husband* and elliptically predicates with *as well as* to convey *My daughter's husband is working.* Walker and Whittaker (1990) used this example to argue that discourse has a hierarchical structure and that the sequence of utterances 65-5 ...65-7 is structurally embedded within 65-4 ...65-10. This explains why C can continue at 65-8 using ellipses and pronominal anaphora to refer to discourse referents realized further back in the discourse.

However, if H had not understood what it was that C was saying in 65, he would have had to make a choice between interrupting or asking C to clarify or repeat what C said. Thus it is plausible that 65-5 still demonstrates understanding of the previous utterance. The inference of mutual understanding is a default, and can be defeated by subsequent events. For example, at 65-7 H could say *Now what were you saying about your daughter?*

In sum, the distributional analysis provides good evidence for the ATTITUDE LOCUS and this in turn provides support for the application of the COLLABORATIVE PRINCIPLE to infer acceptance in the absence of evidence to the contrary.

6.5.4 Cognitive Basis for the Collaborative Principle

Earlier I noted that the ATTITUDE LOCUS, as the sequential position for displaying understanding, acceptance and rejection, has been characterized as a convention of conversation (Sacks, Schegloff, and Jefferson, 1974; Clark and Schaefer, 1989). Here I briefly present a few arguments suggesting that there are cognitive processing motivations for the site of ATTITUDE LOCUS.

First, psychological studies on the limits of attention/memory have found that the verbatim content of an utterance is retained for a very short period of time (Sachs, 1967; Bransford, Barclay, and Franks, 1972; Anderson, 1974). The Bransford etal. work also shows that conversants can't distinguish propositional content from some subset of simple inferences that are derived from that content. The content of an utterance is typically semantically integrated into memory very soon after the utterance is completed. Thus any action which attempts to operate on the verbatim content of an utterance is constrained to occur very soon after that utterance.

Second, at the point when an utterance is semantically integrated into memory, inferences based on the propositional content of that utterance may be made and stored in memory. Since inferences are based on salient beliefs, most inferences are made at this point.

However, studies in psychology on belief revision and decision making have shown that inferences based on facts that are untrue persist even when the original fact has been retracted (Tversky and Kahneman, 1982; Ross and Anderson, 1982). Thus if an agent misunderstands another agent and

then bases inferences on this misunderstanding, s/he may have trouble rectifying his/her belief state. Similarly if one agent infers (falsely) that another agent accepted her assertion, she may have trouble retracting all the beliefs that were inferred from that false inference. These observations on the human inference mechanism are incorporated in Harman's PRINCIPLE OF POSITIVE UNDERMINING, also discussed in section 5.1 (Harman, 1986; Galliers, 1990):

> PRINCIPLE OF POSITIVE UNDERMINING:
> Only stop believing a current belief if there are positive reasons to do so, and this does not include an absence of justification for that belief.

An operationalization of this in terms of a belief model means that each inferred belief, added as an inference from a false belief, must independently be explicitly challenged in order to be retracted. This shows that conversants should ensure that they have understood and been understood in the intended way **at the time** the propositional content of the utterance is initially added to the discourse model. Furthermore, if inferences are based on the acceptance of an utterance by an agent, agents will be more efficient if they are sure about what is accepted at the time the utterance is said. In other words, there is strong motivation for **local management** of potential misunderstandings and disagreements and for **signals** by conversants as to their beliefs about the current propositional content.[9]

6.6 Attitude in Design World

This section compares a strategy that communicates Attitude explicitly, the Explicit-Acceptance strategy, with the All-Implicit strategy. The strategies are compared for each task situation, and under different assumptions about processing costs as discussed in section 5.10.

Figure 6.6 shows the discourse action protocol for an Explicit-Acceptance agent. Openings and Closings are left implicit, but acceptance is always communicated explicitly. Figure 6.7 shows a dialogue sequence composed of the discourse actions shown in figure 6.6. This shows how a dialogue between two Explicit-Acceptance agents might go. In general, each discourse act can be composed of a number of utterance acts, however the Explicit-Acceptance agent realizes each discourse act with one utterance act. Figures 6.8 and 6.9 illustrate the relationshiop between the discourse acts and the cognitive processes for each agent that results in this type of dialogue sequence.

[9] I noted earlier that these signals are often IRUs, so that their content is not conventionally indicated, or presented as new information (Walker, 1993c). Thus these signals are not as explicit as they might be. However the way acceptance is signalled presumably is motivated by the multiple functions Attitude IRUs can achieve in terms of updating not only the Acceptance hypothesis but also the assumptions related to Understanding.

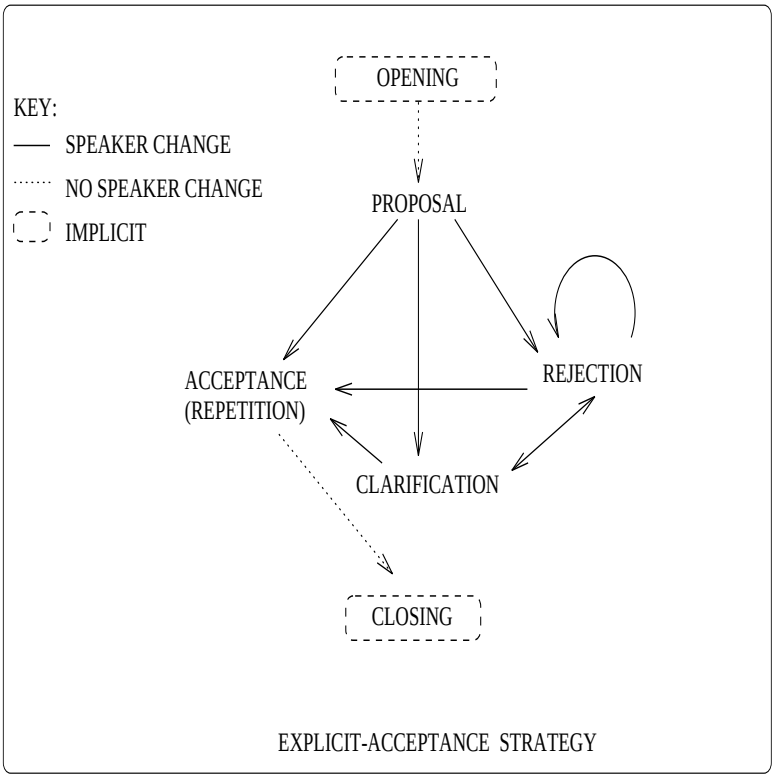


Figure 6.6: Discourse Action protocol for agents using the Explicit-Acceptance strategy

Agent A	Agent B
PROPOSAL 1	
	ACCEPTANCE 1
PROPOSAL 2	
	REJECTION 2,3
ACCEPTANCE 3	
...	...
	PROPOSAL N
ACCEPTANCE N	

Figure 6.7: Sequence of Discourse Acts for Dialogue in 66

The Explicit-Acceptance strategy implements just one of a number of possible variations in Attitude strategies; an agent parametrized with the Explicit-Acceptance strategy **repeats** each proposal as s/he accepts it. Other variations on the content of ACCEPT utterances could be backchannels such as *uh huh*, paraphrases, or making inferences explicit.

The Explicit-Acceptance strategy will be compared with the All-Implicit strategy that was discussed at length in section 5.8. The hypothesis that is tested using the Explicit-Acceptance strategy is that

Discourse Act	Utterance Act	A PROCESS	B PROCESS
		ME-REASON Put-3 Deliberate Propose, All-Imp	
PROPOSAL	4:(propose A B option-78)		
			ME-REASON Put-3 Check Preconds 78 Deliberate Decide to Accept Store (MS intend 78) Store (Act-Effects 78) Accept, Exp-Acc
ACCEPT-ANCE	5:(accept B A intended-78)		
		Store (MS intend 78) Store (MS intend 78) Store (Act-Effects 78) Store (Act-Effects 78) ME-REASON Put-4 Deliberate Propose, All-Imp	
PROPOSAL	6:(propose A B option-93)		
			Infer Close Put-3 Infer Open Put-4 ME-REASON Put-4 Check Preconds 93 Deliberate Decide to Reject Reject, Exp-Acc

Figure 6.8: PART1: Cognitive Processes for Sequence of Discourse Acts in Dialogue 66: Agent A is All-Implicit and Agent B is Explicit-Acceptance

Attitude repetitions increase robustness by providing explicit indications of acceptance. Due to the way that AWM is designed, the simulation primarily tests whether Attitude IRUs can function as a rehearsal of the repeated information.

6.6.1 The Explicit-Acceptance Strategy

Excerpt 66 demonstrates the Explicit-Acceptance strategy. In each dialogue excerpt, each message is preceded by a gloss of the content of the message. IRUs are shown in CAPS, and other messages are shown in *italics*. Remember that the agents actually communicate with messages in the artificial language as shown just after each gloss.

The dialogue below is for a dialogue between one Explicit-Acceptance agent and one All-Implicit

Discourse Act	Utterance Act	A PROCESS	B PROCESS
REJECTION	7:(propose B A option-97)	ME-REASON Put-4 Check Preconds 97 Deliberate Decide to Accept Store (MS intend 97) Store (Act-Effects 97 ME-REASON Put-5 Deliberate Propose, All-Imp	
PROPOSAL	8:(propose A B option-99)		Infer Acceptance 97 Store (MS intend 97) Store (Act-Effects 97) Infer Close Put-4 Infer Open Put-5 ME-REASON Put-5 Check Preconds 99 Deliberate Decide to Accept Store (MS intend 99) Store (Act-Effects 99) Accept, Exp-Acc
ACCEPT-ANCE	9:(accept B A intended-99)	Store (MS intend 99) Store (MS intend 99) Store (Act-Effects 99) Store (Act-Effects 99)	
	... .		

Figure 6.9: PART2: Cognitive Processes for Sequence of Discourse Acts in Dialogue 66: Agent A is All-Implicit and agent B is Explicit-Acceptance

agent. The agent, EII, who uses the Explicit-Acceptance strategy, indicates acceptance of a proposal explicitly and repeats each proposal as she accepts it. Agent KIM is an All-Implicit agent.[10]

(66) KIM: *Then, let's put the purple couch in the study.*

4:(propose agent-kim agent-eii option-78: put-act (agent-kim purple couch room-1))

EII: OKAY, LET'S PUT THE PURPLE COUCH IN THE STUDY

5:(accept agent-eii agent-kim intended-78: put-act (agent-kim purple couch room-1))

KIM: *Then, let's put the purple chair in the study.*

6:(propose agent-kim agent-eii option-93: put-act (agent-kim purple chair room-1))

EII: *No, instead let's put the fuchsia couch in the study.*

7:(reject agent-eii agent-kim option-97: put-act (agent-eii fuchsia couch room-1))

[10]EII is mnemonic for Explicit attitude, Implicit consequence and Implicit attention.

KIM: *Then, let's put the purple chair in the living room.*

8:(propose agent-kim agent-eii option-99: put-act (agent-kim purple chair room-2))

EII: OKAY, LET'S PUT THE PURPLE CHAIR IN THE LIVING ROOM.

9:(accept agent-eii agent-kim intended-99: put-act (agent-kim purple chair room-2)) .....

Each message consists of an utterance intention, the speaker, the hearer, and the content of the message which is an action or a belief. If the message is a proposal, then its content is an action. The status of this action is an OPTION that the agents might want to pursue. Once the proposal has been accepted, then the status of the action is an INTENTION.

For example, in Kim's proposal shown in 66-4, the put-act is an option that Kim has identified. In 66-5 EII explicitly communicates her acceptance of Kim's proposal. The acceptance of the proposal changes the status of the action under discussion from an option (option 78) to an intention (intended-78). The effect on AWM of each Repetition Attitude IRU shown here is to duplicate in memory the fact that the proposal has been made and accepted. In addition, the inference that follows from the acceptance of a proposal, that the agent no longer has the piece of furniture used in the proposal, is also duplicated in AWM.

EII doesn't always accept Kim's proposals. In 66-7, EII rejects Kim's proposal and makes a counter-proposal which Kim implicitly accepts in 66-8 by making a new proposal. EII must infer in this case that Kim has accepted her proposal, because no evidence to the contrary was provided. The dialogue continues until the Kim and EII have achieved a collaborative plan. The Appendix includes an example of a complete dialogue between EII and Kim. Figures 6.8 and 6.9 shows the relationship between discourse acts, utterance acts and cognitive processes for dialogue 66.

The Explicit-Acceptance strategy is contrasted with the All-Implicit strategy discussed in section 5.8. While the endorsement on acceptance is a default for the All-Implicit agent's implicit acceptance and linguistic for the explicit acceptance of the Explicit-Acceptance agent, as discussed in section 6.3, there is no possibility of misunderstanding or implicit rejection in this artificial situation.

Sections 6.6.2 and 6.6.3 discusses the differences between the two strategies that depend on parameters of the communication situation such as the communication cost, inference cost and retrieval cost. Section 6.6.4 discusses the differences between the two strategies that depend on how the task is defined.

6.6.2 Explicit-Acceptance produces Rehearsal Benefits

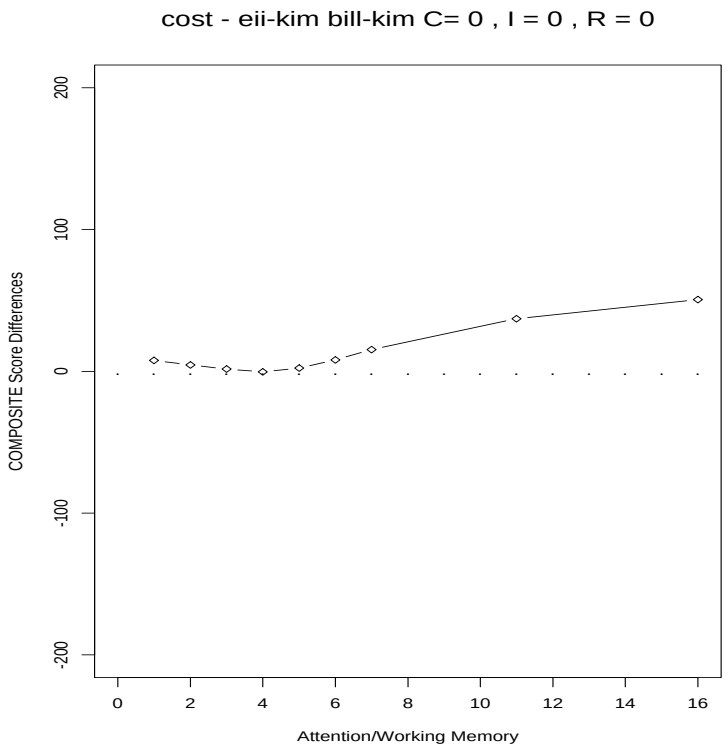


Figure 6.10: Explicit-Acceptance produces Rehearsal Benefits for AWM above 6. Strategy 1 is Explicit-Acceptance with All-Implicit and strategy 2 is two All-Implicit agents, Task Definition = Standard, commcost = 0, infcost = 0, retcost = 0.

Figure 6.10 shows that as AWM increases we begin to see some benefits for the Explicit-Acceptance strategy (KS $>$ .19 for AWM of 7,11,16, p $<$ .05). This is a rehearsal benefit of rehearsing that the agents have agreed on an action, which also has the effect that the effect of that action is duplicated in memory. This rehearsal benefit reduces the number of invalid steps in the collaborative plan and thus produces benefits at higher AWM. The reason that agents are more likely to make mistakes at higher AWM is that they are more likely to retrieve information that is inconsistent with the current state of the world.

If communication cost is 1, inference cost is 1 and retrieval cost is .0001, we still get a benefit for this strategy at AWM of 11 and 16. However if retrieval cost is increased to .01, there is no difference between this strategy and the All-Implicit strategy. Explicit-Acceptance increases the number of retrievals by displacing score propositions so that they take more effort to retrieve. Once the cost of retrieval is high enough, this cost dampens the benefits of the strategy. The next section will discuss situations in which communication cost dominates the other costs of the interaction.

6.6.3 Explicit-Acceptance can be detrimental if communication is expensive

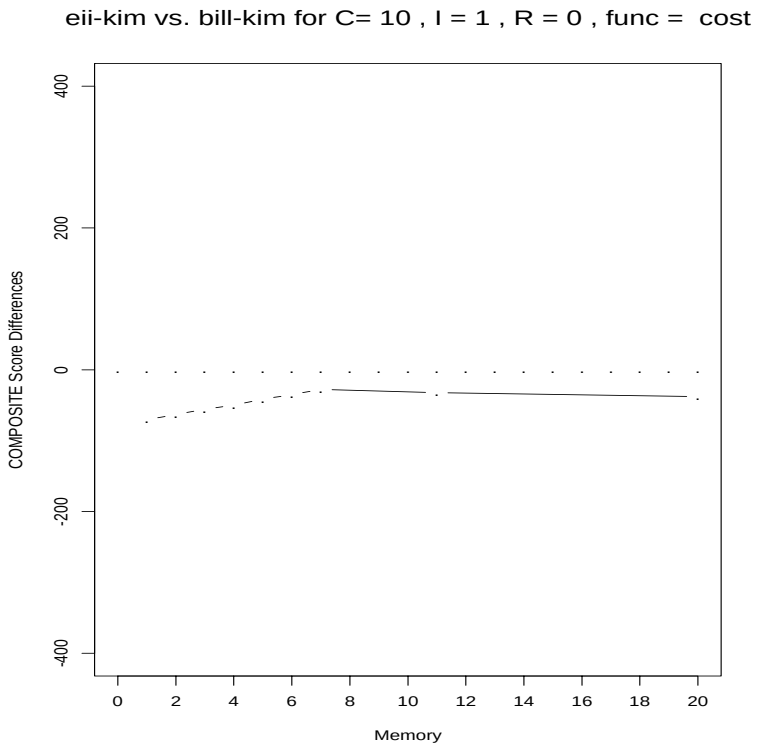


Figure 6.11: If communication is expensive Explicit-Acceptance is detrimental for AWM < 7. Strategy 1 is the combination of one Explicit-Acceptance agent with one All-Implicit agent and strategy 2 is two All-Implicit agents, Task Definition = Standard, commcost = 10, infcost = 1, retcost = 0.

As with all explicit strategies (ones that include IRUs), if communication cost is high relative to inference cost, performance is lower. This is because strategies that include IRUs always use more messages to communicate the same content. IRUs make inferences explicit or tell agents facts that they already know or that they could retrieve from memory given enough resources. Figure 6.11 shows that differences exist at low AWM values if communication cost is 10 times more expensive than inference cost and retrieval is free (KS $> .23$, p$< .01$ for AWM < 7). The next section examines the role of the task in determining when a strategy is beneficial.

6.6.4 Explicit-Acceptance prevents Errors

Section 5.10 discussed two different ways of defining the task that require a greater level of agreement or precision. These are the Zero-Nonmatching-Beliefs and the Zero-Invalid versions of the task. The Explicit-Acceptance strategy has no effect when the task definition is Zero-Nonmatching-Beliefs.

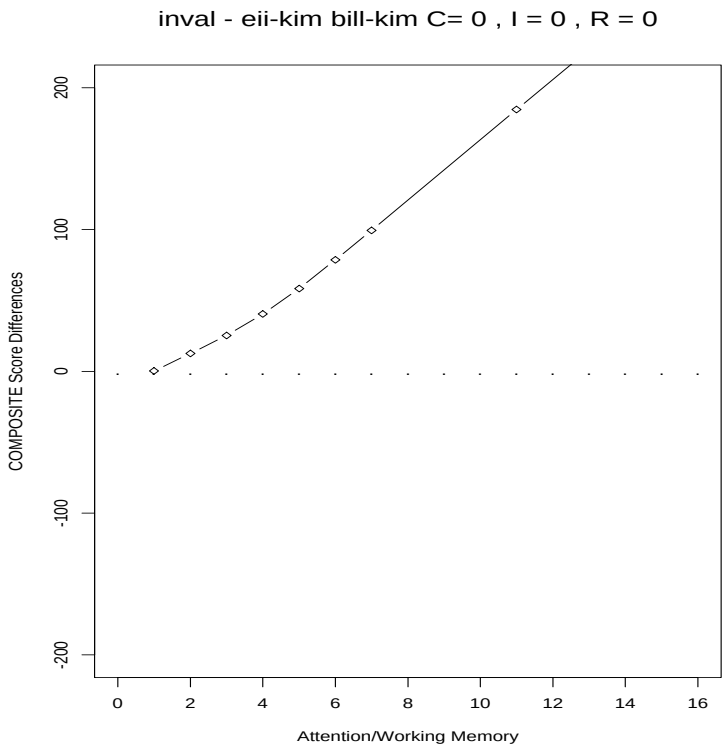


Figure 6.12: Explicit Acceptance is beneficial for the Zero-Invalid task for AWM > 5. Strategy 1 is the combination of one Explicit-Acceptance agent with one All-Implicit agent and strategy 2 is two All-Implicit agents, Task Definition = Zero-Invalid, commcost = 0, infcost = 0, retcost = 0.

However, the major benefit of the Explicit-Acceptance strategy within the limits of Design-World is in the Zero-Invalid version of the task, i.e. the Explicit-Acceptance strategy helps agents avoid making mistakes as shown in figure 6.12 (KS > .23 for AWM > 5, p < .01).

6.6.5 Summary: Attitude in Design-World

This section has demonstrated that the Explicit-Acceptance strategy does benefit agents, even when communication is not uncertain. This is because it improves the robustness of memory for what the agents have agreed: essentially functioning as a rehearsal strategy. I have also shown that the benefits of the strategy depend on other variables in the communication situation.

The differences in figures 6.10, 6.11 and 6.12 show that the benefits of a strategy depend on the task situation and the relative costs of communication, inference and retrieval, which can depend on the communication situation. If agents are penalized for invalid steps in their plans, then the Explicit-Acceptance strategy can help them avoid such mistakes.

6.7 Attitude:Summary

This chapter presents a theory of the function of Attitude IRUs in supporting mutual supposition. I present an account of the inference of mutual supposition of understanding and acceptance in which Attitude IRUs make these inferences more robust and less defeasible. I showed through the distributional analysis that Attitude IRUs can be characterized by the two SALIENCE parameters of Adjacent and Other. Furthermore, I showed that Attitude IRUs tend to be repetitions in this corpus.

The COLLABORATIVE PRINCIPLE redefines the notion of cooperativity in situations of coordinated action. In contrast to other accounts (Allen and Perrault, 1980; Cohen, 1978; Grosz and Sidner, 1990; Perrault, 1990; Litman and Allen, 1990), here cooperativity means that one must make conflicts or potential problems obvious to one's conversational partner. An agent does not need to accept whatever s/he is told, but the collaborative principle can license the inference of acceptance whenever no evidence to the contrary is provided in the ATTITUDE LOCUS (Heeman and Hirst, 1992).

A COLLABORATIVE PLAN is similar to, and in many respects compatible with, the SharedPlan formalism of Grosz and Sidner (Grosz and Sidner, 1990). However I don't assume that a collaborative plan is a primitive. The underlying account of mutual belief as mutual supposition (Prince, 1978; Nadathur and Joshi, 1983; Lewis, 1969; Walker, 1992b) also distinguishes this account from Grosz and Sidner's and the Joint Intention account of Cohen and Levesque (Cohen and Levesque, 1991). Another difference is the incorporation of deliberation via maximizing utility which reflects agents' **autonomy**. Even if agents have agreed to carry out a dialogue to do a collaborative task, they do not necessarily accept every proposal that another agent makes. Each agent deliberates about the utility of each proposal from his/her own perspective. In addition, in the framework presented here *okay* does not convey acceptance directly (Grosz and Sidner, 1990; Levesque, Cohen, and Nunes, 1990; Cohen and Levesque, 1991), because utterances such as *okay, fine, alright, yea* are simple variations on the backchannel *uh huh*, which does not communicate anything stronger than evidence of attention.

The account presented here on the relationship between understanding and acceptance predicts the fact that there is a systematic ambiguity as to whether a form communicates understanding or acceptance. While certain forms only convey attention, the fact that more wasn't said can support the inference of acceptance, if the conversants believe that acceptance is necessary for some shared intention. If the conversants don't believe that acceptance is necessary, then the inference of acceptance isn't licensed.

In addition, section 6.6 presented results from Design-World simulations that demonstrate that Attitude IRUs have clear benefits in improving robustness and avoiding errors in planning. These benefits have an associated expense of an increase in the number of messages, which does not decrease performance unless communication is much more expensive than other operations. In addition, redundant messages can cause agents to forget other facts that are relevant to the task, and thus in certain circumstances can be detrimental. Whether or not an Attitude explicit strategy is beneficial depends on parameters of the communication situation such as the cost of communication and the penalties for making errors or for not agreeing on each aspect of the task.

Chapter 7

Attention

7.1 Introduction

Attention IRUs manipulate the locus of attention of the discourse participants by making a proposition salient. A speaker can intend to make a proposition salient only if it is not currently salient, so Attention IRUs are defined as those IRUs whose antecedents are not currently salient.

Making a proposition salient is a vague functional characterization that can have many specific effects. According to the DISCOURSE INFERENCE CONSTRAINT, salience constrains which propositions are available for interpretation and reasoning.[1] The Attention IRU in 67b supports inferential processes.

(67) a. Clinton has to take a stand on abortion rights for poor women.

b. HE'S THE PRESIDENT.

This is a Deliberation IRU. Deliberation IRUs make a proposition salient in order to support deliberation about whether to accept or reject other currently salient propositions. For example, 67b is a reason why the hearer should accept the speaker's assertion in 67a.

Salient entities also activate or cue retrieval of semantically related propositions or entities (Ratcliff and McKoon, 1988; Collins and Quillian, 1969; Anderson and Bower, 1973). This could be the primary function of 68 repeated here from chapter 1:

[1] This is a type of domain restriction (Roberts, 1993b), but research on domain restriction has typically focused on restrictions on the set of discourse referents, $\mathcal{D}$, for domains of quantification or anaphora resolution. In this case, the domain of propositions, $\mathcal{P}$, available for inference is restricted.

(68) Frieda, YOU'RE A PSYCHOLOGIST. What do you think about this case of a ten year old boy kidnapping a two year old? (fg 4/14/93)

This IRU restricts the domain of propositions that are relevant to answering the question. It is **as a psychologist** that Frieda is asked; thus she should formulate her answer using facts relevant in psychology (Ratcliff and McKoon, 1988). This is an Open Segment IRU, which to my knowledge, have not been previously noted in the literature. An additional class of Attention IRUs, Close Segment IRUs, will be discussed in section 7.4.

In the rest of this section, I will examine the distributional parameters that define Attention IRUs and their distributional correlates. Then I will propose three potential discourse functions for Attention IRUs. In the remainder of the chapter, I will discuss each distributionally distinct class with respect to the hypothesized functions.

7.1.1 Distributional Correlates of Attention IRUs

The class of Attention IRUs is defined by a single parameter: the antecedent for the IRU is Remote, and thus no longer salient (see section 4.3). Figure 7.1 shows the distribution of Attention IRUs by HEARER OLD category as compared with other IRUs.

	Repetitions	Paraphrases	Inferences	Unused	Presuppositions
Attention (56)	6	41	7	4	0
Not Attention (150)	61	42	32	0	15

Figure 7.1: Distribution of Attention IRUs by Hearer Old Category

	Open Segment	Close Segment	Other
Attention (56)	10	15	33
Not Attention (150)	0	12	138

Figure 7.2: Open and Close Segment correlated with Attention

In figure 7.1 Inferences are the combination of Entailments and Implicatures. The small number of Attention Repetitions in figure 7.1 is due to the fact that speakers don't remember the verbatim

form of a proposition and that form is determined by context (Prince, 1985). Thus a proposition re-realized remotely is unlikely to be realized in the same form. The small number of Attention Inference IRUs is explained by the DISCOURSE INFERENCE CONSTRAINT. There is no difference in the distribution of paraphrases and only Attention IRUs can be Unused by definition. Thus none of the HEARER OLD distinctions determine differences in function for Attention IRUs.

One distributional characteristic of Attention IRUs is the tendency to occur at loci characterized as discourse segment boundaries by the discourse theories discussed in chapter 3 (Hobbs, 1979; Polanyi, 1987; Grosz and Sidner, 1986). In section 4.5, I discussed criteria for determining whether an IRU is at a segment boundary. This defines two subclasses of Attention IRUs: Open Segment IRUs and Close Segment IRUs. Figure 7.2 shows that Open and Close Segment IRUs are more likely to be Attention IRUs ($\chi^2 = 35.88, p < .001$, df $= 1$).

Thus we can use the distributional analysis to define three distributional classes: Open Segment, Close Segment and Deliberation, where Deliberation IRUs are those Attention IRUs that are not Open or Close Segment.

Neither distributional class nor the characterization of Attention IRUs as making a proposition salient define specific discourse related functions that Attention IRUs might achieve. These effects may rely on underlying cognitive processes that are not specific to the interpretation of IRUs, or they may be based on some conventional use of IRU. In section 7.1.2, I propose three plausible functions for these classes. Then in the following sections, each distributional class will be examined to determine which of the hypothesized functions the class supports.

Deliberation IRUs will be discussed in section 7.2. Open and Close Segment IRUs will be discussed in sections 7.3 and 7.4. Then in section 7.5 the evidence that each distributional class provides for the hypothesized discourse functions introduced in section 7.1.2 will be reviewed. Finally, section 7.6 will discuss Design-World experiments on discourse strategies that incorporate Deliberation and Open and Close Segment IRUs.

7.1.2 Discourse Functions of Attention IRUs

This section introduces three hypotheses about more specific discourse functions of Attention IRUs. Then in the following sections each distributional class identified above will be examined to see whether it can achieve one or more of the hypothesized functions below:[2]

Discourse Functions of Attention IRUs:

[2]As with other IRUs, it is also plausible that the IRU functions to rehearse the proposition that it realizes. The AWM model in Design-World always tests this hypothesis.

1. COORDINATION HYPOTHESIS: IRUs are a conventional means of signaling discourse structure. In particular IRUs can be used to indicate the opening and closing of discourse segments, and help the hearer infer where the current segment fits in the overall discourse structure.

2. DISCOURSE INFERENCE HYPOTHESIS: A proposition from a prior context is selected and realized in the current context because a process in the current context, such as inference or deliberation, requires that proposition to be salient.

3. RETRIEVAL CUE HYPOTHESIS: the IRU's main function is as a retrieval cue. The content of the IRU serves as a retrieval cue for the retrieval of propositions in a prior context.

The COORDINATION HYPOTHESIS is meant to tease apart some of the claims of previous work. One claim has been that the function of Close Segment IRUs is to explicitly signal discourse structure by marking the Closing of a discourse segment. In mixed-initiative dialogue, Close Segment IRUs can be characterized as a 'bid' by the speaker to close the segment (Schegloff and Sacks, 1977; Hobbs, 1979). The view implicit in the COORDINATION HYPOTHESIS is that hearers are trying to infer a hierarchical discourse structure and so speakers rely on particular conventions to indicate structure, as a way of helping hearers' inference processes. Thus the COORDINATION HYPOTHESIS means that Attention IRUs are like cue words and have a conventional rather than a cognitive function. Since this function relies on convention, the IRU must be recognized as redundant.

Another view of how conversants coordinate is possible, i.e. that conversants locally manage topic transitions and no higher level structure is inferred. For example, the Close Segment function of some IRUs does not rely on convention according to Whittaker and Stenton. The recognition that the utterance has 'no new information' is interpreted straightforwardly as an indication that the speaker has nothing new to say about the subject (Whittaker and Stenton, 1988; Walker and Whittaker, 1990). To avoid boredom, the conversation moves on.

The DISCOURSE INFERENCE HYPOTHESIS reflects the surface function since saying an utterance makes its content salient. According to the DISCOURSE INFERENCE CONSTRAINT, a proposition can only be used in inference or deliberation if it is currently salient. Recognizing the IRU as redundant is not necessary to achieve this function.

The RETRIEVAL CUE HYPOTHESIS can be contrasted with the DISCOURSE INFERENCE HYPOTHESIS because it formulates the role of the IRU as a pointer to a previous context. Rather than the IRU alone being functionally related to the current context, the IRU indicates that a whole prior context is functionally related to the current context. Unless this retrieval happens automatically, the IRU must be recognized as redundant to achieve this function.

The RETRIEVAL CUE HYPOTHESIS was also briefly discussed in chapter 3 where I discussed the relationship between theories of discourse structure and AWM. There I suggested that one way in which AWM could achieve the same effects as Grosz and Sidner's stack model of attentional state (Grosz, 1977; Sidner, 1979; Grosz and Sidner, 1986) is by formulating a POP to a prior context as a retrieval from that context. If this use of retrieval were developed in more detail, then the use of Open Segment IRUs for COORDINATION would be identical with their use as RETRIEVAL CUES.

The COORDINATION HYPOTHESIS, RETRIEVAL CUE HYPOTHESIS and DISCOURSE INFERENCE HYPOTHESIS need not be mutually exclusive. If both retrieval and inference follow fairly automatically from making a proposition salient, then it is possible that Attention IRUs perform all of these functions without being functionally ambiguous.

These hypothesized functions will be discussed in relation to the distributional classes given in the following sections. I will conclude that it is not possible to provide unequivocal support for a single one of these functions, but will show that the DISCOURSE INFERENCE HYPOTHESIS is viable for all three distributional classes. Because AWM is too simple to support the COORDINATION HYPOTHESIS this hypothesis cannot be tested in Design-World. However section 7.6 presents experiments that support the DISCOURSE INFERENCE HYPOTHESIS.

7.2 Deliberation IRUs

Deliberation IRUs are those Attention IRUs that are neither Open nor Close Segment IRUs. This does not mean that their discourse function might not overlap or be the same as Open and Close Segment IRUs. I will examine this class as distributionally defined, and then see whether it supports the DISCOURSE INFERENCE HYPOTHESIS, the RETRIEVAL CUE HYPOTHESIS or the COORDINATION HYPOTHESIS.

Deliberation IRUs are used in contexts evocative of argumentation or logical proof, and the IRU appears to be a premise that is re-evoked as part of laying out the structure of the argument (Levinson, 1979; Cohen, 1987; Sadock, 1978; Webber and Joshi, 1982). Thus Deliberation IRUs appear to support the DISCOURSE INFERENCE HYPOTHESIS and occur simply because of the DISCOURSE INFERENCE CONSTRAINT, i.e. both inference and belief revision operate on propositions that are currently salient.

These SUPPORT and WARRANT relations defined in section 3.3 are dependent on the type of the propositions being related and define two types of Deliberation IRUs: WARRANT IRUs and SUPPORT IRUs. In both cases, the point of the IRU can be characterized as providing a premise for the

inference of a relation between the IRU and another proposition salient in the current context. I will discuss WARRANT IRUs in section 7.2.1, and SUPPORT IRUs in section 7.2.2.

Both WARRANT and SUPPORT IRUs may support either the DISCOURSE INFERENCE HYPOTHESIS or the RETRIEVAL CUE HYPOTHESIS. I will leave this question aside while I discuss the examples, and then in section 7.2.3 I will argue for the DISCOURSE INFERENCE HYPOTHESIS. Furthermore, in order to argue against the RETRIEVAL CUE HYPOTHESIS, I will suggest that the structure of an argument is often based on ACCOMMODATION of an inference rule rather than retrieval (Lewis, 1979; Thomason, 1990).

7.2.1 Warrant IRUs

As an example of a WARRANT relation between two propositions, one describing an intention and the other a belief, consider the following excerpt that was part of a discussion about where to eat lunch:

(69) (1) Listen to Ramesh.
(2) HE'S INDIAN. (DH 11/5/91)

The point of the discussion was for the group to agree on an instantiation of X in the proposition *We should eat at restaurant X of type Indian.* The speaker intended the addressees' adoption of the intention to do γ in 69-1 to CONTRIBUTE to achieving this agreement. The IRU *Ramesh is Indian*, is a WARRANT for adopting the intention to do γ. In other words, the addressees were meant to infer that the proposition conveyed by 69-2 is a reason for adopting an intention to do γ. This example supports the DISCOURSE INFERENCE CONSTRAINT; even though the proposition conveyed by 69-2 is already HEARER OLD, the inference that 69-2 is a warrant for 69-1 depends on saying 69-2 to make that proposition salient.[3]

As another example of the WARRANT relation, consider the dialogue in 70. Here M tells H that she and her husband are both retired in 70-43, and a number of other facts about their financial situation in successive dialogue from 70-43 to 70-45.

(70) (42) H: Now what is your income situation
(43) M: *We're both retired* and our income for the year is about um 24 about 26 thousand
(44) H: Have you other securities than stock? have you any bonds or certificates?
(45) M: Yes yes we do. We have some certificates oh about uh 15 - 20 thousand, not much we're not rich - and we have a house completely paid for, have some land in the poconos, completely

[3] In the Design-World simulation discussed below, I will explore the difference between it being **necessary** to say 69-2 and it being **more efficient** for the speaker to say 69-2 than to expect his audience to retrieve the proposition conveyed by 69-2 from memory.

paid for - and uh that's actually the extent of our uh
(46) H: Ok. On the proceeds of that GM stock,
(47) M: yes
(48) H: I'd like to see you put that into two different southern utilities
.
(clarification of southern utilities)
.
(58) H: and those, I think that you will find, that that will give you a good return,
YOU ARE RETIRED and that is primarily what you are looking for, in addition I think that eventually those stocks will rise as well.....

Beginning with 70-46, H suggests a course of action. In 70-58 he paraphrases M's assertion from 70-43, *you are retired.* This paraphrase in this context leads to the inference of a warrant relation between the paraphrased proposition and the proposed course of action. In other words, H implies that *X is retired* warrants *X wants a good return on her investment.* More precisely let:

A = Invest proceeds of GM stock in 2 different southern utilities

P = You are retired

G = Get a good return on your investment

What M is supposed to infer is that: (1) P WARRANTS G. In fact H's statement *that is primarily what you are looking for* comes close to making this explicit. H asserts that A CONTRIBUTES to achieving G. Since M believes P, she may adopt G as a higher level intention, and then intend to do A as a good way of achieving G. The WARRANT relation is based on a different information relation than in 69: in this case the warrant proposition describes which aspect of the situation is a reason for M adopting a goal of a good return on her investment.

7.2.2 Support IRUs

The SUPPORT relation holds between two beliefs whenever believing one is a reason for believing the other as in 71:

(71) Owen knows who Percy Sledge is. HE'S AMERICAN.

To make sense of 71, an inference rule must be retrieved or accommodated that *All Americans know who Percy Sledge is.* Consider the variations on 71 in 72.

(72) a. You should know who Percy Sledge is. YOU'RE AMERICAN. (OR 3/4/93)

b. Owen will know who Percy Sledge is. HE'S AMERICAN.

In 72a, what is inferred is again that being American is **why** you should know, but additionally there is an inference that this is **why** the speaker makes the assertion. In the actual context in which this utterance was said, the speaker had presupposed that his audience knew who Percy Sledge was, and was surprised when questioned. While in each case a specific inference rule must be retrieved or accommodated, the exact character of the relation of this rule to deliberation varies depending on exactly what is asserted. Thus 72b can be easily seen to be part of reasoning about a future directed intention. For example, if a goal in the context is *find out who Percy Sledge is* and one of the conversants proposes the intention of *Ask Owen*, then the sequence in 72b is a reason why we might want to **ask** Owen, rather than someone else, who Percy Sledge is.

7.2.3 Accommodation, Inference or Retrieval

Sadock noted that *The appropriate contribution to a conversation is often not merely a statement, but one backed up by some evidence.* For example, the appropriate answer to a question such as 73a is not merely a yes or a no, but rather a reasoned argument such as 73b:

(73) a. Will Bilandic win?

b. Yes. He's the machine candidate.

The response in 73b relies on the hearer to retrieve or accommodate an inference rule such as *If X is the machine candidate, then X will win.* While Sadock doesn't note that this reasoned argument often includes IRUs, it seems plausible that 73b could have been known to both conversants.

A reasoned argument for adopting a belief or intention can involve any proof strategy or type of proof rule; the hearer must recognize what rule and strategy combination the speaker intends (Sadock, 1977; Cohen, 1981; Webber and Joshi, 1982; Cohen, 1987). For example consider the RULE OF INFERENCE FRAMES in figure 7.3 from (Cohen, 1987) (see also (Webber and Joshi, 1982)):

Sadock points out that any argument can occur in MODUS BREVIS form, that is only some of the premises need be made explicit. Figure 7.4 shows all the possible MODUS BREVIS forms for an argument based on MODUS PONENS.

Cohen states that the Missing Major form is the most popular form. The Missing Major form is the one that will be discussed in the remainder of this section. This form is also exemplified by 73 where the 73b is the Minor premise and the Major premise (the inference rule) must either be retrieved or accommodated; a similar process is operating in the examples in sections 7.2.1 and 7.2.2. According to Cohen, in order to recognize the argument structure the hearer must:

Inference Rule	Major Premise	Minor Premise	Conclusion
Modus Ponens	$P \rightarrow Q$	P	Q
Modus Tollens	$P \rightarrow Q$	$\neg Q$	$\neg P$
Modus Tollendo Ponens	$P \vee \neg Q$	Q	P
Modus Ponendo Tollens	$P \vee Q$	Q	$\neg P$

Figure 7.3: Rule of Inference Frames

Form	Given Premises	Conclusion
Normal	$P \rightarrow Q$, P	Q
Missing Minor	$P \rightarrow Q$	Q
Missing Major	P	Q
Only Major	$P \rightarrow Q$	(assume rest)
Only Minor	P	(assume rest)

Figure 7.4: Modus Brevis Forms of a Modus Ponens Argument

1. Identify the missing premises.
2. Verify the plausibility of these missing premises.

If the missing premises are plausible, then the hearer believes that s/he has correctly identified the structure of the speaker's argument. To carry out the steps above, Cohen proposes that the hearer can try the following strategies (in order):

- Identify the missing premise within a knowledge base of shared knowledge.
- Identify a 'relaxed version' of the missing premise within own private knowledge.
- Identify the missing premise within a model of the speaker's beliefs.
- Judge the beliefs of a hypothetical third party, which could be simplified as believe it unless there's reason to strongly doubt from within one's own beliefs.

The first three strategies rely on retrieval of an inference rule or a premise from various 'knowledge bases' in memory. Thus Cohen's formulation is related to the RETRIEVAL CUE HYPOTHESIS; the IRU (Minor premise) serves as a cue for retrieving the inference rule (Major premise).

In the case of examples such as 69, this would mean that each rule of the form *If a person is Indian, then they know which Indian restaurants are good*, *If a person is Chinese, then they know which Chinese restaurants are good*, etc. are stored explicitly in memory and retrieved as part of recognizing the argument structure. Because this seems implausible, this remainder of this section explores the possibility that a very general rule such as *Someone who knows about X should be consulted about X* is applied, and that the existence of these general rules supports the accommodation of more specific rules. Cohen's last strategy can be viewed as a type of accommodation. The remainder of this section argues against the RETRIEVAL CUE HYPOTHESIS as a function for Deliberation IRUs, by providing a sketch of how the inference rules might be accommodated rather than retrieved.

Assume simply that hearers generally deliberate about incoming assertions and attempt to determine SUPPORT or WARRANT relations between them. If the hearer believes only that the speaker is using **some** proof strategy in an argument, the hearer can use a suitably modified version of the logical rule of IMPLICATION INTRODUCTION given informally in 74 (Allwood, Andersson, and Dahl, 1977):[4]

(74) a. Assume: A

b. Assume: B

c. Therefore: C

d. Implication Introduction: If A and B then C

This rule allows the hearer to infer or accommodate inference rules, albeit with some missing premises. No retrieval of inference rules is needed.

Additional support for this idea is that all of the minor premises and the conclusion are salient in each example in the corpus. This is consistent with the DISCOURSE INFERENCE HYPOTHESIS. Furthermore Sadock and Cohen noted that Missing Major was the most common form of MODUS BREVIS, but usually the Conclusion is also available in the context (Sadock, 1978; Cohen, 1987). In 69, the inference of the warrant relation involves the audience either retrieving or accommodating

[4]Modified to reflect the fact that the relation between the premises and the conclusion may be SUPPORT or WARRANT, not entailment.

a speaker belief that someone who is Indian knows which Indian restaurants are good. The salient propositions are:

(75) a. The conversants want to select an Indian restaurant.

b. Ramesh is Indian.

c. Listen to Ramesh.

It seems very plausible that rather than **retrieving** a rule, such as *If X is Indian then X knows which Indian restaurants are good*, that the hearer can **infer** the inference rule *If A and B then C* using IMPLICATION INTRODUCTION. The inference of this rule makes the speaker's argument cohere. If the hearer wishes s/he can then **accommodate** this inference rule.

The implication introduction rule is very general and abstracts away from the type of A, B, and C. In example 69, A is a goal, B is a belief and C is an action, and B is a reason why the audience should intend C. The types of the propositions are all that is required to determine whether they are related by the SUPPORT relation (two beliefs) or the WARRANT relation (a belief and an intention). Implication introduction applies whether it is beliefs or intentions that are being deliberated.

7.2.4 Role of Redundancy

IRUs are often used in arguments and most of the time they consist of simple predicative statements such as *You're his mother*, *You're a psychologist*, or *You are retired*. Each of these statements defines a set which one can easily imagine as the restrictor in a syllogistic inference rule. Thus, it is plausible that both the form of these utterances and the fact that they are IRUs triggers the application of the IMPLICATION INTRODUCTION rule.

This view can be contrasted with the RETRIEVAL CUE HYPOTHESIS in which the function of the IRU is to trigger the retrieval of the inference rule from memory, and the IRU is matched against the left hand side of the inference rule.

Cohen argues that using propositions that are already mutually believed (i.e. IRUs) helps hearers' identify the argument structure. This can only hold for Deliberation IRUs and Affirmation IRUs. Cohen claims that the redundancy means that the IRU cannot be used as the conclusion of an inference process, so it must be used as a premise. Cohen's view might be a version of the COORDINATION HYPOTHESIS because it requires recognizing the IRU as redundant, and this recognition plays a role in recognizing the argument structure.

Another plausible motivation for using IRUs in arguments is to reduce the time devoted to deliberation. As Webber and Joshi point out, the hearer's recognition of the structure of the speaker's argument is **no guarantee** that the hearer will ACCEPT the speaker's argument (Webber and Joshi, 1982). However, if IRUs are used to support an argument, the validity of the support or warrant propositions is already accepted, and this makes it more likely that the hearer will accept the argument.

For example, assume that A and B are both necessary and sufficient reasons for concluding C. Assume that the speaker asserts C, because A and B, and the hearer believes that the asserted relation holds between A, B and C but in fact doesn't believe that A and B hold. Then the speaker must provide evidence for A and B. In contrast, if A and B are IRUs then there is no need for this recursive step to determine acceptance.

7.3 Open Segment

Open Segment IRUs are identified by the discourse correlate for Open Segment discussed in section 4.5. Their function will be discussed below with respect to the hypothesized functions in section 7.1.2.

7.3.1 Open Segment IRUs support the Coordination Hypothesis

Open Segment IRUs would, apriori, seem to be stellar support for the COORDINATION HYPOTHESIS. Open Segment IRUs often return to an earlier discussion and so can be analyzed as providing a cue that the current segment is subordinate to or part of the same segment as that earlier discussion. They may be used when an intention which has not been satisfied is returned to, after any type of intervening segment. Example 76, repeated here from section 3.6.1 (example 32), illustrates this with the two IRUs in 76-22. In this dialogue E has been telling H about how her money is invested, and asks her question in 76-3:

(76)
(3) E: .....
– and I was wondering – should I continue on with the certificates or
(4) H: Well, it's difficult to tell because we're so far away from any of them. But I would suggest this – if *all of these are 6 month certificates and I presume they are*
(5) E: *Yes*
(6) H: *Then I would like to see you start spreading some of that money around*
(7) E: uh hu
(8) H: Now in addition, how old are you?
.
(discussion and advice about starting an IRA)

.

(21) E: uh huh and

(22a) H: But as far as the certificates are concerned,

(22b) I'D LIKE THEM SPREAD OUT A LITTLE BIT -

(22c) THEY'RE ALL 6 MONTH CERTIFICATES

(23) E: Yes

(24) H: And I don't like putting all my eggs in one basket - and I would suspect that February 25 would be a good time to put it into something that runs for 2 and a half years. That first one that comes due. Call me on the others as they come due .....

In this case, it could be argued that the utterances in 76-22 conventionally indicate a return to the segment ending with 76-6, and thus support the COORDINATION HYPOTHESIS. The repetition or paraphrase of information that was part of an earlier context helps the hearer determine the relation of the information and intentions of the current discourse segment with that prior context.

Open Segment IRUs also support the DISCOURSE INFERENCE HYPOTHESIS, because 76-24 is inferentially related to the two IRUs in 76-22. Having all your certificates as 6 month certificates constitutes having all your eggs in one basket. An alternate course of action is described of spreading out the certificates and having some in two and a half year certificates.

Since all of the relevant information from the prior segment is restated in 22, it seems unlikely that the IRUs are retrieval cues. Thus this example provides no support for the RETRIEVAL CUE HYPOTHESIS.

7.3.2 Open Segment IRUs support the Discourse Inference Hypothesis

While 76 may be functionally ambiguous, the IRU in 77-58 supports only the DISCOURSE INFERENCE HYPOTHESIS.

(77) (41) E: I see: Oh and also *we have uh 34000 in the passbook.*

(42) H: *you have what?*

(43) E: *34 thousand in a passbook.*

(44) H: You're gonna make me cry.

(discussion of why they have so much in a passbook and what to do about it)

(49) E: Well, that'll be fine. *I have 40 thousand in in uh money market funds.*

(50) H: *You have 40 thousand in a money market fund?*

(discussion of why they have so much in a money market fund)

(58) H: AND YOU HAVE 40 IN A MONEY MARKET FUND AND 34 IN A PASSBOOK. Is that everything?

(59) E: Oh yes uh that's everything mhm.

(60) H: Ok. Let's first of all take the money in that passbook. Leave only a thousand in there. .....

In 77-58 H paraphrases two propositions, first asserted in (41) and (50). Each of these propositions were plausibly part of two distinct prior segments, each subparts of the description of E's investments; (58) summarizes what E has said so far. This example doesn't appear to argue for the RETRIEVAL CUE HYPOTHESIS, because the two IRUs represent exactly what has been discussed and there are no related propositions to retrieve from the prior segments. d to retrieve other propositions which were discussed in the segments in which these propositions were first introduced. These utterances do 'return' to the enumeration of E's investments which was in progress when H interrupted E at 77-50, but the segment starting with (58) is not subordinate to those prior segments and the information in (58) is used immediately in (59) and (60). Thus the DISCOURSE INFERENCE HYPOTHESIS is also an adequate explanation of the data.

The discussion of 76 and 77 shows that we can't definitively select among the three functional hypotheses on the basis of the distributional analysis. However, it appears that the RETRIEVAL CUE HYPOTHESIS is less plausible than the others because there is no evidence that other information from the context in which the antecedent occurred is needed. The following section discusses the distributional class of Close Segment IRUs with respect the three hypothesized discourse functions.

7.4 Close Segment

The final distributionally distinct class of Attention IRUs are Close Segment IRUs. The simplest type of closing statement is *okay* (Schegloff and Sacks, 1977). A slightly more contentful version from the financial advice corpus is *That was my question, okay.* In neither of these cases is the content of the statement important. However, Close Segment IRUs often include information that is important to remember or to understand properly. The main points of a proposed course of action are repeated, sometimes with propositional relations like causality made explicit, or reasons why another course of action cannot be taken are given. This section explores factors that might determine which information is repeated in Closing statements. As in sections 7.3 and 7.2, I will discuss the functional hypotheses presented in section 7.1.2 with respect to this distributionally defined class.

7.4.1 Close Segment IRUs support the Coordination Hypothesis

One advantage of an IRU as a Closing statement as compared with a less contentful utterance such as *Okay* is that the content of the IRU indicates the scope of the closing. In a hierarchical model of discourse, the scope of a closing indicates exactly how many of the stacked FOCUS SPACES would be popped off the stack. Close Segment IRUs provide support for the COORDINATION HYPOTHESIS because their occurrence can be correlated with aspects of the hierarchical model that make coordination difficult. Consider the IRU in 78-21:

(78)
(6) R: ..... or uh *Does that income from the certificate of deposit rule her out as a dependent*
(7) H: *Yes it does*
(8) R: *It does*
(9) H: *Yup, that knocks her out.*
Now there is something you can do. Do you support her in any way?
(10) R: Yes, I mean she yeah, we supply everything, heat, light, food. In other words, we, you know, she pays nothing as far as the uh upkeep of the home
(11) H: The only amount you have spent then is indirect.
(12) R: uh
(13) H: There's nothing direct
(14) R: yea
(15) H: then that you spend.
(16) R: Food
(17) H: The rest of her support comes out of her three thousand and social security.
(18) R: Yeah whatever clothes she needs er or uh dental care, that kind of thing
(19) H: Well, the medical and dental care you can deduct, provided you can establish that you have provided more than half support.
(20) R: uh huh
(21) H: BUT THE DEPENDENCY YOU CANNOT CLAIM
(22) R: um hm (breath) I see. Ok. uhh, Alright, the second question...

The antecedents for the IRU in 78-21 are 78-6 through 78-9. This example supports the COORDINATION HYPOTHESIS because it is plausible that the repetition of the answer in 78-21 is intended to close two segments: the current one starting at 78-9, *now there is something you can do* and the embedding segment starting with the question in 78-6. The scope of the Closing is clearly indicated by the content of the IRU because it coheres both propositionally and lexically with the original question in 78-6 (Halliday and Hasan, 1976; Morris and Hirst, 1991; Hearst, 1993).

This example also supports the DISCOURSE INFERENCE HYPOTHESIS because 78-19 is related to the IRU by a type of set-based contrast to be discussed more fully in section 8.2. The contrast arises because *dependency, medical deductions, dental deductions* are all members of the set of deductions. The contrast underlies an argument structure relation between 19 and 20 because 19 is an alternate

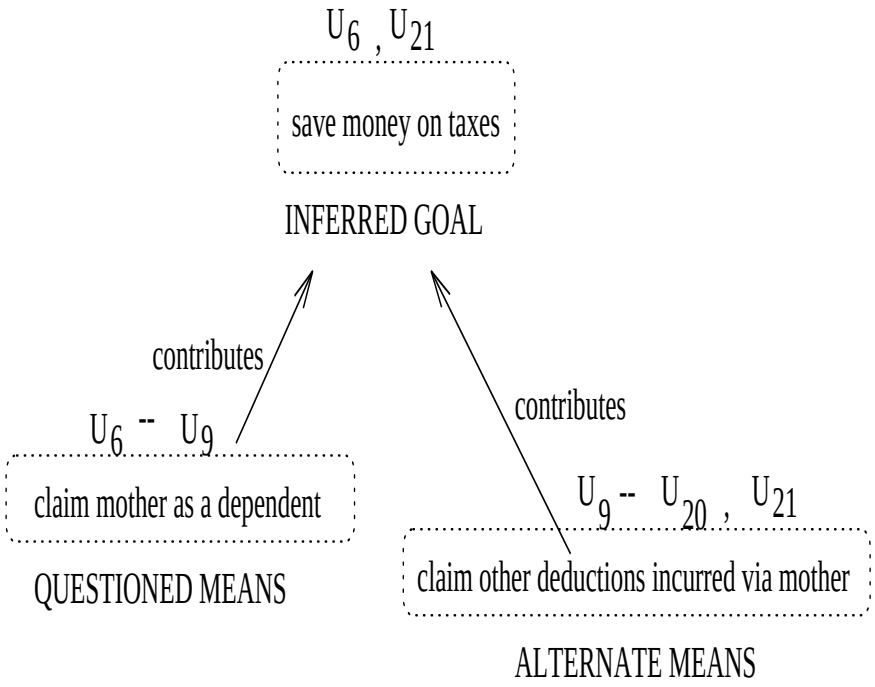


Figure 7.5: Relation of Means and Intentions in Dialogue 78

means to achieve the intention underlying the original question (Pollack, Hirschberg, and Webber, 1982).

This is illustrated more clearly in figure 7.5 which shows a plausible relationship between the utterances in 78 and the underlying intention structure. In hierarchical terms, 78-9 to 78-20 contributes to the intention inferred from 78-6. 78-21 is related to 78-19 because they describe two alternate ways to achieve the same intention; this is marked by the parallel topicalizations in these two utterances.

7.4.2 Close Segment IRUs support the Discourse Inference Hypothesis

The previous section discussed how Close Segment IRUs may be used conventionally to indicate discourse structure, thereby supporting the COORDINATION HYPOTHESIS. I also showed that 78 could support the DISCOURSE INFERENCE HYPOTHESIS. This section discusses another Close Segment IRU that supports the DISCOURSE INFERENCE HYPOTHESIS.

In 79, C is discussing how to save his daughter and her husband money on their taxes by having his daughter pay his wife for the child care that his wife provides for their grandchild. The amount of payment under discussion is two thousand dollars. The fact that his daughter could get 400 dollar

tax credit is first established in 79-20 ... 79-23, discussed again in 79-26, and finally paraphrased in 79-30.

(79) (20) H: Right. The maximum amount of credit that you will be able to get will be 400, that THEY will be able to get will be 400 dollars on their tax return.
(21) C: 400 dollars for the whole year?
(22) H: Yeah it'll be 20 percent.
(23) C: Um hm.
(24) H: Now if indeed they pay the 2000 dollars to your wife, that's great.
(25) C: um hm
(26) H: So we have 4 hundred dollars.
.....
(discussion how to avoid paying taxes on the $ 2000 his wife would then have)
.....
(30) H: You could do that too.
If the 2000 went into the I R A then you'd be completely protected,
and you'd not pay a tax on it,
and THEY WOULD GET A 400 DOLLAR CREDIT.

The Closing statement in 79-30 summarizes, and thereby ends, the discussion of how to avoid paying taxes on the 2000 dollars potential income. Thus, this example supports the COORDINATION HYPOTHESIS. The summary in 79-30 also reintroduces the fact that his daughter would get a tax credit; this is the result of the whole discussion and the answer to the caller's original question. Thus this example parallels that in 78. However the discussion in 79 is continued in 80:

(80) (31) C: um hm
(32) H: There's a child care credit right on the form.

In 80-32, H continues discussing the credit reintroduced as part of the Closing statement in 80-30.[5] We cannot determine whether the credit was discussed in 80-30 because H intended to say more about it in 80-32, or whether the fact that C merely backchannels in 80-31 prompts H to say more, or whether both the form and content of 80-32 shows that H assumes that the course of action has been agreed and that points of execution can now be discussed. Because it is possible that H reintroduced this proposition in order to talk about it in 80-32, this example also supports the DISCOURSE INFERENCE HYPOTHESIS.

[5]80-30 is classified as a Closing statement on the basis that 32 is a presentational there sentence.

In sum, Close Segment IRUs do not unambiguously support either the COORDINATION HYPOTHESIS or the DISCOURSE INFERENCE HYPOTHESIS. However, none of the Close Segment IRUs appear to require the RETRIEVAL CUE HYPOTHESIS.

7.5 Retrieval, Discourse Inference or Coordination?

In section 7.1.2, I introduced three different hypotheses about the more specific discourse functions of Attention IRUs. In sections 7.3, 7.4 and 7.2 I discussed the various distributionally identified classes of IRUs with respect to the hypothesized discourse functions.

Open Segment IRUs are best explained by either the DISCOURSE INFERENCE HYPOTHESIS or the COORDINATION HYPOTHESIS. Close Segment IRUs are also best explained by the DISCOURSE INFERENCE HYPOTHESIS or the COORDINATION HYPOTHESIS. The RETRIEVAL CUE HYPOTHESIS was not plausible for these IRUs because no other information from the prior context was relevant to the current intention. Deliberation IRUs can be explained by either the DISCOURSE INFERENCE HYPOTHESIS or the RETRIEVAL CUE HYPOTHESIS. In sum, none of the hypotheses about discourse function are eliminated, but the DISCOURSE INFERENCE HYPOTHESIS is a viable explanation in every case.

The fact that it is not possible to distinguish the COORDINATION HYPOTHESIS from the DISCOURSE INFERENCE HYPOTHESIS suggests that perhaps Open and Close Segment IRUs simultaneously achieve both functions and that in some circumstances the speaker may be primarily concerned with one function while in other circumstances the speaker is primarily concerned with the other. The speaker's intentions might be determined by aspects of the task that are not measured here, e.g. whether the hearer is already familiar with the task structure or the degree of complexity of the task.

One final consideration is whether the function of Attention IRUs depends on recognizing them as redundant. According to the DISCOURSE INFERENCE HYPOTHESIS they need not be recognized as such. According to the COORDINATION HYPOTHESIS, some IRUs are conventionally used to help the hearer recognize the discourse structure (task structure). Use of this convention requires the hearer to recognize redundancy. However it is possible that that coordination would happen without recognition simply because these IRUs propositionally and lexically cohere with a previous segment. If retrieval doesn't happen automatically (Ratcliff and McKoon, 1988), then the RETRIEVAL CUE HYPOTHESIS is dependent on the assumption that IRUs are recognized as redundant.

Figure 7.5 shows that Attention IRUs are not marked as HEARER OLD information in the same way that Other IRUs are. Other IRUs tend to be realized with phrase final Mid $\chi^2 = 4.435, p < .05$.

	Phrase Final Mid	Phrase Final Low
Attention	11	20
Not-Attention	38	27

Figure 7.6: Not-Attention as a Predictor of Mid

Therefore, figure 7.5 provides weak evidence against the RETRIEVAL CUE HYPOTHESIS. However, I have not shown that Attention IRUs cannot be distinguished from utterances consisting wholly of new information. It is possible that prosodic contour and not boundary tones are used to mark IRUs as HEARER OLD information. It is also plausible that IRUs are recognized as redundant independently of their prosody.

In addition, if the COORDINATION HYPOTHESIS were true, it would be plausible that Open and Close Segment IRUs would also be marked with cue words that have been claimed to indicate discourse structure. An examination of cue words on Open Segment IRUs such as *now, okay, but* and on Close segment IRUs such as *so, then, well* shows that neither of these classes of IRUs are reliably marked with such cues (Polanyi and Scha, 1984; Grosz and Sidner, 1986; Schiffrin, 1987; Hirschberg and Litman, 1987). However, it is possible that the IRUs are such good coordination cues that no other cues are needed. Thus this is weak evidence against the COORDINATION HYPOTHESIS.

In sum, an examination of the corpus does not rule out any of the hypotheses discussed here. Furthermore, due to the simplicity of Design-World, the simulation results discussed in section 7.6 cannot eliminate any of the hypothesized functions. However the simulation results will show that the DISCOURSE INFERENCE HYPOTHESIS is viable.

7.6 Attention in Design World

The COORDINATION HYPOTHESIS cannot be tested in Design-World because the discourse structure is too simple. There are no ambiguities as to which segments are currently open and which should be closed. This means that the potential return and subordination effects of Open statements and the scoping effects of Closing statements cannot be tested. Closing statements cannot help the agent infer the task structure because both agents know the structure of the task. Furthermore, the RETRIEVAL CUE HYPOTHESIS cannot really be tested in Design-World because agents know which inferences rules to use and what facts to use in means-end reasoning. Design-World can provides a testbed for the DISCOURSE INFERENCE HYPOTHESIS, and the DISCOURSE INFERENCE CONSTRAINT.

In Design-World, Attention strategies include Attention IRUs, those whose antecedents are no longer salient. As with other Design-World strategies, the range of possible Attention strategies is limited. One set of limitations comes from the domain and the methods that can be developed for selecting the content of IRUs. The strategy for selecting propositions to realize as Open Segment and Deliberation IRUs is to use those that the speaker is actively using in reasoning, and vary how many of these propositions are realized explicitly. The hypothesis is that IRUs allow the other agent to duplicate the speaker's reasoning process with less effort. The speaker's IRU can potentially save the hearer the processing involved with (1) retrieving facts from memory that are to be used in reasoning, (2) the reasoning required to determine that those facts should be retrieved, and (3) drawing potentially relevant inferences. Using the speaker's own reasoning as the basis for selecting the content of IRUs can be contrasted with using a model of the hearer's reasoning. For simplicity, no user modeling was attempted.

Another set of limitations comes from the structure of the discourse. As shown in figure 5.4 each primitive put-act has an implicit or an explicit Opening, Proposal, Acceptance and Closing statement. Attention IRUs cannot occur in the Acceptance position since by definition none of the examples of Attention IRUs consist of salient information. This leaves possible Attention loci as Opening, Proposal or Closing.

Because Opening statements are juxtaposed with Proposals (see figure 5.4), we would not expect the effects of IRUs in an Opening statement to be different than those at the beginning of a Proposal. The only difference will be that Proposals include different types of information than those in explicit Opening statements. The information in Opening statements is general information that might be useful in means-end reasoning about the put-act intention for that segment. IRUs as part of proposals consist of information specifically related to the proposal that is made.

Three strategies were tested: (1) the Open-Best strategy includes the next best option that the speaker has identified during means-end reasoning in an opening statement at the beginning of each Put-Act segment; (2) the Explicit-Warrant strategy includes facts that are warrants for the proposal with each proposal; and (3) the Matched-Pair-Premise strategy includes IRUs in proposals that can be used as premises for making matched pair inferences in the MATCHED-PAIR version of the task. No Close Segment IRU strategies are discussed here because there is no benefit for them in the Design-World task and because section 8.4.1 presents results of Close Segment IRUs that make inferences explicit.

Section 7.6.1 will show that strategies incorporating Open Segment IRUs are not beneficial. The conclusion is that if the task is easy enough, making propositions salient is not useful: IRUs should be targeted on achieving specific inferences. Section 7.6.2 will present Design-World results showing

that DELIBERATION IRUs can reduce retrieval costs, simplify deliberation, and achieve a high level of agreement as evaluated by the ZERO NONMATCHING BELIEFS evaluation function. Section 7.6.3 will show that when inferential complexity is increased, as in the MATCHED-PAIR task, making premises salient increase agents' ability to make inferences.

7.6.1 Open Segment in Design-World

The simplest Opening statement is:

(Say A B (Open (Achieve A&B (Design Room-1))))

A naturally-occurring utterance most similar to this would be *Let's figure out how to do the first room*, which could be left out of the dialogue because an utterance such as *Let's put the red couch in the first room* would license the inference that the speaker's intention is to work on the first room. In other words, a proposal to (Put A&B Red-Couch Room-1 Later) can be easily inferred to be an action that CONTRIBUTES to the goal (Achieve A&B (Design Room-1)) and there is no need to have this goal made explicit beforehand.

Strategies with Open statements such as these were tested and shown to have no benefit in Design-World because both agents know the structure of the task. Thus they will not be presented here. A variant will be presented below, which was hypothesized to potentially have more benefit because it includes Open Segment IRUs with propositional content such as those found in the corpus.

7.6.1.1 Open-Best Strategy: IRUs for Means-End Reasoning

Figure 7.7 shows the discourse action protocol for Open Best agents. These agents leave Acceptance and Closing implicit but produce explicit Opening statements which consist of additional information at the Opening of each discourse segment. This additional information consists of an IRU that could be used in means-end reasoning for the current segment. The Open-Best strategy is illustrated by the Opening statement of OPnc in 81-1 and the the Opening statment by OPnc2 in 81-3:

(81) OPnc: AGENT-BILL HAS A GREEN COUCH.
1:(say agent-opnc agent-opnc2 bel-10: has (agent-bill green couch))
OPnc: *Then, let's put the green rug in the study.*
2:(propose agent-opnc agent-opnc2 option-10: put-act (agent-bill green rug room-1))
OPnc2: AGENT-KIM HAS A PURPLE COUCH.
3:(say Agent-opnc2 agent-opnc bel-36: has (agent-kim purple couch))

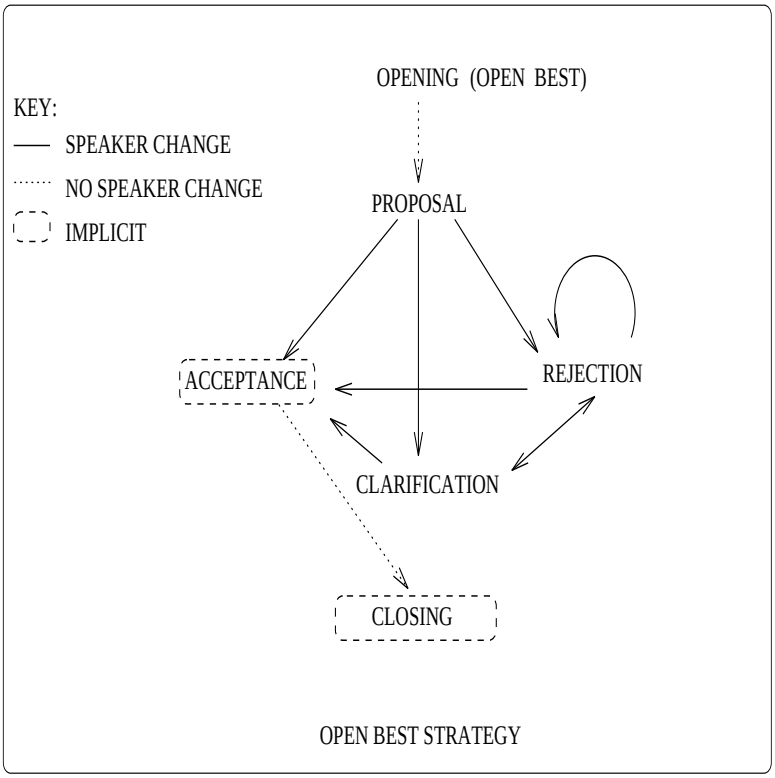


Figure 7.7: Discourse Action protocol for agents using the Open Best strategy

OPnc2: *Then, let's put the green lamp in the study.*
4:(propose agent-opnc2 agent-opnc option-35: put-act (agent-kim green lamp room-1))

The IRUs used in the Open-Best strategy are based on the options identified by an agent's means-end reasoning. The IRU that is included is that the agent has another (high scoring) piece that could potentially be used in the room being designed. Figure 7.8 shows what a sequence of discourse acts might be for a dialogue between two Open-Best agents.

Figure 7.9 shows the discourse acts, utterance acts and the cognitive processing for the two agents for the dialogue excerpt in 81.

This Opening strategy is compared with the All-Implicit strategy described in chapter 5, example 46. The results of simulations using this strategy are discussed below.

7.6.1.2 Open Best Strategy is Detrimental

Figure 7.10 shows that the Open-Best strategy provides no benefits over the All-Implicit strategy even when communication, inference and retrieval are free (KS NotSig at all AWM). The inclusion

Agent A	Agent B
OPENING 1 PROPOSAL 1	
	OPENING 2 PROPOSAL 2
REJECTION 2,3	
	OPENING 4 PROPOSAL 4
...	...
	OPENING N PROPOSAL N

Figure 7.8: Sequence of Discourse Acts for Two Open-Best Agents as in Dialogue 81

of the IRUs apparently makes agents forget facts that would have been just as useful as those made salient by the IRUs.

In addition, figure 7.11 shows that if communication, inference and retrieval have some cost, then the Open-Best strategy is detrimental. The increased number of messages and the displacement of facts used in deliberation have the largest effect at higher values of memory, but can be seen by AWM of 4 (KS of 4 and above $> .2$, $p < .05$).

That this detrimental effect is related to the cost of retrieval can be seen by increasing that cost as shown in figure 7.12 where the detrimental effect is even more pronounced (KS $> .29$ for AWM 4 and above, $p < .01$).

The Open-Best strategy is also detrimental if the evaluation function is ZERO-INVALID or ZERO-NONMATCHING-BEL, but the difference graphs are not presented here. Rather strategies which are beneficial will be discussed in the following sections and contrasted with the Open-Best strategy.

7.6.2 Deliberation IRUs in Design World

Design-World focuses on achieving agreement about intentions rather than beliefs, so Deliberation IRUs in Design-World are WARRANT IRUs. The following sections will first describe a strategy that includes WARRANT IRUs and then present the results of evaluating that strategy in different task situations.

Discourse Act	Utterance Act	A PROCESS	B PROCESS
		ME-REASON Put-1 Deliberate Open, Open Best	
OPENING	1:(say A B bel-10)	Store bel-10	
			Infer Open Put-1 Store bel-10
PROPOSAL	2:(propose A B option-10)		
			ME-REASON Put-1 Check Preconds 10 Deliberate Decide to Accept Store (MS intend 10) Store (Act-Effects 10) ME-REASON Put-2 Deliberate Open, Open Best
OPENING	3:(say B A bel-36)		Store bel-36
		Infer Acceptance 10 Store (MS intend 10) Store (Act-Effects 10) Infer Close Put-1 Infer Open Put-2 Store bel-36	
PROPOSAL	4:(propose B A option-35)		
		ME-REASON Put-2 Check Preconds 35 Check Matches Deliberate Decide to Accept Store (MS intend 35) Store (Act-Effects 35) ME-REASON Put-3 Deliberate Open, Open Best	
	... .		

Figure 7.9: Cognitive Processes for Sequence of Discourse Acts in Dialogue 81: Agent A and B are Open Best Agents

7.6.2.1 The Explicit-Warrant strategy

The Explicit-Warrant strategy includes a WARRANT IRU along with every proposal, ie. every proposal includes Points information as shown below:

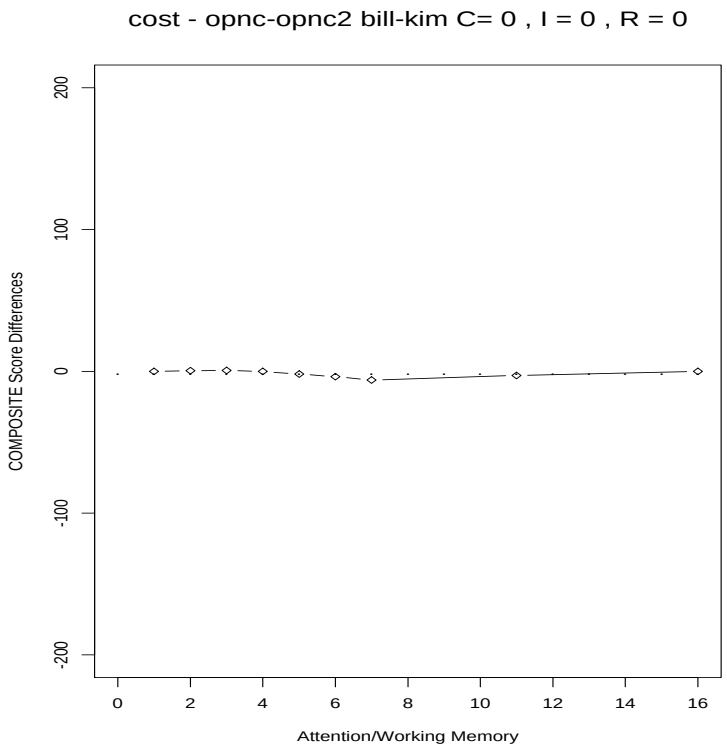


Figure 7.10: Open Best is not Beneficial. Strategy 1 is two Open-Best agents and strategy 2 is two All-Implicit agents, Task = Standard, commcost = 0, infcost = 0, retcost = 0

(Say A B (Points Red-Couch 30))

(Propose A B (Put A&B Red-Couch Room-1 Later))

The Explicit-Warrant strategy is representative of the Deliberation IRUs discussed in section 7.2 because the points information is used by the hearer to deliberate about whether to accept or reject the proposal. Thus in this particular case, the fact that the red couch is worth 30 points is meant to be used in reasoning about whether to adopt the intention of putting the red couch in room-1. This parallels 69 where the fact that Ramesh is Indian was meant to be used as a reason for adopting an intention to listen to Ramesh. Figure 7.13 shows the discourse action protocol for EXPLICIT WARRANT agents.

In 82, both agents use the Explicit-Warrant strategy. The WARRANT IRUs are shown in CAPS:

(82) IEI: PUTTING IN THE GREEN RUG IS WORTH 56.
1:(say agent-iei agent-iei2 bel-10: score (option-9: put-act (agent-bill green rug room-1) 56))
IEI: *Then, let's put the green rug in the study.*
2:(propose agent-iei agent-iei2 option-10: put-act (agent-bill green rug room-1))

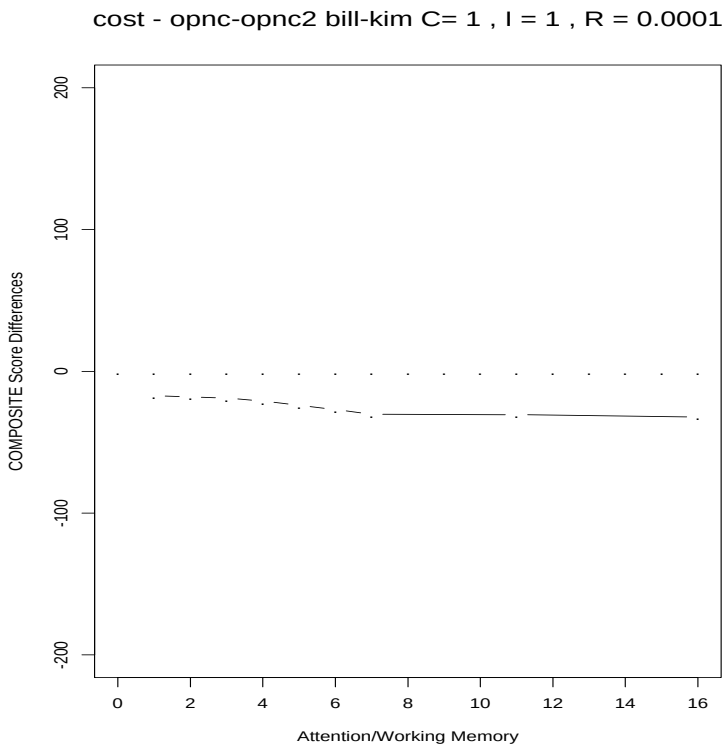


Figure 7.11: Open-Best can be Detrimental: Strategy 1 is two Open-Best agents and strategy 2 is two All-Implicit agents, Evaluation Function = COMPOSITE-COST, commcost = 1, infcost = 1, retcost = .01

IEI2: PUTTING IN THE GREEN LAMP IS WORTH 55.

3:(say agent-iei2 agent-iei bel-34: score (option-22: put-act (agent-kim green lamp room-1) 55))

IEI2: *Then, let's put the green lamp in the study.*

4:(propose agent-iei2 agent-iei option-33: put-act (agent-kim green lamp room-1))

Figure 7.14 provides an abstraction of the dialogue in 82. Figure 7.15 shows the discourse acts, the utterance acts, and the cognitive processes for both agents for dialogue 82. The rest of this section will discuss the results of comparing the Explicit-Warrant strategy with the All-Implicit strategy.

7.6.2.2 Explicit Warrant reduces Retrievals

Dialogues in which one or both agents use the Explicit-Warrant strategy are more efficient in certain discourse situations. As shown in figure 7.16, the Explicit-Warrant strategy is detrimental at AWM of 3,4,5 for the standard task if retrieval from memory is free (KS 3,4,5 > .19, $p < .05$).

However, contrast the case discussed above where retrieval is free with the differences in scores shown in figure 7.17 when retrieval is one tenth the cost of communication and inference. Here, by

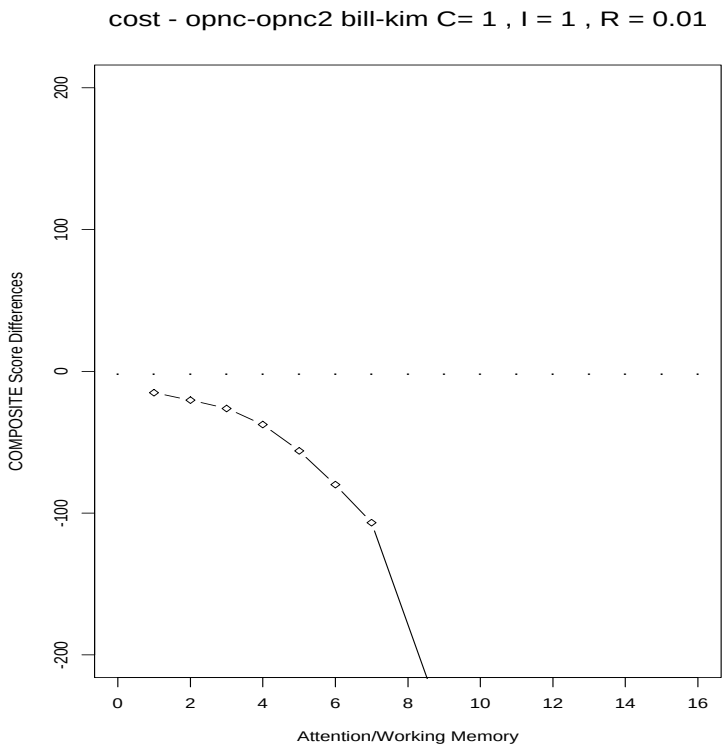


Figure 7.12: Open Best increases Retrievals. Strategy 1 is two Open-Best agents and strategy 2 is two All-Implicit agents, Evaluation Function = COMPOSITE-COST, commcost = 1, infcost = 1, retcost = .01

AWM values of 3, we see an improvement in scores because the beliefs necessary for reasoning are made available in the current context with each proposal (KS for AWM of 3 and above $> .23$, $p < .01$). At AWM parameter settings of 16, where agents can search a huge belief space for beliefs to be used as warrants, the saving in processing time is substantial.

7.6.2.3 Explicit Warrant is no benefit if Communication is Expensive

Figure 7.18 shows that if communication is expensive, e.g. 10 times as much as inference or retrieval, then the Explicit-Warrant strategy is not beneficial (KS for AWM 1 to 5 $> .23$, $p< .01$).

This is because providing explicit warrants increases the number of utterances required to perform the task; it doubles the number of messages in every proposal. If communication is expensive compared to retrieval, communication cost can dominate the other benefits.

7.6.2.4 Explicit Warrant Achieves a High Level of Agreement

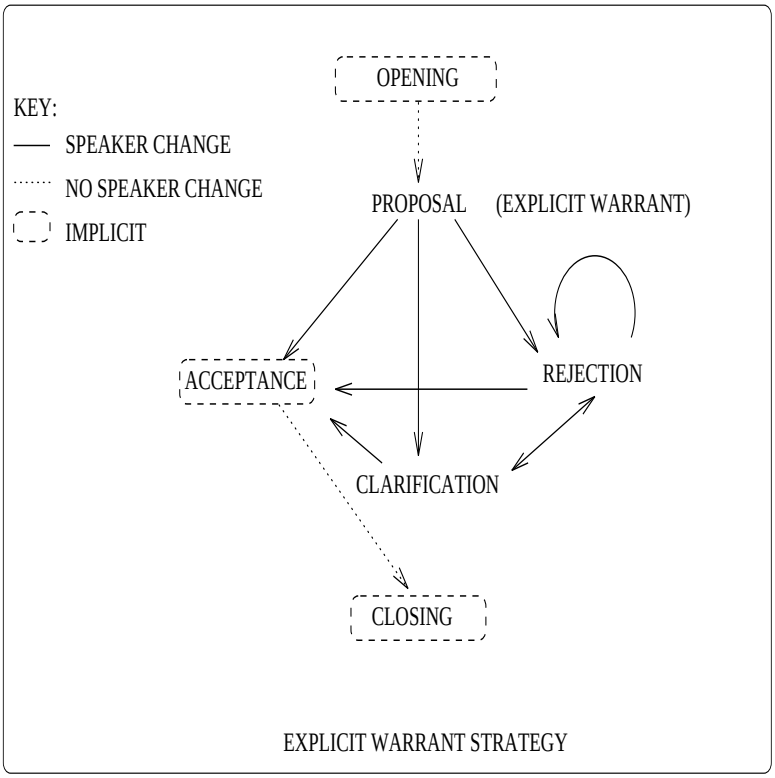


Figure 7.13: Discourse Action protocol for agents using the Explicit-Warrant strategy

Agent A	Agent B
PROPOSAL 1	
	PROPOSAL 2
REJECTION 2,3	
	PROPOSAL 4
...	...
	PROPOSAL N

Figure 7.14: Sequence of Discourse Acts for Two Explicit Warrant Agents in Dialogue 82

The Explicit-Warrant strategy is also beneficial in achieving a high degree of agreement. This is shown when the task is Zero-Nonmatching-Beliefs; figure 7.19 shows that the Explicit-Warrant strategy is beneficial even if retrieval is free ($KS > .23$ for AWM from 2 to 11, $p < .01$). The benefits are because the warrant information that is redundantly provided is exactly the information that is needed in order to achieve matching beliefs about the warrants for actions under discussion. The strategy virtually guarantees that the agents will agree on the reasons for carrying out a particular course of action.

Discourse Act	Utterance Act	A PROCESS	B PROCESS
		ME-REASON Put-1 Deliberate Propose Include Warrant	
PROPOSAL	1:(say A B bel-10)	Store bel-10	
			Infer Open Put-1 Store bel-10
	2:(propose A B option-10)		
			ME-REASON Put-1 Check Preconds 10 Deliberate Decide to Accept Store (MS intend 10) Store (Act-Effects 10) ME-REASON Put-2 Deliberate Propose Exp-Warr
PROPOSAL	3:(say B A bel-34)		Store bel-34
		Infer Acceptance 10 Store (MS intend 10) Store (Act-Effects 10) Infer Close Put-1 Infer Open Put-2 Store bel-34	
	4:(propose B A option-33)		
		ME-REASON Put-2 Check Preconds 33 Deliberate Decide to Accept Store (MS intend 33) Store (Act-Effects 33) ME-REASON Put-3 Deliberate Propose Exp-Warr	
	... .		

Figure 7.15: Cognitive Processes for Sequence of Discourse Acts in Dialogue 82: Agent A and B are Explicit Warrant Agents

For agents with high AWM, the Explicit-Warrant strategy is also beneficial when the task is fault intolerant (KS $>$ .23, for AWM 11 and 16, $p < .05$). This is due to the fact that agents with high AWM tend to make mistakes about which pieces they still have since they don't distinguish between outdated beliefs and recent beliefs. The belief deliberation component relies on the fact that agents are unlikely to retrieve beliefs that are inconsistent with recent events, however agents with high AWM can do so. If the belief deliberation component were modified so that temporal

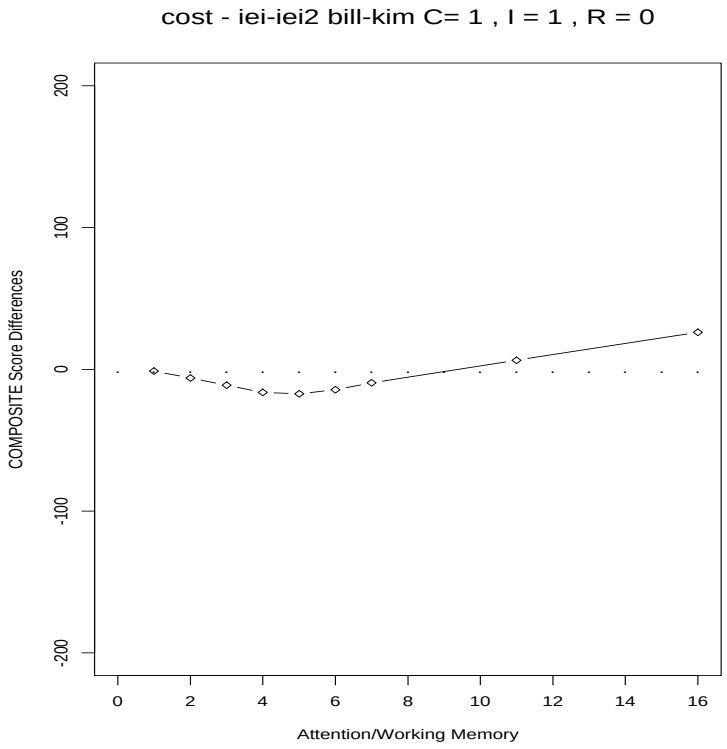


Figure 7.16: If Retrieval is Free Explicit-Warrant is detrimental at AWM of 3,4,5: Strategy 1 of two Explicit-Warrant agents and strategy 2 of two All-Implicit agents: Evaluation Function = COMPOSITE-COST, commcost = 1, infcost = 1, retcost = 0

information was considered when deliberating about beliefs, then the Explicit-Warrant strategy would no longer be beneficial for the Zero-Invalid task.

7.6.2.5 Summary: Explicit Warrant

The most striking result of the evaluation of the EXPLICIT WARRANT strategy is that the benefits of the strategy are highly situation specific. Whether the strategy is good or bad depends on parameters of the task situation. This shows that there is no task independent way of defining cooperative behavior. Cooperative strategies are specific to the communicative situation and should be adaptable to the situation at hand. Furthermore, the experiments clearly show that IRUs are not beneficial simply for rehearsal. The benefits of the Explicit-Warrant strategy are greatest when the strategy specifically targets aspects of the task at hand.

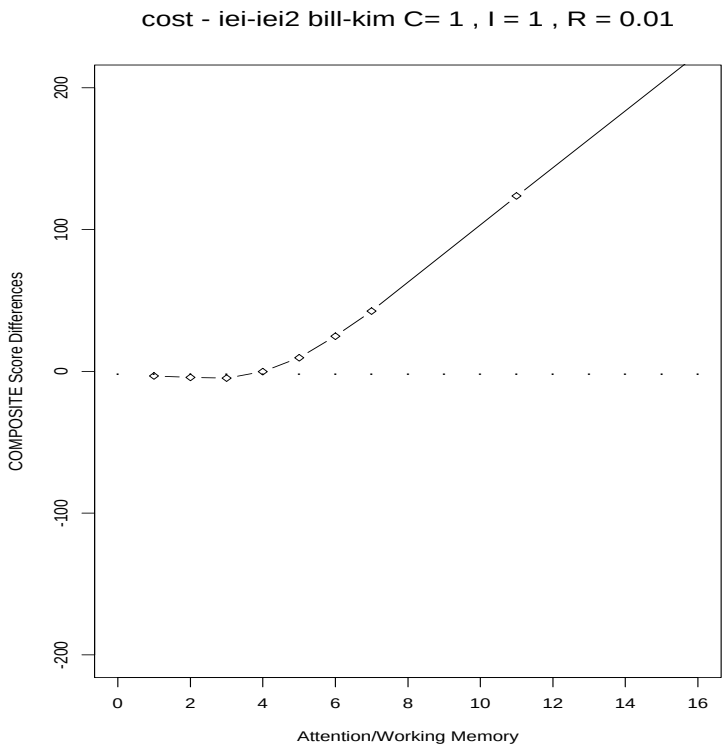


Figure 7.17: Retrieval costs: Strategy 1 is two Explicit-Warrant agents and strategy 2 is two All-Implicit agents: Evaluation Function = COMPOSITE-COST, commcost = 1, infcost = 1, retcost = .01

7.6.3 Increasing Inferential Complexity: Attention IRUs

7.6.3.1 The Matched-Pair-Premise strategy

As discussed earlier, there is a version of the task called the MATCHED-PAIR task that increases inferential complexity. The Attention strategy targeted at improving performance on the MATCHED-PAIR task is called the Matched-Pair-Premise strategy: agents using this strategy make premises salient that could potentially be used to infer matched pair goals. A proposal in the Matched-Pair-Premise strategy consists of a statement about another matching piece along with a proposal:

(Say A B (Has A Green-Couch))
(Propose A B (Put A&B Green-Chair Room-1 Later))

The information about matching pieces is made explicit even though the other agent could retrieve this information. In addition, this doesn't mean that agent A has already inferred a matched-pair: matched pair inferences depend on current intentions. In order for the agents to achieve a matched pair with a green chair and a green couch, the proposal about the green chair would have to be

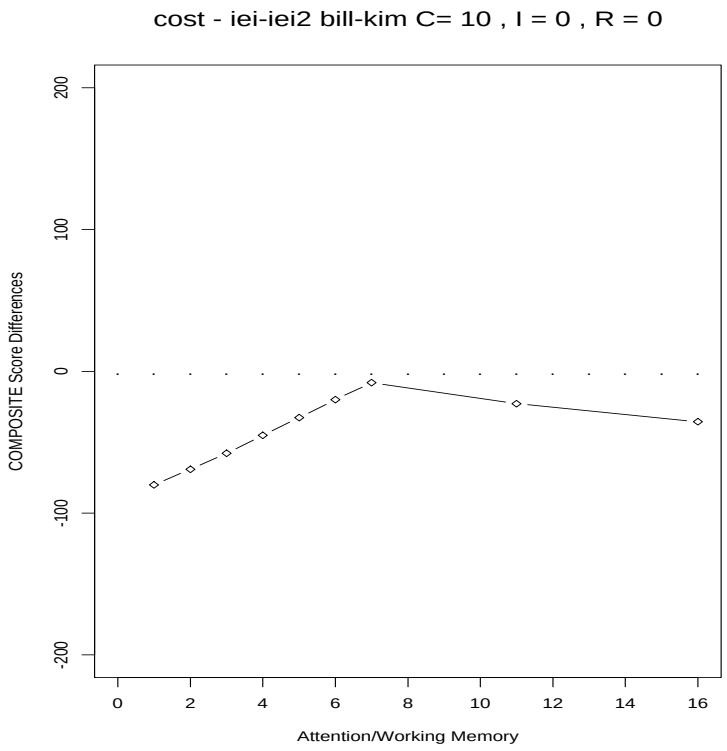


Figure 7.18: If Communication is Expensive: Communication costs can dominate other costs in dialogues. Strategy 1 is two Explicit-Warrant agents and strategy 2 is two All-Implicit agents: Evaluation Function = COMPOSITE-COST, commcost = 10, infcost = 0, retcost = 0

accepted as well as a proposal about the green couch and the matched pair inference would have to be made by both agents.

The Matched-Pair-Premise strategy is illustrated by the proposal made by agent IBI in 83-1:

(83) IBI: *Agent-Bill has a green couch.*
1:(say agent-ibi agent-ibi2 bel-10: has (agent-bill green couch))
IBI: *Then, let's put the green rug in the study.*
2:(propose agent-ibi agent-ibi2 option-10: put-act (agent-bill green rug room-1))
IBI2: AGENT-BILL HAS GREEN COUCH.
3:(say agent-ibi2 agent-ibi bel-36: has (agent-bill green couch))
IBI2: *Then, let's put the green lamp in the study.*
4:(propose agent-ibi2 agent-ibi option-35: put-act (agent-kim green lamp room-1))

Figure 7.20 shows the discourse action protocol for a Matched Pair Premise agent. Openings, Closings and Acceptances are left implicit, but Proposals are structured using the Matched-Pair-Premise strategy. Figure 7.21 shows a dialogue sequence composed of the interaction of a two Matched-Pair-Premise agents. Figure 7.22 shows the dialogue sequence, the utterance acts and the cognitive processes for dialogue 83 between two Matched-Pair-Premise agents.

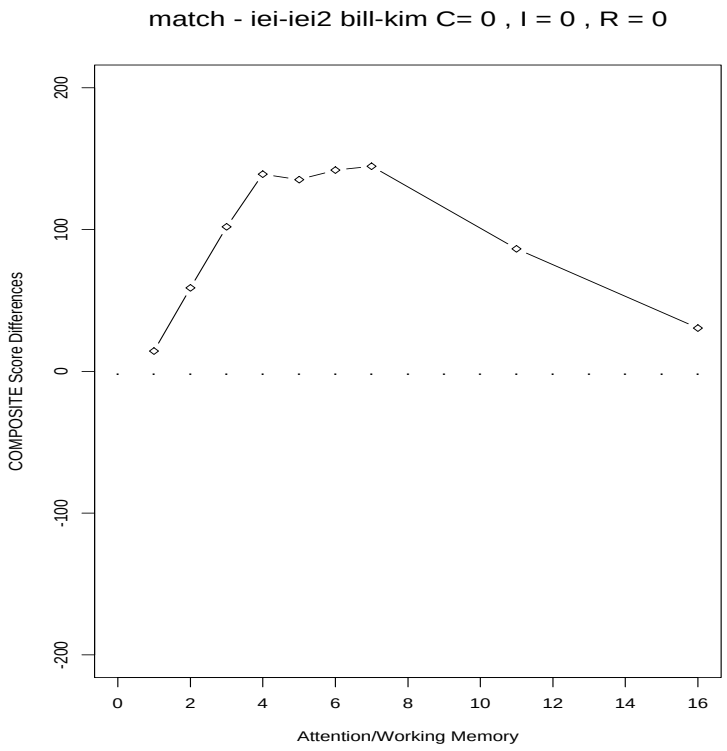


Figure 7.19: Beliefs Match with Explicit-Warrant: Strategy 1 is two Explicit-Warrant agents and strategy 2 is two All-Implicit agents: Evaluation Function = ZERO NONMATCHING-BELIEFS, commcost = 0, infcost = 0, retcost = 0

7.6.3.2 Salient Premises for Matched Pairs improves Performance

As usual, the Matched-Pair-Premise strategy will be compared with the All-Implicit strategy. Figure 7.23 shows that the Matched-Pair-Premise strategy is beneficial at low AWM by increasing agents' ability to make matched pair inferences (KS(3) = .23, $p < .01$; KS(4) = .2, $p < .05$).

This result contrasts sharply with the negative results for the Open-Best strategy. The Matched-Pair-Premise strategy includes IRUs which are not **guaranteed** to be useful for making matched pair inferences, they are only **likely** to be useful. The Open-Best strategy was similar in including IRUs reminding other agents of the existence of high scoring pieces. However in this situation the strategy improves performance. The improvement is due to the slightly increased inferential complexity of the MATCHED-PAIR task.

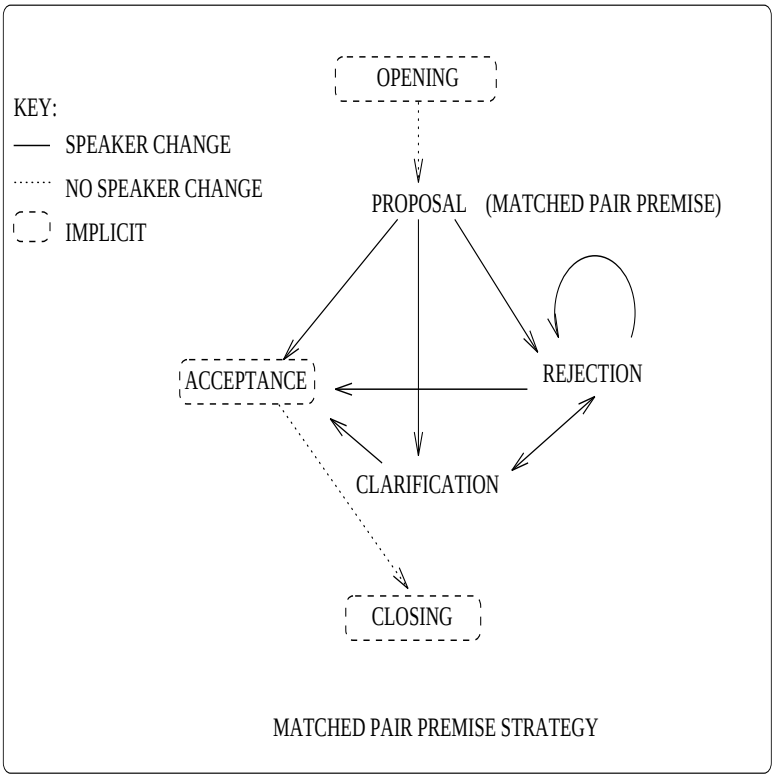


Figure 7.20: Discourse Action protocol the Matched-Pair-Premise strategy

Agent A	Agent B
PROPOSAL 1	
	PROPOSAL 2
REJECTION 2,3	
	PROPOSAL 4
...	...
	PROPOSAL N

Figure 7.21: Sequence of Discourse Acts for Two Matched-Pair Premise Agents in Dialogue 83

7.6.3.3 Salient Premises for Matched Pairs increases Retrievals

Because the agents are carrying out the standard version of the task in addition to the MATCHED-PAIR optional goals, they still must retrieve score propositions when they deliberate their own options or proposals made by the other agent. Figure 7.24 shows that the Matched-Pair-Premise strategy increases the overall number of retrievals. This is because the propositions which are said in order to increase the likelihood of inferring a matched pair displace the warrant propositions (beliefs about the scores of pieces). Consequently, more effort is required to retrieve them. The

Discourse Act	Utterance Act	A PROCESS	B PROCESS
		ME-REASON Put-1 Deliberate Check Future Match Propose, use MPP	
PROPOSAL	1:(say A B bel-10)	Store bel-10	
			Store bel-10 Infer Open Put-1
	2:(propose A B option-10)		
			ME-REASON Put-1 Deliberate Decide to Accept Store (MS intend 10) Store (Act-Effects 10) ME-REASON Put-2 Check Matches Check Future Match Deliberate Propose, use MPP
PROPOSAL	3:(say B A bel-36)		Store bel-36
		Infer Acceptance 10 Store (MS intend 10) Store (Act-Effects 10) Infer Close Put-1 Infer Open Put-2 Store bel-36	
	4:(propose B A option-35)		
		ME-REASON Put-2 Deliberate Decide to Accept Store (MS intend 35) Store (Act-Effects 35) ME-REASON Put-3 Check Matches Check Future Match Deliberate Propose, use MPP	
	... .		

Figure 7.22: Cognitive Processes for Sequence of Discourse Acts in Dialogue 83: Agent A and B are Matched-Pair-Premise Agents

following section describes a strategy that attempts to address this problem by also including IRUs for score propositions.

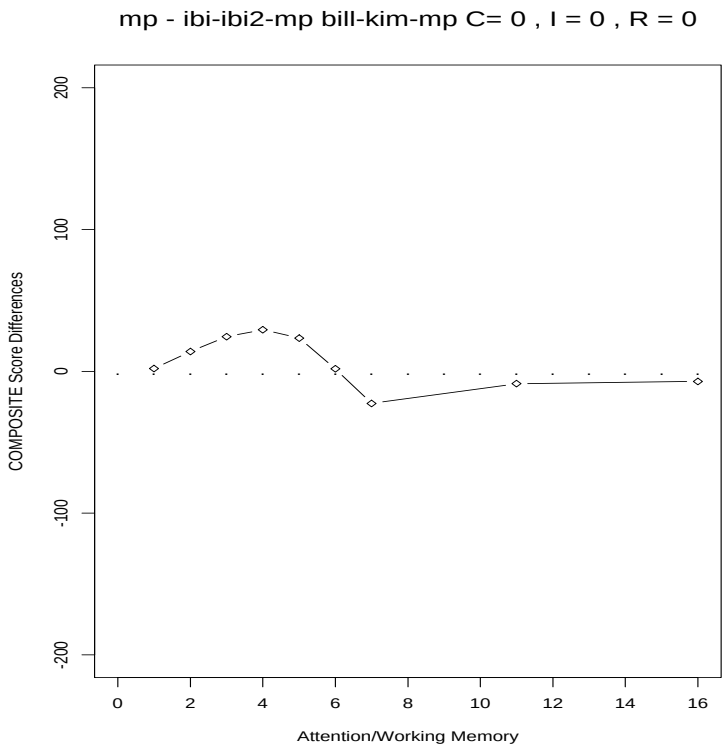


Figure 7.23: The Matched-Pair-Premise strategy increases the number of Matched Pair inferences. Strategy 1 is the two Matched-Pair-Premise agents and Strategy 2 is two All-Implicit agents. Task = MP, commcost = 0, infcost = 0, retcost = 0

7.6.3.4 There are tradeoffs between Retrieval and Communication

In an attempt to develop a strategy that would combine the benefits of the Matched-Pair-Premise strategy with the reduction in retrieval of the Explicit-Warrant strategy, a second strategy for use in the MATCHED-PAIR task was tested. This strategy is the Matched-Pair-Premise-Warrant strategy: a proposal in this strategy includes both a premise for a potential matched pair inference and the warrant associated with that potential matched pair inference. This is illustrated by IBSI's proposal in 5,6 and 7 in 84:

(84) IBSI: AGENT-BILL HAS A FUCHSIA RUG.
5:(say agent-ibsi agent-ibsi2 bel-59: has (agent-bill fuchsia rug))
IBSI: PUTTING IN THE FUCHSIA RUG IS WORTH 52.
6:(say agent-ibsi agent-ibsi2 bel-60: score (option-58: put-act (agent-bill fuchsia rug room-1) 52))
IBSI: *Then, let's put the fuchsia couch in the study.*
7:(propose agent-ibsi agent-ibsi2 option-62: put-act (agent-bill fuchsia couch room-1))
IBSI2: *Then, let's put the fuchsia rug in the study.*
8:(propose agent-ibsi2 agent-ibsi option-86: put-act (agent-bill fuchsia rug room-1))

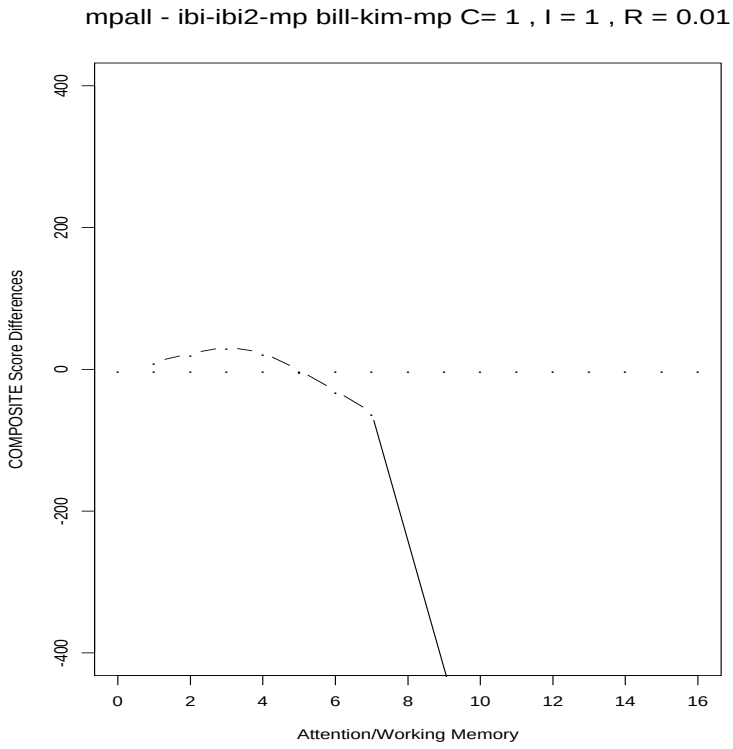


Figure 7.24: The Matched-Pair-Premise strategy increases number of retrievals. Strategy 1 is the two Matched-Pair-Premise agents and strategy 2 is two All-Implicit agents. Evaluation function = COMPOSITE-COST, commcost = 1, infcost = 1, retcost = .01

As shown by the proposal by IBSI2 in 84-8, IBSI2 appears to be 'convinced' by this extra information into proposing to use the fuchsia rug and thereby achieving a matched pair.

Figure 7.25 shows that at LOW AWM (3 and 4) this strategy succeeds at increasing agents' ability to get matched pairs ($KS(3,4) > .19$, $p < .05$).

Like the Matched-Pair-Premise strategy, the Matched-Pair-Premise-Warrant makes premises salient that are **likely** to be useful, but not guaranteed to be so. Unlike the Open strategies that were tested, both of these strategies are helpful. In this situation, unlike the situation in which the Open strategies were tested, the inferences are more difficult to make. Thus a strategy that is only likely to help can be beneficial.

In order to see whether including the additional score information means that retrieval is reduced as it is in the Explicit-Warrant strategy, figure 7.26 compares retrievals only, with a retrieval cost of .01 for Matched-Pair-Premise strategy with the Matched-Pair-Premise-Warrant strategy. As the figure shows, including the warrants does reduce the amount of retrieval. However, as before, in

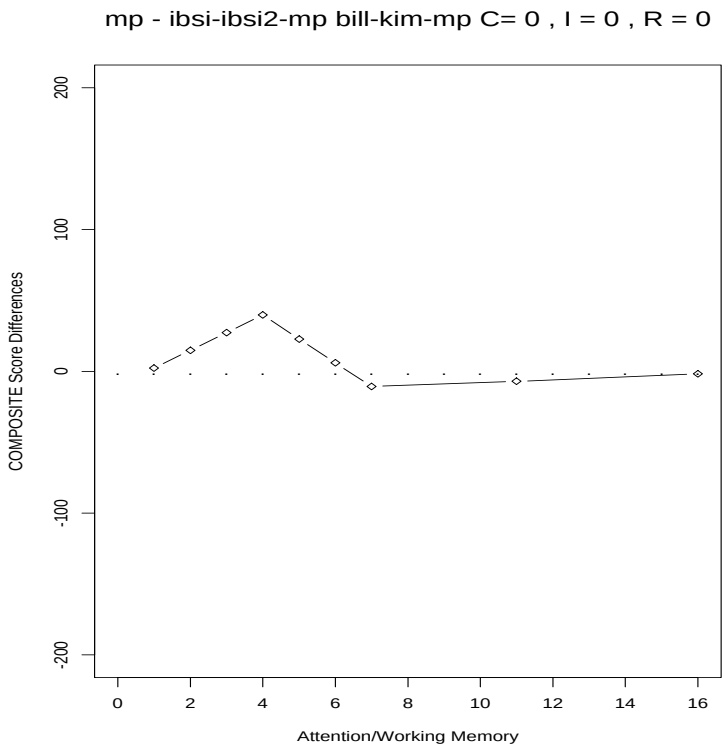


Figure 7.25: The Matched-Pair-Premise-Warrant strategy increases the number of Matched Pair Inferences. Strategy 1 is two Matched-Pair-Premise-Warrant agents and strategy 2 is two All-Implicit agents. Evaluation function = MP, commcost = 0, infcost = 0, retcost = 0

situations in which communication is costly, the reduction in retrievals might not be worth the additional communication cost.

7.6.4 Summary: Attention in Design-World

The Design-World experiments discussed in this section were based on the DISCOURSE INFERENCE HYPOTHESIS. I tested an opening strategy called Open-Best which makes premises salient that might be beneficial in means-end reasoning. This strategy was shown to be not beneficial. Section 7.6.2 presented the EXPLICIT WARRANT strategy, in which agents include the warrant for their proposal along with every proposal. This strategy was shown to be beneficial in achieving matching beliefs for the warrant for proposals, even when retrieval is free, as might be expected. In addition, this strategy reduces retrieval costs and thus provides a benefit whenever retrieval is not free. Finally, section 7.6.3 presents a strategy used in the MATCHED PAIR version of the task. This strategy makes premises explicit that might be useful in making matched pair inferences. Unlike the Open-Best strategy, the Matched-Pair-Premise strategy increases agents' ability to make matched

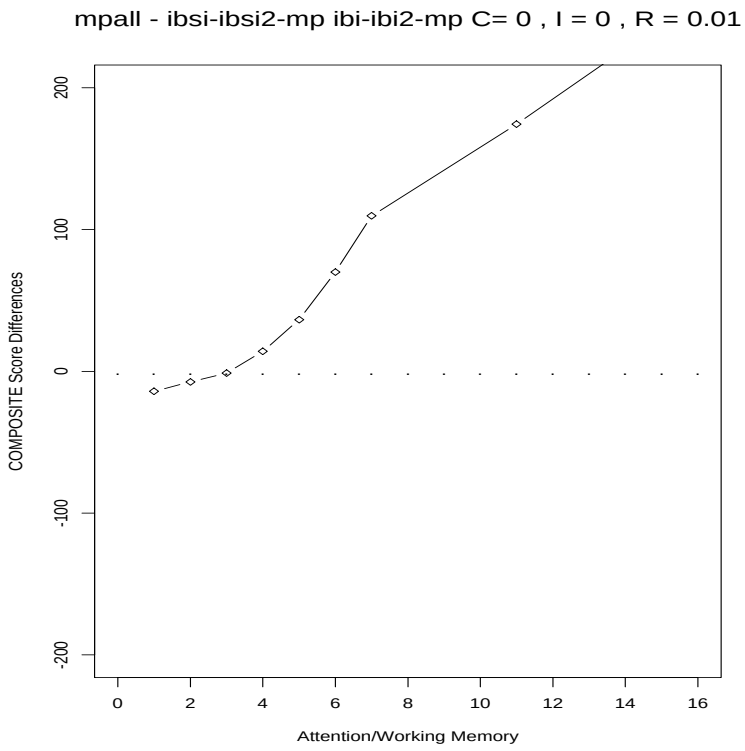


Figure 7.26: Strategy 1 is the Matched-Pair-Premise-Warrant strategy and Strategy 2 is the Matched-Pair-Premise strategy. Evaluation function = MPALL, commcost = 0, infcost = 0, retcost = .01. The Matched-Pair-Premise-Warrant strategy reduces retrieval

pair inferences. In sum, the main results are that Attention IRUs are beneficial when they are targeted specifically at particular inferences, when inference is complex, or when retrieval isn't free.

7.7 Attention: Summary

This chapter discussed the function of Attention IRUs. Attention IRUs are defined by the information status location parameter of Remote. Their function is to make a proposition salient, and the defining parameter simply picks out cases of IRUs whose antecedent is not currently salient.

Section 7.1.2 describes three hypotheses about the underlying cognitive processes which motivate Attention IRUs: the COORDINATION HYPOTHESIS, the RETRIEVAL CUE HYPOTHESIS and the DISCOURSE INFERENCE HYPOTHESIS. Althought, the analysis of the corpus does not eliminate any of these hypotheses, it shows that the DISCOURSE INFERENCE HYPOTHESIS is viable as the function of all distributional types of Attention IRUs. I argued that the function of Attention IRUs follows from the DISCOURSE INFERENCE CONSTRAINT, which states that inferences are derived from currently salient propositions. Because interpretation is highly dependent on what is salient, conversants

coordinate what is salient. Coordinating what is salient for both conversants reduces uncertainty about which inferences both conversants can derive. If agents were logically omniscient or if everything agents knew was always salient and available, then there would be no need for coordination, and no need for Attention IRUs.

Note also that Open Segment IRUs such as 76 appear to function similarly to the Deliberation IRU in 69. In each case a proposition is made salient to be used in reasoning. In the case of 76 the proposition is used in means-end reasoning, whereas in 69 the proposition is used in deliberation. Except for these differences in the type of reasoning, there seems little reason to distinguish between these two types of Attention IRUs.

Section 7.6 tests the DISCOURSE INFERENCE HYPOTHESIS for Open Segment and Deliberation IRUs and showed that Attention IRUs can be beneficial in three situations: (1) when the content of the IRU consists of information that is targeted at specific inferences that the agent is performing at that point in the discourse; (2) when inferential complexity is increased; and (3) when retrieval is not free.

Chapter 8

Consequence

8.1 Introduction

In chapter 1, I proposed that the function of Consequence IRUs is to support inferential processes. There are two classes of Consequence IRUs. First, some make an inference explicit from available premises, such as the inference in 85c in the context of 85a and 85b:

(85) a. If John goes to France, he'll miss Barbara's birthday.

b. John is definitely going to France.

c. THEN JOHN WILL MISS BARBARA'S BIRTHDAY.

This type of Consequence IRU is called an Inference-Explicit IRU. These are defined by the single distributional parameter that the HEARER OLD relation of the IRU to its antecedent is a type of inference, such as an implicature, entailment or other type of common-sense inference. The IRU makes an inference explicit that was implicit in what had already been said.

The second type of Consequence IRU are Affirmation IRUs, as shown in 86, previously characterized as supporting the inference of 'rhetorical contrast' (Horn, 1991).

(86) *I like you* Lizzy. I don't know *why I like you.* But I LIKE YOU. (CS, 3/4/92)

Unlike the Inference-Explicit IRUs, the content of an Affirmation IRU such as that in 86 is already explicit in the context. I will argue below that the affirmation of *I like you* is related to underlying processes of inference and deliberation.

Section 8.2 will discuss Affirmation IRUs and section 8.3 will discuss Inference-Explicit IRUs. These IRUs have little in common in terms of function. Affirmation IRUs have not been implemented in Design World, but Design-World experiments for Inference-Explicit IRUs when agents are logically omniscient or inference limited and when inferential complexity is increased will be discussed in section 8.4.

8.2 Affirmation IRUs

Affirmation IRUs function simultaneously at two levels: intentional and informational (Moore and Pollack, 1992; Walker, 1993a). At the intentional level, speakers affirm propositions for multiple reasons. At the informational level, it appears that Affirmation IRUs meet an information structure constraint of rhetorically contrasting with another salient proposition.

Example 86 illustrates a common schema for Affirmation IRUs in which a proposition Q is asserted, then a proposition P, which 'rhetorically contrasts' with Q is asserted, then Q is affirmed. The affirmation of Q is typically prefaced with the discourse particle *but*, which conventionally implicates that there is some contrast between P and Q.

Affirmation IRUs are not easily identifiable according to taxonomic criteria. The occurrence of *but* is a cue for Affirmation IRUs but some Affirmation IRUs are not prefaced by *but*. Figure 4.6 shows how Affirmation IRUs distribute in terms of HEARER OLD and SALIENCE parameters. The figure shows that Affirmation IRUs are not taxonomically selected by any of these parameters. Affirmation IRUs are selected by the occurrence of *but*, Argument Opposition and Focus/Open proposition constructions that rely on a Salient Set.

Unlike Deliberation IRUs, the proposition realized by an Affirmation IRU is frequently already salient. Figure 4.6 shows that the antecedents of 8 Affirmation IRUs are not currently salient, but 15 occur in the same turn as their antecedent and 8 of these are actually adjacent to their antecedent.[1] The remainder of this section will address the following issues in the analysis of Affirmation IRUs:

- Are there semantic/pragmatic felicity conditions on Affirmation IRUs and how should they be characterized?
- Why is P asserted and why Q is affirmed?
- Does the hearer need to recognize why P was asserted for the Affirmation to be felicitous?
- What cognitive effects might the Affirmation have?

[1] Adjacent-Self IRUs occur in the same turn and are adjacent to their antecedents.

First, in section 8.2.1, I will discuss a pragmatic/semantic felicity condition on Affirmation IRUs proposed in previous work (Ward, 1985; Horn, 1991; Ward, 1990). I will suggest that this condition needs to be extended to account for the full range of Affirmation IRUs. Then, in section 8.2.2, I will consider possible intentions that a speaker of an Affirmation IRU might have. I will argue that the communicative intention underlying Affirmation IRUs is to support inferential processes and the deliberation of beliefs and intentions. Because the function of Affirmation IRUs is complex, and because there is no natural use for Affirmation IRUs in Design-World, Design-World experiments on Affirmation IRUs are not reported here.

8.2.1 Rhetorical Contrast

8.2.1.1 Argumentative Distinctness

As briefly discussed in section 2.3, semantic/pragmatic felicity conditions on Affirmation IRUs have been proposed in previous work. Affirmation IRUs are related to a verb-preposing construction in 87 called Proposition Affirmation, first discussed by Ward (1985):

(87) Tchaikovsky was one of the most tormented men in musical history. In fact, one wonders how *he managed to produce any music at all.* BUT PRODUCE MUSIC HE DID. [WFLN Radio, Philadelphia] (Ward's 96, (Ward, 1990))

The proposition *produce music he did* is an IRU because *manage* is a semi-factive predicate which presupposes the proposition realized as the complement of manage, *Tchaikovsky produced music* (Karttunen, 1973; Gazdar, 1979).

Ward argues that this verb preposing subset of Affirmation IRUs are felicitous because they 'can convey a particular propositional attitude on the part of the speaker, one of *surprise* or *unexpectedness*' (Ward, 1985), p. 227. However, Horn (1991) points out that the 'surprise' condition is not necessary, since 88 is perfectly felicitous:

(88) It's unfortunate that *it's cloudy in San Francisco this week*, but CLOUDY IT IS – so we might as well go listen to the LSA papers.

Horn generalizes Ward's account to non-preposed cases of affirmation and claims that the condition that characterizes felicitous affirmation is RHETORICAL CONTRAST. Horn suggests that the typical schema for Affirmation IRUs is concession/affirmation: I concede P and affirm Q where Q may follow logically from P, but contrasts rhetorically with it. Furthermore, 'both presuppositions and

entailments, may be felicitously, though redundantly affirmed, if the affirmatum can be introduced with *but* rather than *and*', since *but* supports the conventional implicature that there is some contrast between P and the informationally redundant affirmation Q.

Horn proposes that many cases of RHETORICAL CONTRAST can be characterized via the ARGUMENTATIVE DISTINCTNESS CONDITION (Horn, 1991; Anscombre and Ducrot, 1983):

(89) ARGUMENTATIVE DISTINCTNESS CONDITION:
An informationally redundant affirmation Q will be discourse acceptable if it counts as ARGUMENTATIVELY DISTINCT from P in the sense that where P counts as an argument for a conclusion R, Q represents or argues for an opposite conclusion R'.

This formulation predicts when affirmation is not felicitous as shown in the examples from Horn and Ward in 90 and when it is, as exemplified by those in 91:[2]

(90) a. #It isn't odd that *dogs eat cheese*, {and/but} THEY DO (eat cheese)

b. #I know why *I love you*, {and/but} I DO.

c. #He doesn't regret that *he said it*, {and/but} HE DID SAY IT.

d. #*He won by a large margin*, {and/but} WIN HE DID.

(91) a. It's odd that *dogs eat cheese*, but THEY DO (eat cheese)

b. I don't know why *I love you*, but I DO.

c. He regrets that *he said it*, but HE DID SAY IT.

d. *He won by a small margin*, but WIN HE DID.

For example, 91d meets the condition of argumentative distinctness since P, *He won by a small margin*, can argue for an R such as the relative lack of a popular mandate for Mr. X, while the Q of his winning per se, *win he did* argues for the opposite conclusion (Horn, 1991).

Horn also points out that the condition of argumentative distinctness captures only a subset of the full range of cases involving rhetorical contrast, since 92 involves rhetorical contrast, which is not due to either surprise or argumentative distinctness (Horn's 31):

(92) a. I'm unhappy/#happy *they fired him*, but FIRE HIM THEY DID.

[2]Note that the embedding verbs here are factives (Gazdar, 1979).

b. I'm sorry/#glad *I said it*, but SAY IT I DID.

Horn suggests that perhaps examples like 92 support rhetorical contrast through the sociolinguistic edict that the speaker must try to avoid negative face (Brown and Levinson, 1987).

The argumentative distinctness condition is a good starting point for a semantic and pragmatic condition on felicitous affirmation, but leaves several issues open. In what follows I will first make this condition more precise so it is clearer what cases it covers and then suggest extensions to the characterization of rhetorical contrast.

First, it is not clear what it means for P to 'count as an argument for' R (Horn, 1989). The relation between P and R cannot be entailment, or speakers that produced affirmations would be contradictory. Second, the argumentative distinctness condition must be a condition on the hearer: (1) the hearer must recognize that P and Q present arguments for opposition conclusions, and (2) the hearer must be able to ACCOMMODATE some R, R', such that P counts as an argument for R and Q counts as an argument for R' (Lewis, 1979). Finally, there must be some restrictions on which R, R' can be accommodated. Below we will consider a restriction that the R, R' must be evoked in the discourse situation.

Some additional precision may be gained by noting that the argumentative distinctness condition represents an argument schema that can be subsumed under the Only Minor schema of Sadock's MODUS BREVIS, shown in figure 7.4. If we reformulate argumentative distinctness as a type of MODUS BREVIS, we see that the condition is met when there are two inferences supported in the context. The first consists of the minor premise P, and major premise P $\rightarrow$ R, supporting the conclusion R. The second consists of the minor premise Q, major premise Q $\rightarrow$ R', supporting the conclusion R'.

One possible inference rule is a trivial one of P $\rightarrow$ P or Q $\rightarrow$ Q. Thus we might also want to consider cases where P argues directly against Q rather than against an accommodated R'. In addition, since the inference rules of P $\rightarrow$ R and Q $\rightarrow$ R' cannot be rules of entailment, the inference rules in question must be commonsense, default or heuristic inference rules. We will show below that both SUPPORT and WARRANT relations can be the basis for these inference rules.

As in the MODUS BREVIS cases discussed in section 7.2.3: the hearer must be able to recognize the argument structure based on these default rules of inference. This means that it must be fairly easy to recognize or accommodate the fact that the rule applies or that the speaker thought that that the hearer would think that the rule applies (Cohen, 1987). With Affirmation IRUs, like Deliberation IRUs, the relevant inference rule may be currently salient, or the hearer may retrieve or accommodate it.

The following sections use the MODUS BREVIS formulation to argue for other ways of defining rhetorical contrast in addition to the argumentative distinctness condition. Section 8.2.2 will then discuss the range of intentions that motivate Affirmation IRUs.

8.2.1.2 Set-Based Contrast

Many cases of Affirmation IRUs are felicitous without requiring the hearer to retrieve or accommodate specific inference rules. For example, affirmation is felicitous whenever there is a set-supported contrast. In 93, Viv is talking about a recent vacation to Mexico where due to a drought, the hotel had trouble supplying water consistently to all the rooms.

(93) *We always had water* (in that room).
I think we were the only ones.
WE NEVER RAN OUT OF WATER.
hot water, we ran out of.
but WE ALWAYS HAD WATER. (Viv 3/20/92)

Viv repeats her main point, that *we always had water*, three times and each repetition is supported by a different set. First {*We, the others*} is enumerated with the affirmation and negation of having water. Then a second set, {*hot water, cold water*} is enumerated with the affirmation/negation of running out of it. The sets are evoked in the context by the use of the term *the only ones* and by the topicalization of *hot water* (Prince, 1981a).

Similarly the affirmation in example 94 is supported by a set of relevant retirement plans {*IRA, Keogh Plan*}.

(94) (30) H: Is he self-employed?
(31) L: Yes
(32) H: I'm sorry, i missed that
(33) L: Yes he is.
(34) H: Ok. Well *why not start for him a Keogh plan*? You can't get an IRA after 70, but YOU CAN GET A KEOGH PLAN

The set {*IRA, Keogh Plan*} is evoked by the parallel open propositions in (34). The fact that *you can't get an IRA after 70* doesn't necessarily argue against *You can get a Keogh Plan* at the logical

level. Rather {*IRA, Keogh Plan*} are alternate ways of achieving the same goal of avoiding paying taxes.

In 95, the participants must have agreed by utterance (9) that the R cannot claim his mother as a dependent. This proposition is affirmed by H in (17). It doesn't seem to be necessary to infer an R,R' as opposite conclusions from (15) and (17), since R could have both deducted the medical and dental care as well as claim his mother as a dependent.

(95) (6) R: ..... or uh *does that income from the certificate of deposit rule her out as a dependent*
(7) H: *yes it does*
(8) R: *it does*
(9) H: *yup, that knocks her out.*
now there is something you can do. do you support her in any way?
(discussion about whether r supports his mother)
(15) H: well the medical and dental care you can deduct, provided you can establish that you have provided more than half support.
(16) R: uh huh
(17) H: BUT THE DEPENDENCY YOU CANNOT CLAIM
(18) R: um hm (breath) I see.
ok. uhh, alright, the second question...

The affirmation contrasts with (15) because *dependency* and *medical and dental care* are members of a set of deductions. This set is evoked by the use of the topicalizations in (15) and (17) and is easily recognized by the hearer.

These are all classic examples of SET-BASED CONTRAST, defined by Prince as an inference arising under the following conditions (Prince, 1981b; Prince, 1986):

- There is an OPEN PROPOSITION (OP) taken to be salient shared knowledge,
- The OP is predicated on the members of a set of entities,
- the variable in the OP is instantiated differently for each member,
- the difference in the instantiation is considered relevant.

In the case of 94, the OP is *you can/can't get a Y* and the set of entities that the proposition is predicated on is {*IRAs, Keogh plans*}. The variable in the OP, affirmation/negation, is instantiated

differently for each member of the set, and the difference in the instantiation is relevant precisely because a way of putting aside money for retirement is what is under discussion.

When the set-contrast condition is met, *but* may not be needed as a marker of contrast. In 93, the first affirmation of *We never ran out of water* was not prefaced by *but*. The contrast here was induced by the phrase *the only ones*, similar to the use of *the only* in 96 from Schiffrin (1982):

(96) (1) See this one right here?
(2) He's smart.
(3) He himself don't think he's smart,
(4) but he's smart.
(5) *He came in first* in plumbing,
(6) out of a hundred thirty five,
(7) He was the only Jewish kid.
(8) HE CAME IN FIRST.

The affirmation in 96-8 is neither surprising nor argumentatively distinct. Prefacing (8) with a *but* completely changes the meaning, and indicates that (6) and (7) are somehow concessive. While there seems little to concede here, 96 is easily characterized as a case of set-based contrast in which the one child under discussion is contrasted with the one hundred and thirty five children in his class.

These examples show that Prince's characterization of the conditions for contrast provide an environment for felicitous affirmation that is independent of the argumentative distinctness condition.

8.2.1.3 Lack of Support

There are an additional class of examples in which the the conceded proposition P does not seem to argue against either the affirmed proposition Q or any interesting R' that could be inferred from Q. Consider 97, said by a speaker hiking through the woods.

(97) *Something has been through here.*
I don't know if it was a deer or what,
but SOMETHING HAS.

Since the argumentative distinctness condition leaves potential R and R' completely open, a possible default inference rule is *something has been through here* (Q) $\rightarrow$ *I know what it is* (R'). The

assertion of P, *I don't know if it was a deer or what* defeats the inference of R'. While this inference rule is possible, it does not seem plausible and the affirmation is felicitous without consciously accommodating the suggested R,R' when reading 97. The same argument holds for 98, asserted by a passenger in a vehicle in response to the driver's comment that the heavy traffic was unexpected:

(98) *There's something on fire* up there.
I can't see *what's on fire*,
but SOMETHING IS. (LW 6/12/92)

In 98, a possible R is *I'll tell you what's on fire.* The assertion of P, *I can't see what's on fire*, argues for ¬ R. Again this inference rule does not seem convincing. Furthermore, the original assertion of *Something is on fire* already argues for ¬ R if we assume that speakers provide as much information as possible that is currently relevant (Quantity Maxim). In other words, if the speaker had known what was on fire, she would have said so at the beginning rather than asserting the less informative *There's something on fire up there.*.

Another way of viewing these examples is that concession and affirmation function at the metalinguistic level. The relation between P and Q then is that P fails to provide support for the assertion of Q, nevertheless the speaker is asserting Q. Examples of this type are characterized by verbs referring to typical sources of evidence for propositions, e.g. *see, hear, say*, as well as mental state verbs such as *know, remember.* Consider 99:

(99) (25) M: So I just put that on the little part there on that new form there
(26) H: Well you'll have to list that as interest earned, and then you can knock that amount off, somewhere on schedule B.
There's a line for it. I don't remember what the line number is, but IT'S THERE

The statement that H doesn't remember the line number simply fails to provide support for the fact that there is a line number, but it doesn't argue against the existence of the line. Perhaps R' is a proposition such as *you will be able to locate the line number.* But the original assertion of *there's a line for it* would have supported R'.

Thus while it is possible in many cases to come up with a possible pair of default inference rules, these rules may not be plausible as ones that the hearer would recognize and it is not clear that the speaker intended the hearer to do so. I will return to these examples in section 8.2.2.1, when I discuss the intentions underlying the assertion of P and subsequent affirmation of Q.

8.2.1.4 Information Structure Constraints

In order for the ARGUMENTATIVE DISTINCTNESS CONDITION to hold, the hearer must be able to retrieve or accommodate two inference rules: P → R and Q → R'. Horn and Ward argued that the reason that 100 is infelicitous is that no such rules are possible.

(100) #*He won by a large margin*, {and/but} WIN HE DID.

However if we imagine a situation in which the hearer was part of the competition for a promotion and now must work for the winner, then 100 is much better. If the R of *he humiliated you* is made explicit in the dialogue as in 101, then the affirmation becomes completely felicitous.

(101) *He won by a large margin.*
He humiliated you.
But HE DID WIN.

This suggests that we need an additional condition on the application of argumentative distinctness that the inference rules in question must be salient in the context, or easily evoked or accommodated.

In sections 8.2.1.2 and 8.2.1.3, I argued that the argumentative distinctness condition was not **necessary** for affirmation, but so far it appears to be a **sufficient** condition for felicitous affirmation. However, information structure provides an additional critical constraint. As discussed above, 102a is felicitous. Surprisingly 102b and 102c are not.

(102) a. It's unfortunate that you failed, but you did.

b. # *You failed* unfortunately but YOU DID.

c. # Unfortunately *you failed*, but YOU DID.

d. *You failed*. It's unfortunate that *you failed*, but YOU DID.

What is the difference? The proposition P realized by 102b and 102c is identical to that in 102a and thus should be argumentatively distinct as well. The only difference is that in 102a, the proposition *you failed* is backgrounded in the first clause, and the lexical items *unfortunate* and *did* are both marked as focal by pitch accents. Both syntactic structure and prosody contribute to establishing a focus/open proposition information structure where the focus is a propositional attitude. In the case of 102a this structure is UNFORTUNATE(P) but TRUE(P).

In 102b and 102c the proposition P is part of the focus because it is not subordinated and must receive a pitch accent. The lexical item *unfortunately* sounds parenthetical rather than focal. The effect is that P cannot be felicitously affirmed in an adjacent clause. Note that 102d shows that even if the proposition was previously asserted, the affirmation is still felicitous as long as it is not asserted in the concessive clause.

Further evidence against the sufficiency of the argumentative distinctness condition is provided by the asymmetry in the pairs below:

(103) a. *John managed to win*, but IT WAS DIFFICULT.

b. #*John managed to win*, but HE DID WIN.

In 103a, the contrastive structure is TRUE(winning) but DIFFICULT(winning). In 103b, even though *manage to X* conventionally implicates the difficulty of X'ing, this implicature does not provide the basis for a DIFFICULT(winning) but TRUE(winning) rhetorical contrast. Presuppositions of a proposition are inferences from a proposition, but not part of the focal structure of the proposition.

These examples support the conclusion that there is an additional information structure constraint on felicitous affirmation and that it is possible to affirm propositions that are in the background in a previous utterance with appropriate prosody.[3] Further elucidation of these constraints must be left to future work.

8.2.2 Deliberation and Inference

The previous section discussed the semantic/pragmatic felicity conditions on felicitous affirmation. These constraints hold at the information structure level. This section proposes speaker intentions underlying Affirmation IRUs. What I would like to explain is why the speaker asserts the concessive proposition P as well as affirming Q. The intentions that I argue for are: (1) to support deliberation, (2) to defeat undesired inferences, and (3) to indicate commitment to an asserted proposition. These intentions do not map isomorphically onto the information structure classes of Affirmation IRUs (Moore and Pollack, 1992).

8.2.2.1 Coherence of Beliefs

In chapter 1, I introduced the ATTITUDE ASSUMPTION:

[3]Levinson noted that 'some presuppositions and even entailments may be reinforceable with heavy stress' (Levinson, 1983).

ATTITUDE ASSUMPTION:

Agents deliberate whether to ACCEPT or REJECT an assertion or proposal made by another agent in discourse.

Because of the ATTITUDE assumption, speakers want to strategically provide information which they believe will make it more likely that other agents will accept their assertions and proposals. Increasing the likelihood of acceptance is related to the underlying deliberation process. This is reflected in the COHERENCE ASSUMPTION repeated here from section 3.3:

COHERENCE ASSUMPTION:

Beliefs and intentions that are subject to deliberation are evaluated for their coherence with other beliefs via the relations of SUPPORT and WARRANT.

The COHERENCE ASSUMPTION explains why P is asserted; P is asserted to ensure coherence via consistency in beliefs by defeating unintended inferences, to indicate the basis for assertions or proposals, or to compare two possible bases for intentions, or two possible bases for beliefs.

But once P is asserted why is Q affirmed? The AFFIRMATION ASSUMPTION below posits a plausible function for the affirmation of Q:

- AFFIRMATION ASSUMPTION: **Repeating** a proposition is a weak type of SUPPORT that provides evidence of the speaker's commitment to the truth of the proposition.

The AFFIRMATION assumption means that the occurrence of an affirmation is a cue that the speaker believes that s/he must provide additional SUPPORT for a proposition Q recently asserted. This speaker belief could be motivated by the perception that a currently salient proposition P argues against Q (Ward, 1990; Horn, 1991). In other words, if a proposition Q is affirmed in a context C, something in C must either support or warrant an opposite conclusion, or fail to support or warrant Q. This would motivate example 87 as defeating the common-sense inference that torment leads to unproductivity by affirming productivity. The following section will discuss how these examples are related to deliberation.

8.2.2.2 Supporting or Demonstrating Deliberation

According to the ATTITUDE ASSUMPTION both beliefs and intentions are deliberated. If the ATTITUDE ASSUMPTION didn't hold then there would be no need to support assertions or to provide warrants for proposals. Often the best support or warrant that a speaker can provide is an AFFIRMATION of the relevant fact in the face of other information. Affirmations motivated by SUPPORT

are characterized by verbs referring to typical sources of evidence for propositions being deliberated, e.g. *see, hear, say*, as well as mental state verbs reflecting deliberation, *know, remember*. Affirmations motivated by WARRANT refer to intentionality, costs, or benefits of a course of action.

In 98, the speaker perceives a need to support the assertion that something is on fire. The concessive clause concedes lack of support, and the affirmation clause demonstrates her commitment to the truth of her assertion, despite the fact that she can provide no evidence to support her claim. Typically the kind of evidence the speaker would like to have to support the assertion of Q is made explicit in the context. In 98 the speaker stated that visual support was desirable. In 99, H is searching his memory/belief structure.

The most convenient way to represent the structure of examples such as 98 is with a relation of NOT-SUPPORT, since the assertion of *I don't know why I like you* does not strictly argue against *I like you*. A similar example was given in 97. In example 104, the relevant relation seems to be NOT-WARRANT:

(104) He didn't make a profit *from doing it*, but HE DID IT.

Deliberating over alternate courses of action can take many kinds of information into account. In 105, the speaker concedes that efficiency might dictate one course of action but preference warrants another.

(105) *I don't like to go down that way.* // It may be shorter, // but I DON'T LIKE IT. (L 3/10/92)

The source of the argument structure in 105 is intentional rather than informational. Different types of information structure can also be related to deliberating intentions. For example, in 94, the fact that a plausible alternate is not possible, *You can't get an IRA after 70* is an argument for getting a Keogh plan. In 106, the Affirmation is only indirectly related to the concession at the information structure level, but provides the basis for comparing possible intentions in what follows.

(106) H: Sure, if you pay off, they give you all kinds of numbers, that if you pay the the mortgage off early, you p.. you make a principle payment now, you save umpteen dollars in interest. That's great, but THEY'RE TALKING ABOUT DOLLARS. Let's compare comparable things. Let's compare the percentage. The percentage that you're making on your money, if it exceeds 14 %, you don't pay off the mortgage.

In 107, J demonstrates what supporting information she is using to deliberate whether to file a tax return.

(107) H. Jennifer I understand what you're saying and I'm sorry I have to tell you that, I really am.
J. Well, I'm, I have more faith in you than what he told me,
HE SAID I DIDN'T HAVE TO FILE,
BUT THEN YOU JUST TOLD ME I DO
H. Yes. and I wouldn't want to see you get in trouble.

J has received verbal advice from two different sources, and the IRUs provide evidence that the **source** of these two opposed beliefs is what is determining the result of deliberation. It appears that J's intention is to make explicit how she is resolving the inconsistency between two pieces of linguistic evidence.[4]

8.2.2.3 Negative Face as Preference

Horn's negative face examples such as 88 are not well characterized by the argumentative distinctness condition. However, these can be subsumed under the COHERENCE ASSUMPTION if we augment the endorsement types for beliefs to reflect agents' preferences (Kahneman, Slovic, and Tversky, 1982; Galliers, 1991a):

- PREFERENCE ASSUMPTION: Agents' beliefs are partially determined by their **preferences** about what to believe, which may have a nonlogical or nonutility basis.

Then, a combination of the ATTITUDE and PREFERENCE assumptions motivate 88, where what is relevant is that the speaker believes that the hearer may not want to accept the assertion of P, preferring to believe ¬ P. It is possible that P in 88 conflicts with the hearer's view of herself as extremely intelligent, or that the acceptance of P would lead the hearer to infer a number of conclusions which she would prefer not to derive. Factive predicates for other relations that express the difficulty of accepting a proposition are *odd, strange, surprising, amazing, I'm sorry that, It's a wonder that*, and all of these also license affirmation.

8.2.2.4 Defeating Inferences

The ARGUMENTATIVE DISTINCTNESS CONDITION suggests that one of the speaker intentions underlying Affirmation can be to defeat inferences. If the relationship between P and Q can be

[4]This may show that it is important to include source information as an endorsement type on supporting beliefs (Galliers, 1990; Cohen, 1985). On the other hand, perhaps J must demonstrate her reasoning precisely because it is difficult to choose between two oppositing beliefs that are both supported by a linguistic endorsement.

characterized by argumentative distinctness, where P counts as an argument for a conclusion R, and Q argues for an opposite conclusion R', one intention that the speaker could have is to defeat the inference of either R or R'. Q is affirmed in order to defeat the inference of R or to make it clear that Q holds despite the fact that R' doesn't follow.

We discussed argumentation-based inferences, but defeating plausible inferences of GENERALIZATION is also possible. Consider 108 and 109:

(108) *Sabine doesn't make syntax mistakes in English*, maybe in German, BUT NOT IN ENGLISH. (LL 5/93)

(109) *The agent they have right now is pretty cute.* It can't do much, but IT'S CUTE. (LW 11/6/93)

In neither of the discourses that these examples are drawn from were the conceded propositions of *Sabine may make syntax mistakes in German* or *The agent can't do a lot* controversial or up for discussion. Rather, in 108 it seems that the speaker intends to keep the hearer from inferring something like *Sabine doesn't make syntax mistakes at all*, and in 109 the capabilities of the agent under discussion is related to its appearance by a set-relation of desirable attributes. The speaker believes that if she asserts that the agent is cute, the hearer may infer (falsely) that the agent is both cute and can do a lot, or that it is cute because it can do a lot.

A diagnostic for whether a felicitous affirmation depends on defeating inference is to paraphrase the concessive assertion P and the affirmation Q, with *Don't think I mean not P and Q*. For example in 108, the paraphrase would be *Don't think I mean that Sabine doesn't make syntax mistakes in German or in English*. This diagnostic distinguishes the defeating inference Affirmations from the COHERENCE affirmations. Examples such as 98 cannot be paraphrased as: *Don't think I mean that I can see what is on fire and that something is.*

Generalization inferences are similar to examples which bound a scale (Horn, 1989; Hirschberg, 1985), i.e. don't infer too much in either direction of a scale.

(110) (22) b. Are there ah .. I don't think *the ah brokerage charge* will be ah that excessive.
(23) h. No *they*'re not excessive but THERE ARE CHARGES

In 110, a scale is evoked of *no charges, some charges, excessive charges.* The affirmation bounds the scale of charges to be less than some charges but more than no charges. This bounding the scale type of affirmation is also demonstrated by the examples in 111:

(111) a. It's snowing more gently, but it's still snowing. ((Horn, 1991))

b. This happens to me sometimes, not too often, but it does. (2/22/93 Van Pelt)

c. There's also another kind of turtle in here, with a red belly. They are pretty rare, but they ARE in here.(Tinticum)

8.2.2.5 Salience and Rehearsal

A final possibility is that Q is affirmed in order to make Q salient. However there is only one sentence after the first assertion of Q, so by our criteria Q is still salient at the point where it is affirmed. It is plausible that Q is affirmed because the last thing a speaker says in a turn is more salient than earlier assertions, and the speaker wishes to leave Q salient because Q is the 'main point'. It is also plausible that affirmation simply increases the frequency of a proposition in memory. Both of these effects follow automatically from the recency and frequency effects of a memory model like AWM, hold without any recognition by the hearer, and are a general function of all types of IRUs rather than being specific to Affirmation IRUs.

8.2.3 Summary

Affirmation IRUs are subject to information structure constraints and can be characterized at the information level independently of the speaker's intention in producing the Affirmation. They provide evidence that the reasoning process of human beings is affected by their desires and by the desire to maintain coherence among all their beliefs (Harman, 1986; Galliers, 1990). Speakers take care to ensure that hearer's don't draw inconsistent inferences. In addition, Affirmation IRUs show that the hearer's attitude toward a proposition can be a reason not to accept or believe a proposition, even when this attitude has a nonlogical basis. Experiments which test the proposed functions of Affirmation IRUs will not be reported here. The Design-World experiments for Consequence will test the Inference-Explicit IRUs to be discussed in the following section.

8.3 Inference-Explicit IRUs

While the analysis of Attitude and Attention IRUs uses factors related to the utterance context, Inference-Explicit IRUs are classified as such based solely on the HEARER OLD relationship between the IRU and its antecedent. The types of HEARER OLD information discussed in section 4.3 are REPETITION, PARAPHRASE, ENTAILMENT, IMPLICATURE and PRESUPPOSITION. Inference-Explicit

IRUs are those which make an inference explicit: the content of the IRU is derivable as an entailment or an implicature based on linguistic form.

The class of Inference-Explicit IRUs is not homogeneous. One distributional factor that seems to important is whether the antecedents of the IRU are currently salient. This distinction defines two major subclasses of Inference-Explicit IRUs; these will be discussed in the following sections.

8.3.1 Inference-Explicit IRUs with Salient Antecedents

A premise for an inference is **salient** if its antecedent location was Last, Same, or Adjacent.[5] While apriori, we would not expect Inference-Explicit IRUs to be distributionally different than Paraphrases, figure 8.1 shows that the antecedent for a Inference-Explicit IRU is more likely to be salient ($\chi^2 = 4.835, p < .05, df = 1$).

	Inference-Explicit IRUs	Paraphrase IRUs
Salient	24	43
Not Salient	8	39

Figure 8.1: Distribution of Inference-Explicit IRUs as compared with Paraphrases, according to whether their antecedents are currently salient; Salient = Adjacent or Same or Last, Not Salient = Remote

This distributional fact provides evidence for the DISCOURSE INFERENCE CONSTRAINT because in most cases when there is explicit evidence that an inference was derived from an assertion in conversation, the antecedent premises are currently salient.

8.3.1.1 Some Inference-Explicit IRUs are Attitude IRUs

Inference-Explicit IRUs whose antecedent was said by the other speaker and which are adjacent to their antecedent, are a subset of Attitude IRUs, with an additional effect of making an inference explicit. As shown in figure 8.2, 17 out of the 32 Inference-Explicit IRUs with salient antecedents are also Attitude IRUs.

An example of a Inference-Explicit IRU that is in the ATTITUDE LOCUS is given in 112-20:

(112) (13) ..... What did you do with the house in north Jersey?

(14) A: I sold it

[5] See 4.3 for the definition of these parameters.

	Attitude IRUs	Not Attitude
Inference-Explicit IRUs (Salient Ants)	15	9

Figure 8.2: Distribution of Inference-Explicit IRUs with Salient antecedents as Attitude IRUs and Not Attitude; Attitude = Adjacent & Other

(15) H: No, was it rented after you moved out?

(16) A: No.

(17) H: Give me the sequence here.

You built a house in south Jersey 5 yrs ago?

(18) A: Yes.

(19) H: Was that rented?

(20) A: No. I was using it as a summer home.

(21) H: ah THEN AT ONE TIME YOU HAD 2 HOMES.

(22) A: Right.

Here, H makes the inference explicit that if A owned a home in north Jersey and a home in south Jersey simultaneously, and neither home was rented, then A had two homes. A's response in 112-22 indicates that she treats this as a demonstration by H that he had made this inference. Thus in this case the Inference-Explicit IRU also functions as an Attitude Acceptance IRU.

8.3.1.2 Inference-Explicit IRUs Ensure that an Inference is made

As shown in figure 8.2, there are 9 Inference-Explicit IRUs with salient antecedents that are not Attitude IRUs. In these, the speaker explains or draws out an inference based on his/her own previous contribution.[6] The inference is presumably made explicit in order to ensure that the other conversant makes the inference. For example consider 113:

(113) (27) C: See my comment was *if we should throw even the 2000 dollars into an I R A* or something for her

....

(other effects of the hypothetical course of action)

.....

(36) H: As far as your wife is concerned, she's entitled to an IRA for 81. However in 81,

[6]These are elaborations in (Hobbs, 1979; Mann and Thompson, 1987).

it's 15 percent.[7] SO SHE'D HAVE A MAXIMUM, IF SHE RECEIVED TWO THOUSAND DOLLARS OF THREE HUNDRED THAT SHE PUT INTO IT.

(37) C: um hmm..

In this case, H makes the inference explicit that 15 percent of two thousand dollars is 300 dollars. It is plausible that H thought that by making the inference explicit he would save C making the calculation, and that this would be more efficient since he had already calculated the amount. It is also possible that H thought that C was not capable of making the calculation while talking on the radio. It is not very plausible that H thought that C was not capable of making the calculation at all. In the worst case C could make the calculation if given time to access a pocket calculator.

8.3.1.3 Does Sequential Locus matter?

Consider the Close-Segment locus of the Inference-Explicit IRU in 114-26a:

(114) (20) H: Right. The maximum amount of credit that you will be able to get will be 400
that they will be able to get will be 400 dollars on their tax return
(21) C: *400 dollars for the whole year*?
(22) H: *Yeah it'll be 20%*
(23) C: *um hm*
(24) H: Now if indeed they pay the $2000 to your wife, that's great.
(25) C: um hm
(26a) H: SO WE HAVE 400 DOLLARS.
(26b) Now as far as you are concerned, that could cost you more.
(26c) Remember you're gonna have 2 thousand worth of income.
(26d) What's your tax bracket? .....

In 114-26a, H sums up the conclusion of the previous segment, which is that pursuing a potential course of action could result in a gain of 400 dollars. Even though the inference recapitulates what was made explicit earlier, the benefit of this strategy could arise from rehearsing the consequences of a given action, or from indicating that this consequence is relevant for the following segment. In 114, the fact that is repeated will be used in deliberation in the following segment, where a different hypothetical course of action is discussed.

[7] Earlier H had said that in 1982 the percentage deduction was 20 percent.

Does it matter that this IRU occurs at the transition to a new segment. IRUs may be a cue to discourse structure. The potential multiple functional interpretation of IRUs which are distributionally located at discourse segment boundaries was also discussed in chapter 7, where I conjectured that they may be used to indicate task structure when the task structure is unclear to one of the conversants. In 114 however, H is focusing on deliberation structures of one course of action as compared to another and 'rehearsing' the effects of one course of action is an equally valid explanation.

8.3.2 Inference-Explicit IRUs with Non-Salient Antecedents

Figure 8.1 shows that there are 6 cases of Inference-Explicit IRUs with Non-Salient (Remote) antecedents. If the DISCOURSE INFERENCE CONSTRAINT is to hold, then these inferences must have been made when the premises were salient. Therefore, it is not clear whether making an inference explicit when the premises are no longer salient functions differently than a paraphrase or a repetition. I will argue that the intentions underlying the Remote Inference-Explicit IRUs are the same as those for Attention IRUs.

Five of the Remote Inference-Explicit IRUs are introduced into the context and then used as a premise in reasoning. For example, in 115-12, the IRU makes an inference explicit which is then used for deliberating alternate courses of action.

(115) (7) J: ..... and I'm entitled to a lump sum settlement which would be between 16,800 and 17,800 or a lesser life annuity. and *the choices of the annuity um would be $125.45 per month.* That would be the maximum with no beneficiaries
(8) H: You can stop right there: take your money
(9) J: Take the money.
(10) H: Absolutely.
(11) J: How would.....
(12) H: YOU'RE ONLY GETTING 1500 A YEAR. At 17,000, no trouble at all to get 10 percent on 17,000 bucks. .....

H interrupts J at 115-8 and tells her to *take your money.* To provide a WARRANT for this course of action, H makes an inference explicit in 115-12. This inference follows from what she has told him in 115-7, namely *You're only getting 1500 (dollars) a year.* Presumably given enough time J could calculate that $125.45 a month for 12 months amounts to a little over $1500 a year. In addition to making the inference explicit, and thus potentially saving J this calculation, the juxtaposition of this fact against the advice to *take the money* licenses the inference that the fact that *she is only getting 1500 dollars a year*, is a WARRANT for adopting an intention to *take the money.* The

motivation for H's utterance, and the adoption of the advice relies on the assumption that J is deliberating about the utility of performing certain actions as opposed to others. The relevant comparison needed to support deliberation, that it is easy to get 10% on 17 thousand dollars, is also made explicit.

The other Remote Inference-Explicit IRU occurs in a summary at the end of a discourse segment. This may function to coordinate a discourse segment closing as discussed in chapter 7.

The function of reintroducing a proposition into the context and closing a discourse segment are both functions that are discussed as functions of Attention IRUs in chapter 7. So Remote Inference-Explicit IRUs may function exactly like remote repetitions or paraphrases. In order to explore the distinction between making an inference explicit while the premises are salient and making the inference explicit when the premises are no longer salient, Design-World experiments test Inference-Explicit IRUs in both of these situations. These experiments are reported in section 8.4.4.

8.3.3 Recognition of Inference-Explicit

Inference-Explicit IRUs are identified for the purposes of analysis by the fact that they make inferences explicit. However, the fact that the analyst can identify the utterance as a Inference-Explicit IRU is separate from whether in fact the hearer **recognizes** that the utterance makes an inference explicit. While my intuition is that hearers do recognize the Inference-Explicit relation between the Inference-Explicit IRU and its antecedents, there is little evidence to support this intuition.

One source of evidence is that there are no cases where the inference that is made explicit is argued about, i.e. Inference-Explicit IRUs are never controversial assertions.

Another source of evidence is that a plausible strategy is for the speaker to explicitly mark an utterance as inferrable whenever s/he thinks the hearer cannot recognize the inference relation (Pierrehumbert and Hirschberg, 1990; Polanyi and Scha, 1984; Schiffrin, 1987). This could be done via use cue words or prosody. Inference-Explicit IRUs are not reliably marked as inferrable by prosody (Walker, 1993c). Cue words that might mark an inferential conclusion are *then* and *so*, but these only occur on 9 out of 39 Inference-Explicit IRUs and the inferences are not ones that are especially difficult.

8.3.4 Summary

A functional distinction between types of Inference-Explicit IRUs is whether the antecedents for the IRU are salient or not. A second distinction among IRUs with salient antecedents is whether the speaker of the IRU said the antecedent or whether the antecedent was said by the other conversant.

Inference-Explicit IRUs normally have salient antecedents. When the antecedent is both Adjacent and produced by the Other speaker then Inference-Explicit IRUs also function as Attitude IRUs. The distributional analysis cannot distinguish those cases where the speaker made an inference explicit to help the other conversant make an inference and when the inference is made explicit as a demonstration that the speaker made the inference.

When Inference-Explicit IRUs have Remote antecedents, they function like Attention IRUs. In order to test whether this is the case, Design-World experiments will test Inference-Explicit IRUs in two different situations: when the antecedents of the Inference-Explicit IRU are currently salient and when they are not.

8.4 Consequence in Design World

As I pointed out in section 5.8, only a limited number of inference types are possible in Design-World. In the simplest version of the Design-World task there are three types: (1) means-end reasoning, ie. what actions can be performed to achieve the task; and (2) act effect inferences, ie. once an agent uses a piece of furniture in the task then that piece of furniture is no longer available; (3) inference of acceptance of a proposal when acceptance is implicitly communicated.

As discussed before, processing can be data-limited or inference limited. The variations in AWM provide a data-limit on inference because agents can only make inferences on facts that are currently salient. Limitations in means-end reasoning are a result of data-limitations. Every strategy that is tested in Design-World is tested against variations in AWM from 1 to 16.

In addition, as discussed in section 5.9.2, inference-limited processing is possible by parameterizing agents' inference capabilities. Agents can be parameterized so that (1) they are ALL-INFERENCE agents, ie. logically omniscient; or (2) they are HALF-INFERENCE agents, or (3) they are NO-INFERENCE agents. The inferences that are limited are the act-effect inferences.

Design-World also supports two other versions of the Standard task which increase inferential complexity: Matched-Pair and Matched-Pair-Two-Room as discussed in section 5.8. These tasks require making inferences about when two actions will achieve a matched pair in addition to the standard task and are the only inferences in Design-World that require multiple premises. The two

task variations supports exploring the effects of inference-explicit IRUs that have salient antecedents as opposed to those with remote antecedents, The task variations and the Matched-Pair-Inference-Explicit strategy are discussed in section 8.4.4.

The purpose of the simulations is to explore the interaction of discourse strategies with these inference related parameters. As discussed in section 5.8, within the limits of the Design-World domain and the discourse structure, there are a limited number of strategies available. In addition to the limits on types of inferences, there are only a limited number of discourse positions where inferences can be made explicit. These discourse positions are Proposal, Acceptance, or Closing.

As discussed in section 8.3.1, inferences in the corpus are most often made explicit while their antecedents are salient. Thus inferences that follow from the acceptance of a proposal can be made explicit either as part of the Acceptance or as part of an explicit Closing statement for the segment. Since there seems little difference between making an inference explicit in these two situations, a discourse strategy of making an inference explicit in a Closing was tested. This is the Close-Consequence strategy. Section 8.4.1 discusses this strategy and applies it in two situations: when agents are inference limited and when they are not.

The Matched-Pair-Inference-Explicit strategy makes inferences explicit as part of a proposal. One of the constraints on inferences that are part of a proposal is that the inference follows from the acceptance of a prior proposal. This is because no inferences follow from a proposal until it is accepted. The MATCHED-PAIR-INFERENCE-EXPLICIT strategy is presented in section 8.4.4.1 and results of testing this strategy are given in sections 8.4.4.2 and 8.4.4.3. The strategy is tested in situations in which an inference is made explicit when its premises are salient and when they are not.

8.4.1 Close-Consequence Strategy: Making Inferences Explicit

The discourse strategy of making an inference explicit in a closing statement is called the Close-Consequence strategy. This strategy of making inferences explicit at the close of a segment parallels the naturally occurring example given in 114. In both cases an inference is made explicit that follows from what has just been said, and the inference is sequentially located at the close of a discourse segment.

Figure 8.3 shows the discourse action protocol for a Close-Consequence. Openings and Acceptances are left implicit, but Closing is always communicated explicitly. Figure 8.4 shows a dialogue sequence composed of the interaction of a Close-Consequence agent (Agent B) with an All-Implicit agent (Agent A).

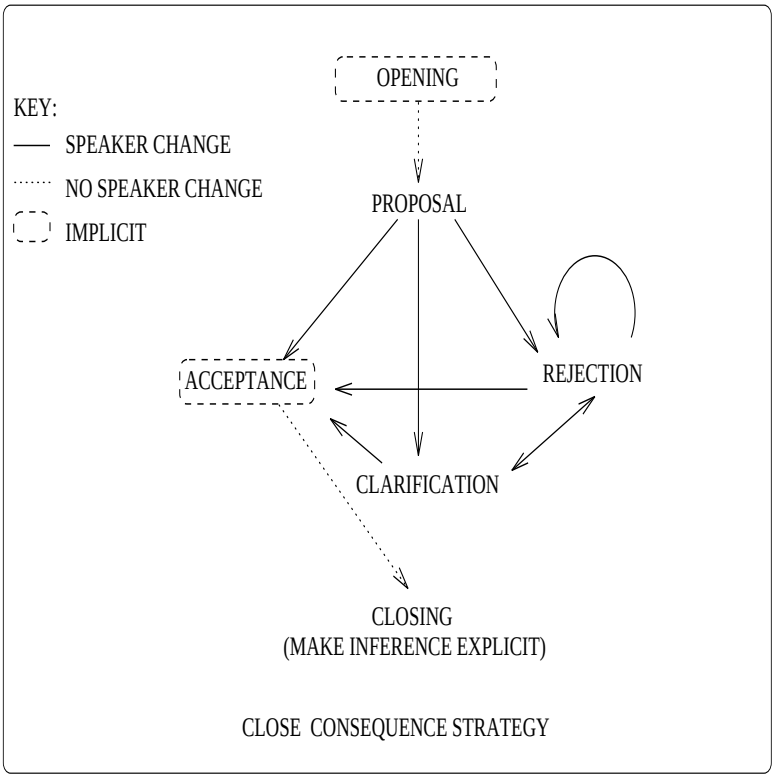


Figure 8.3: Discourse Action protocol for agents using the Close-Consequence strategy

Agent A	Agent B
PROPOSAL 1	
	CLOSING 1
PROPOSAL 2	
	REJECTION 2,3
PROPOSAL 4	
...	...
PROPOSAL N	
	CLOSING N

Figure 8.4: Sequence of Discourse Acts in Dialogue 116 for an All-Implicit agent A and a Close-Consequence agent B

An excerpt of a dialogue where agent CLC uses the Close-Consequence strategy is shown in 116. Agent CLC is a CLose-Consequence agent; CLC makes explicit CLOSING statements on the completion of the intention associated with a discourse segment. In addition to the explicit close statement as shown in 116-2, CLC makes the act-effect inference explicit that follows from what

was just agreed upon. For example, in 116-3, CLC makes the inference explicit that since they have agreed on putting the green rug in the study, Bill no longer has the green rug.

(116) BILL: *Then, let's put the green rug in the study.*
1:(propose agent-bill agent-clc option-30: put-act (agent-bill green rug room-1))
CLC: *So, we've agreed to put the green rug in the study.*
2:(close agent-clc agent-bill intended-13: put-act (agent-bill green rug room-1))
CLC: AGENT-BILL DOESN'T HAVE GREEN RUG.
3:(say agent-clc agent-bill bel-48: has n't (agent-bill green rug))
BILL: *Then, let's put the green couch in the study.*
4:(propose agent-bill agent-clc option-45: put-act (agent-bill green couch room-1))

Figure 8.5 shows the relationship between the underlying cognitive processes and the discourse and utterance acts for an dialogue excerpt given in 116 between an All-Implicit agent and a Close-Consequence agent.

Discourse Act	Utterance Act	A PROCESS	B PROCESS
		ME-REASON Put-1 Deliberate Propose, All-Imp	
PROPOSAL	1:(propose A B option-30)		
			ME-REASON Put-1 Check Preconds 30 Deliberate Decide to Accept Store (MS intend 30) Store (Act-Effects 30) Close, Cl C
CLOSING	2:(close B A intended-30)		
		Infer Acceptance 30 Store (MS intend 30) Store (Act-Effects 30) Close Put-1	
	3:(say B A bel-48)		Store Bel-48
		Store Bel-48 ME-REASON Put-2 Deliberate Propose, All-Imp	
PROPOSAL	4:(propose A B option-45)		
	... .		

Figure 8.5: Cognitive Processes for Sequence of Discourse Acts in Dialogue 116: Agent A is All-Implicit and Agent B is Close-Consequence

The following sections report results from testing the CLOSE-CONSEQUENCE strategy in two different

discourse situations. Section 8.4.2 discusses the results of evaluating this strategy with logically omniscient agents. Section 8.4.3 presents experiments that test the CLOSE-CONSEQUENCE strategy when agents are inference-limited.

8.4.2 Close-Consequence for Logically Omniscient Agents

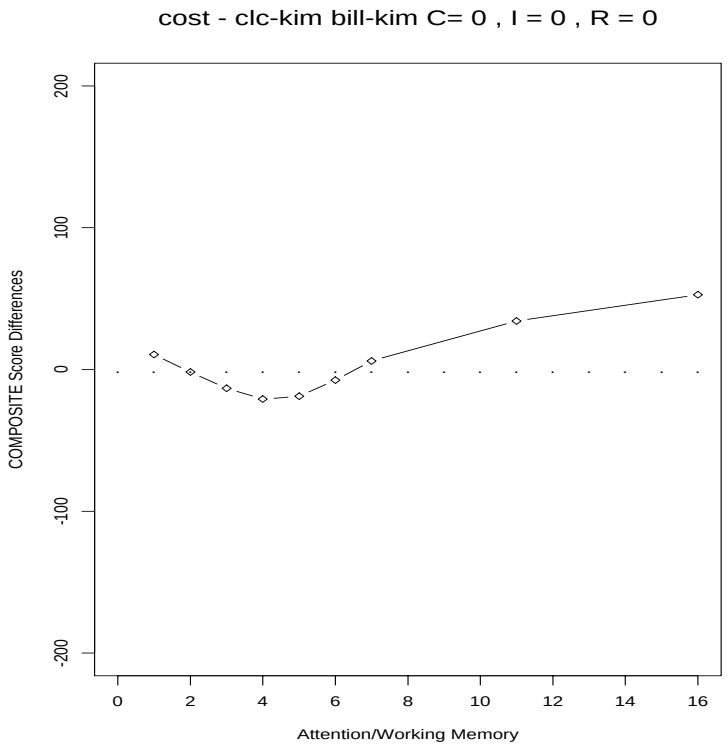


Figure 8.6: Close-Consequence Detrimental at Low AWM. Strategy 1 is the combination of an All-Implicit agent with a Close-Consequence agent and Strategy 2 is two All-Implicit agents, Task Definition = Standard, commcost = 0, infcost = 0, retcost = 0

8.4.2.1 Close-Consequence Detrimental at Low AWM

Figure 8.6 shows a difference plot for dialogues with CLC (Close-Consequence) and the All-Implicit agent Kim and two All-Implicit agents, Bill and Kim. The figure shows that the Close-Consequence strategy is DETRIMENTAL at lower AWM ($KS(4,5) > .19$, $p < .05$), even if all processing is free. This strategy displaces facts about which furniture pieces the agents have, which lowers the number of pieces that they are able to include in their plan, and thus has a detrimental effect at lower AWM. However, the strategy is BENEFICIAL at high AWM settings ($KS(11,16) > 0.52$, $p < .001$).

Why does Close-Consequence help at high AWM? The benefit of this strategy is in reducing the number of invalid steps. When agents can retrieve large amounts of information, they sometimes

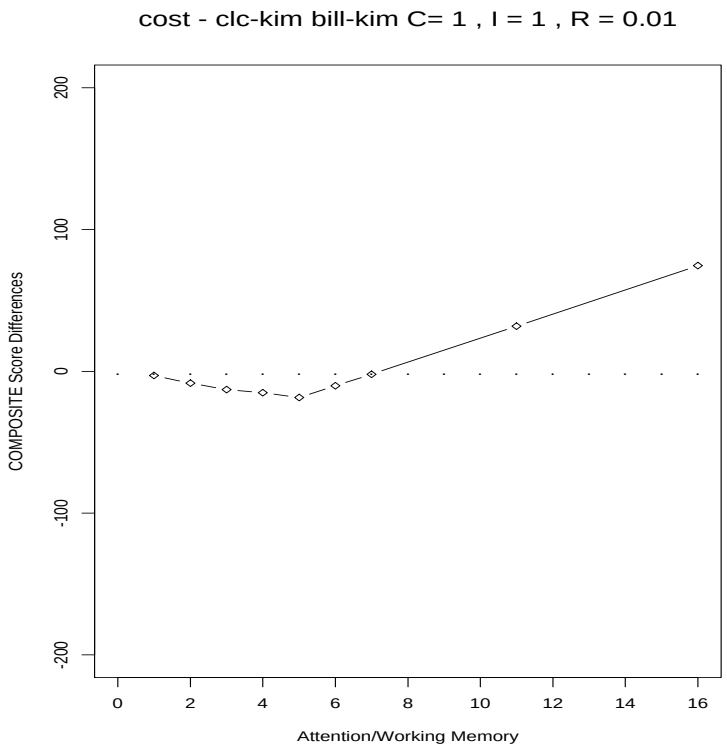


Figure 8.7: Processing Costs can Eliminate Benefits. Strategy 1 is the combination of an All-Implicit agent with a Close-Consequence agent and Strategy 2 is two All-Implicit agents, Task Definition = Standard, commcost = 1, infcost = 1, retcost = .01

retrieve information that is inconsistent with recent events, such as agreements to use a piece of furniture as part of the task. Close-Consequence has the effect of rehearsing act effects as shown by the effects on A's processing of utterance 3 in figure 8.5.

The difference plot in figure 8.7 shows that Close-Consequence is more detrimental if we take the costs of processing into account. Here communication cost is 1, inference cost is 1, and retrieval cost is .01. The distributions show that the CLOSE-CONSEQUENCE strategy is worse at AWM of 1 ... 5 ($KS > 0.19$, $p < .05$). In addition, once processing costs are taken into account this strategy is no longer beneficial at high AWM; the cost of extra communication outweighs the benefits of avoiding invalid steps.[8]

8.4.2.2 Close-Consequence Detrimental if Communication is Expensive

Figure 8.8 shows that when communication cost is high the extra communication required by the Close-Consequence strategy is detrimental ($KS > .19$ for all AWM, $p < .05$). This is similar to

[8]Only AWM of 16 has a significantly different distribution in the positive direction.

the results we achieved for the Explicit-Acceptance strategy. A strategy can be beneficial due to rehearsal effects only as long as the cost of communication is not too high.

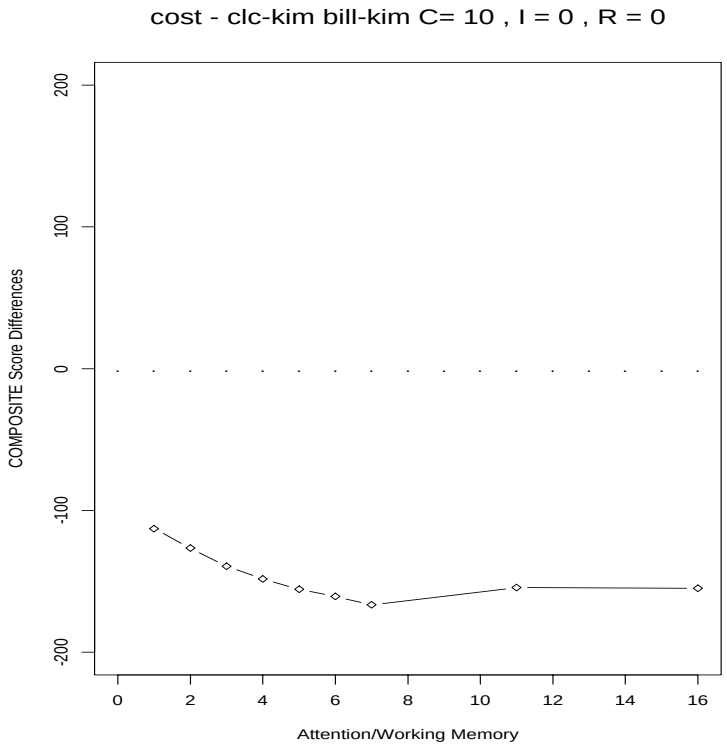


Figure 8.8: Close-Consequence is Detrimental if Communication is Expensive. Strategy 1 is the combination of an All-Implicit agent with a Close-Consequence agent and Strategy 2 is two All-Implicit agents, Task Definition = Standard, commcost = 10, infcost = 0, retcost = 0

8.4.2.3 Close-Consequence Beneficial for Zero-Invalid Task

If the task is the fault-intolerant Zero-Invalids task, then the Close-Consequence strategy is BENEFICIAL at AWM settings of 3 and above. Figure 8.9 demonstrates that strategies which include Consequence IRUs can increase the robustness of the planning process by decreasing the frequency with which agents make mistakes (KS for AWM of 3 and above $> .19$, $p < .05$).

However, the benefits of a strategy must be defined relative to a particular communicative situation. Figure 8.10 shows that the Close-Consequence strategy is detrimental when the task requires agents to achieve matching beliefs on the WARRANTS for their intentions ($KS(7,8) = 0.3$, $p < .01$). This is because IRUs displace other facts from AWM. In this case agents forget the scores of furniture pieces under consideration, which are the warrants for their intentions. Thus here, as elsewhere, we see that IRUs can be detrimental by making agents forget critical information.

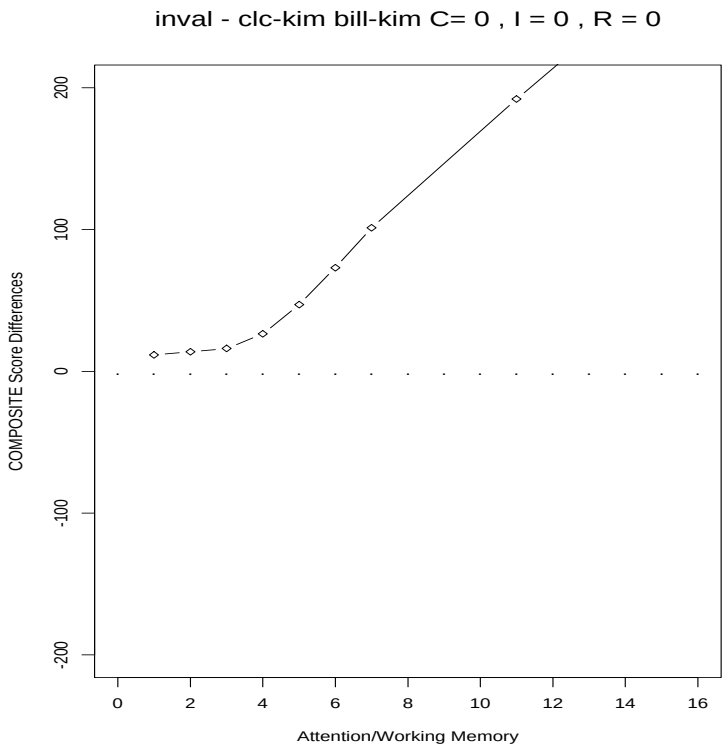


Figure 8.9: Close Consequence is beneficial for Zero-Invalids Task. Strategy 1 is the combination of an All-Implicit agent with a Close-Consequence agent and Strategy 2 is two All-Implicit agents, Task Definition = Zero-Invalid, commcost = 0, infcost = 0, retcost = 0

In sum, this section has shown that even when agents are logically omniscient, the Close-Consequence strategy can be beneficial by reducing the frequency of errors. In task situations that are fault intolerant, Close-Consequence is very beneficial. However if communication is expensive, Close-Consequence is detrimental. Section 8.4.3 presents the results of experiments testing the Close-Consequence strategy with inference-limited agents.

8.4.3 Close-Consequence for Inference-Limited Agents

We saw in section 8.4.2 that a discourse strategy of making inferences explicit is beneficial even for logically omniscient agents when the task is the fault intolerant Zero-Invalid task. We would expect that if agents don't make all the inferences that follow from their beliefs, that this strategy would be even more beneficial. This section validates that prediction.

Two variations of inference limited agents were tested: (1) NO-INFERENCE: severely inference limited agents who don't make any inferences that follow from their beliefs; and (2) HALF-INFERENCE:

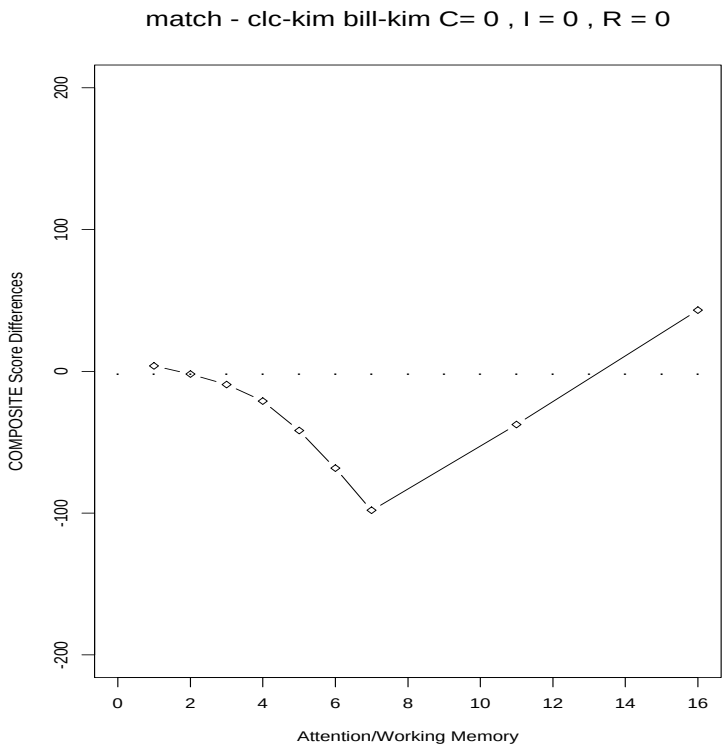


Figure 8.10: Close-Consequence is detrimental for Zero-Nonmatching-Beliefs Task. Strategy 1 is the combination of an All-Implicit agent with a Close-Consequence agent and Strategy 2 is two All-Implicit agents, Task Definition = Zero-Nonmatching-Beliefs, commcost = 0, infcost = 0, retcost = 0

half the time these agents don't make act effect inferences. These inference limitations were discussed more fully in section 5.9.2 and the performance of inference limited agents using the ALL-IMPLICIT strategy was shown in section 5.8 in figures 5.10 and 5.11. The experiments reported in this section tested whether the CLOSE-CONSEQUENCE strategy is beneficial for inference-limited agents.

8.4.3.1 Close-Consequence beneficial for No-Inference agents

One experiment tested NO-INFERENCE agents who used the CLOSE-CONSEQUENCE strategy as compared with NO-INFERENCE agents who didn't. Strategy 1 in figure 8.11 is a dialogue between two NO-INFERENCE agents employing the CLOSE-CONSEQUENCE strategy. Strategy 2 is two NO-INFERENCE agents using the All-Implicit strategy.[9] As we might have expected, figure 8.11 shows that the strategy of making inferences explicit is beneficial in cases where agents don't make these inferences (KS > 0.5 for all AWM, $p < .01$).

[9]The performance of these agents is shown alone in figure 5.10.

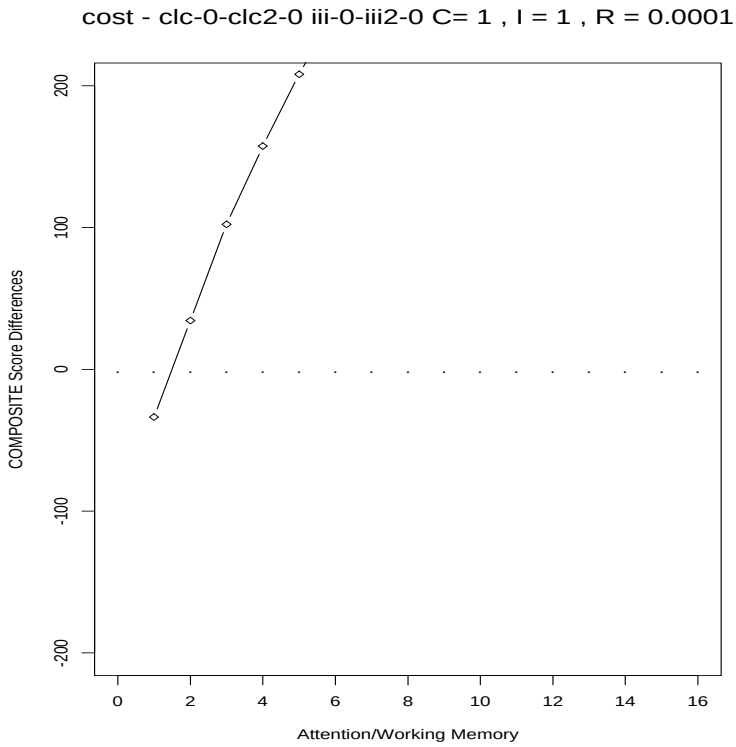


Figure 8.11: Close-Consequence is beneficial for No-Inference agents. Strategy 1 is the two NO-INFERENCE Close-Consequence agents and Strategy 2 is two NO-INFERENCE All-Implicit agents, Task Definition = Standard, commcost = 1, infcost = 1, retcost = .0001

8.4.3.2 Close-Consequence helps No-Inference agents perform like Logically Omniscient Agents

Another experiment compared two NO-INFERENCE agents using the Close-Consequence strategy with logically omniscient agents using the All-Implicit strategy. Figure 8.12 plots the score differences. The figure shows that when both NO-INFERENCE agents employ the Close-Consequence strategy, they actually perform as well as logically omniscient agents at low AWM and better at high AWM KS (7,11) $>$ .19, p $<$ .05).

This demonstrates that situations exist in which NO-INFERENCE agents might be preferred over logically omniscient agents, e.g. those where a large number of inferences may be drawn from a fact when only one is intended or relevant. In this case, the Close-Consequence strategy in combination with NO-INFERENCE agents may be vastly more efficient.

8.4.3.3 Close-Consequence not beneficial for Half-Inference agents

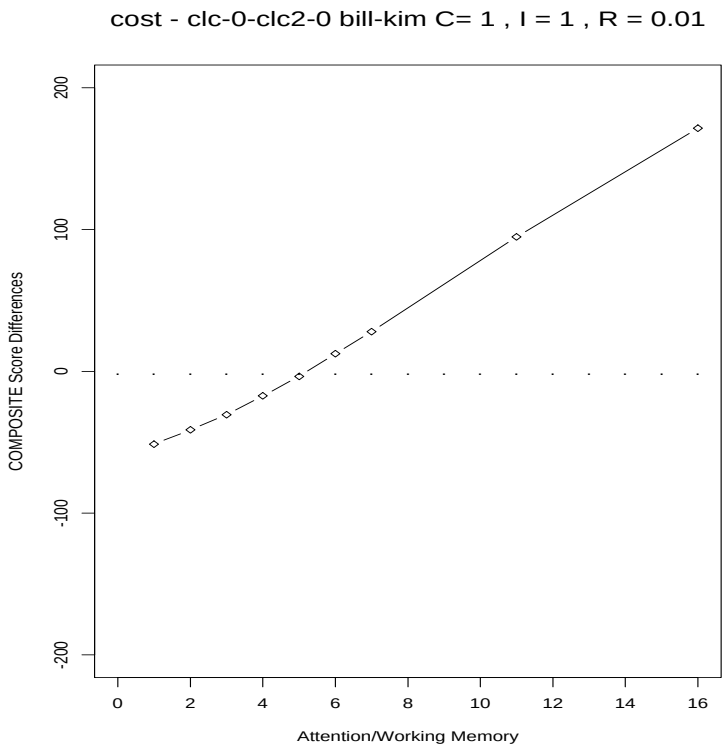


Figure 8.12: Close-Consequence can make NO Inference agents perform as well as ALL inference agents. At AWM < 7, there are NO significant differences when Strategy 1 is two NO Inference Close-Consequence Agents and Strategy 2 is two ALL-Inference All-Implicit agents: Task Definition = Standard, commcost = 1, infcost = 1, retcost = .01

Figure 8.13 is a difference plot between dialogues between two HALF-INFERENCE All-Implicit agents and dialogues between one HALF-INFERENCE Close-Consequence agent and one HALF-INFERENCE All-Implicit agent. In this case, only one agent used the Close-Consequence strategy, because the results shown in figure 8.11 indicate that if both agents use the Close-Consequence strategy, then even NO-INFERENCE agents perform like logically omniscient ones.

The figure shows that if both agents are HALF-INFERENCE agents, then the Close-Consequence strategy has no effect (KS < .19 for all AWM, NS). Since there is no significant difference between the performance of HALF-INFERENCE agents and ALL-INFERENCE agents, there may be no effect because performance may be at near ceiling in these situations.

8.4.4 Increasing Inferential Complexity: Consequence IRUs

In order to investigate the role of Inference-Explicit IRUs in more complex inferential situations, the simulation was designed so that the task can be varied to increase inferential complexity. This is done by defining a version of the task in which agents must also try to achieve matched-pair

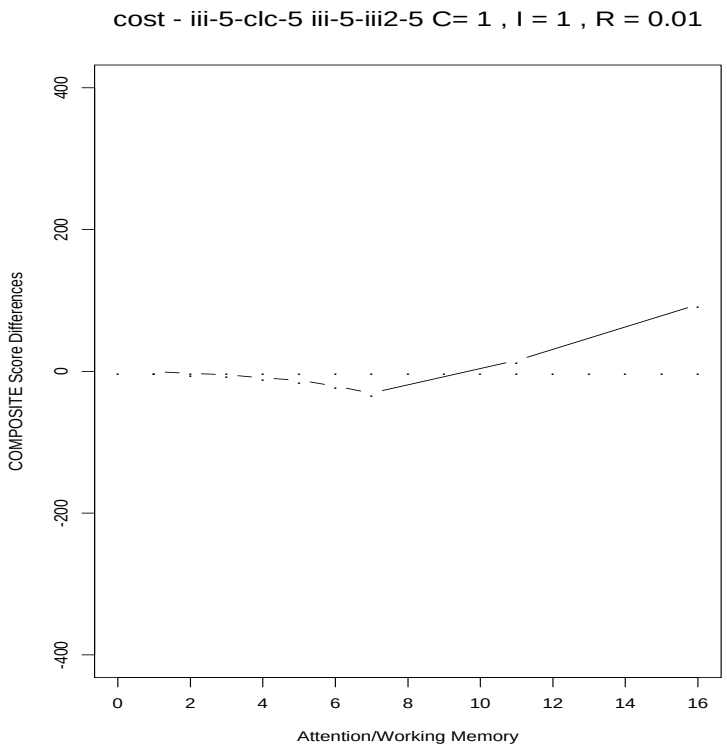


Figure 8.13: No Benefit for Close-Consequence. Strategy 1 is the combination of a HALF-INFERENCE All-Implicit agent with a Close-Consequence agent and Strategy 2 is two HALF-INFERENCE All-Implicit agents, Task Definition = Standard, commcost = 1, infcost = 1, retcost = .01

goals as discussed in section 5.8. Achieving these goals requires making inferences about when two actions will achieve a matched pair.

In order to explore the effects of inference-explicit IRUs that have salient antecedents as opposed to those with remote antecedents, Design-World supports two versions of the matched pair task. In this task, in addition to performing the Standard task, the utility of an action can be increased if it contributes to an optional goals of a matched pair. A matched pair is two furniture items of the same color.

In one task, Matched-Pair, the pieces of furniture that the matched pair consists of must be in the same room. In the second version of the task, Matched-Pair-Two-Room, the pieces of furniture that the matched pair consists of must be in different rooms. The premises for inferences about matched pair goals are what actions the agents have agreed to do, with the resulting inference based on inferred acceptance that an act is INTENDED. In the Matched-Pair task, the premises for the matched pair inferences are either currently salient or inferrable from currently salient beliefs. In the Matched-Pair-Two-Room task, the premises for the matched pair inferences are not likely to

be salient because they are based on agreements that were made while the agents were discussing the other room.

8.4.4.1 The Matched-Pair-Inference-Explicit strategy

The Matched-Pair-Inference-Explicit strategy makes an inference explicit that is used as a premise in a matched pair inference. Figure 8.14 shows the discourse action protocol for a Matched-Pair-Inference-Explicit agent. Openings, Acceptances and Closings are left implicit, but each proposal includes an extra utterance which makes an inference explicit.

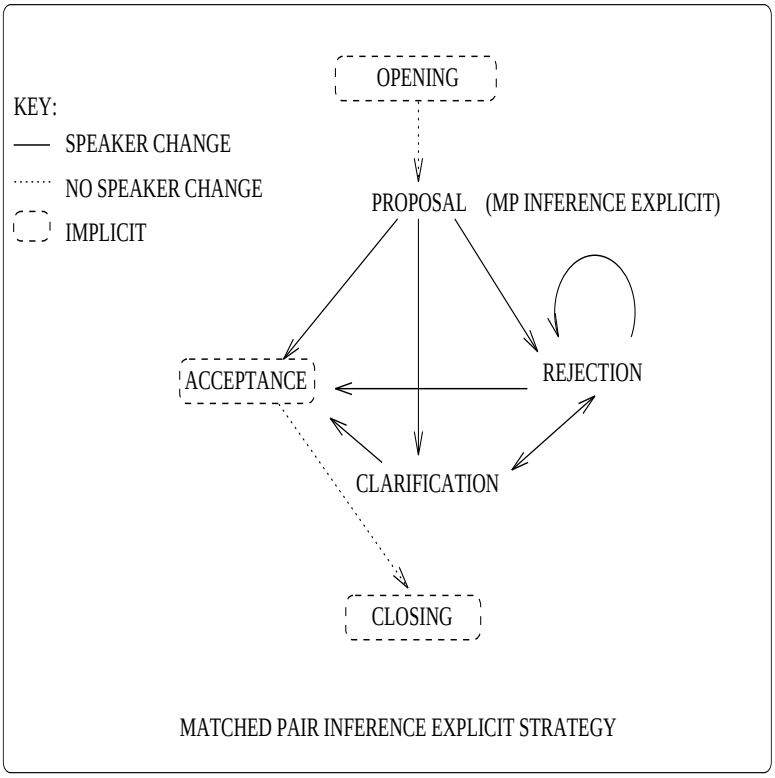


Figure 8.14: Discourse Action protocol for agents using the Matched-Pair-Inference-Explicit strategy

Figure 8.15 shows a dialogue sequence composed of the interaction of two Matched-Pair-Inference-Explicit agents.

The Matched-Pair-Inference-Explicit strategy is exemplified by both agents IMI and IMI2 in 117, where Consequence IRUs are shown in CAPS:

(117) IMI: *Then, let's put the green rug in the study.*
1:(propose agent-imi agent-imi2 option-10: put-act (agent-bill green rug room-1))
IMI2: WE AGREED TO PUT THE GREEN RUG IN THE STUDY.

Agent A	Agent B
PROPOSAL 1	
	PROPOSAL 2
PROPOSAL 3	
	REJECTION 3,4
PROPOSAL 4	
...	...
	PROPOSAL N

Figure 8.15: Sequence of Discourse Acts for two Matched Pair Inference Explicit Agents as in Dialogue 117

2:(say agent-imi2 agent-imi intended-10: put-act (agent-bill green rug room-1))
IMI2: *Then, let's put the green lamp in the study.*
3:(propose agent-imi2 agent-imi option-35: put-act (agent-kim green lamp room-1))
IMI: *Then, let's put the green couch in the study.*
4:(propose agent-imi agent-imi2 option-47: put-act (agent-bill green couch room-1))
IMI2: *No, instead let's put the purple couch in the study.*
5:(reject agent-imi2 agent-imi option-58: put-act (agent-kim purple couch room-1))
IMI: *Then, let's put the green couch in the study.*
6:(propose agent-imi agent-imi2 option-64: put-act (agent-bill green couch room-1))
IMI2: WE AGREED TO PUT THE PURPLE COUCH IN THE STUDY.
7:(say agent-imi2 agent-imi intended-48: put-act (agent-kim purple couch room-1))
IMI2: *No, instead let's put the purple chair in the study.*
8:(reject agent-imi2 agent-imi option-75: put-act (agent-kim purple chair room-1))
IMI: *Then, let's put the green couch in the living room.*
9:(propose agent-imi agent-imi2 option-82: put-act (agent-bill green couch room-2))
IMI2: WE AGREED TO PUT THE GREEN COUCH IN THE LIVING ROOM.
10:(say agent-imi2 agent-imi intended-82: put-act (agent-bill green couch room-2))
IMI2: *Then, let's put the green table in the living room.*
11:(propose agent-imi2 agent-imi option-98: put-act (agent-kim green table room-2))
IMI: WE AGREED TO PUT THE GREEN COUCH IN THE LIVING ROOM.
12:(say agent-imi agent-imi2 intended-79: put-act (agent-bill green couch room-2))
IMI: *No, instead let's put the green chair in the living room.*
13:(reject agent-imi agent-imi2 option-109: put-act (agent-bill green chair room-2))
IMI2: *Then, let's put the purple rug in the living room.*
14:(propose agent-imi2 agent-imi option-118: put-act (agent-kim purple rug room-2))
IMI: *Then, let's put the green table in the living room.*
15:(propose agent-imi agent-imi2 option-123: put-act (agent-kim green table room-2))

In 117, 2, 7, 10 and 12 all make an inference explicit related to what actions are intended. These inferences follow from the implicit acceptance of the proposals given in 1, 5, 9 and 9 respectively. In

this situation, agents always make these inferences. IMI chooses to make these inferences explicit when IMI has inferred that a matched pair is possible.

Figure 8.16 shows the relationship between the underlying cognitive processes and the discourse and utterance acts for an excerpt of a dialogue between two Matched Pair inference Explicit agents such as that in 117. Because this is the MATCHED PAIR version of the task, agents check for matches to calculate potential score increases from achieving matched pairs before deliberation. In order to do this Check Matches procedure they must retrieve from memory what actions they have already agreed upon. The already agreed intentions are the premises for making a matched pair inference.

The task structure for each room represents a discourse segment, if we use task structure to approximate discourse structure (Grosz, 1977; Grosz and Sidner, 1986; Sibun, 1991). In the dialogue in 117, the task is Matched-Pair and IMI makes inferences explicit whose premises are in the current discourse segment. In the dialogue shown in 118, the task is Matched-Pair-Two-Room and IMI makes inferences explicit whose premises were in the previous discourse segment. At most AWM settings, the inferences made while reasoning about the task for the previous segment would no longer be salient.

(118) IMI: *Then, let's put the green rug in the study.*
1:(propose agent-imi agent-imi2 option-10: put-act (agent-bill green rug room-1))
IMI2: *Then, let's put the green lamp in the study.*
2:(propose agent-imi2 agent-imi option-33: put-act (agent-kim green lamp room-1))
IMI: *Then, let's put the green couch in the study.*
3:(propose agent-imi agent-imi2 option-50: put-act (agent-bill green couch room-1))
IMI2: *No, instead let's put the purple couch in the study.*
4:(reject agent-imi2 agent-imi option-61: put-act (agent-kim purple couch room-1))
IMI: *Then, let's put the green couch in the study.*
5:(propose agent-imi agent-imi2 option-69: put-act (agent-bill green couch room-1))
IMI2: WE AGREED TO PUT THE PURPLE COUCH IN THE STUDY.
6:(say agent-imi2 agent-imi intended-51: put-act (agent-kim purple couch room-1))
IMI2: *Then, let's put the purple rug in the living room.*
7:(propose agent-imi2 agent-imi option-80: put-act (agent-kim purple rug room-2))
IMI: WE AGREED TO PUT THE GREEN LAMP IN THE STUDY.
8:(say agent-imi agent-imi2 intended-33: put-act (agent-kim green lamp room-1))
IMI: *No, instead let's put the green chair in the living room.*
9:(reject agent-imi agent-imi2 option-91: put-act (agent-bill green chair room-2))
IMI2: WE AGREED TO PUT THE PURPLE COUCH IN THE STUDY.
10:(say agent-imi2 agent-imi intended-51: put-act (agent-kim purple couch room-1))
IMI2: *Then, let's put the purple chair in the living room.*
11:(propose agent-imi2 agent-imi option-107: put-act (agent-kim purple chair room-2))

Discourse Act	Utterance Act	A PROCESS	B PROCESS
		ME-REASON Put-1 Deliberate	
PROPOSAL	1:(propose A B option-10)		
			ME-REASON Put-1 Check Preconds 10 Deliberate Decide to Accept Store (MS intend 10) Store (Act-Effects 10) ME-REASON Put-2 Check Matches Deliberate Propose, MP IE
PROPOSAL	2:(say B A intended-10)		Store intended-10
		Infer Acceptance 10 Store (MS intend 10) Store (Act-Effects 10) Infer Close Put-1 Infer Open Put-2 Store intended-10	
	3:(propose B A option-35)		
		ME-REASON Put-2 Check Preconds 35 Check Matches Deliberate Decide to Accept Store (MS intend 35) Store (Act-Effects 35) ME-REASON Put-3 Deliberate Check Matches Propose, MP IE	
	... .		

Figure 8.16: Sequence of Discourse Acts for Dialogue in 117 between Two Matched-Pair-Inference-Explicit agents

In 118, 6, 8 and 10 make inferences explicit of what actions are already intended. These inferences followed from the implicit acceptance of the proposals given in 4, 2 and 4 respectively.

Section 5.8 showed a sample dialogue for two All-Implicit agents in the MATCHED PAIR version of the task. The performance of two agents using the MATCHED-PAIR-INFERENCE-EXPLICIT strategy will be compared with the performance of two All-Implicit agents for both the Matched-Pair task and the Matched-Pair-Two-Room task in the following sections. All-Implicit agents can generally

do quite well at achieving matched pairs, thus the strategies to be discussed below must consistently achieve more matched pairs than the All-Implicit agents to perform better on the task.

8.4.4.2 Matched Pairs with Salient Antecedents

This section discusses the performance of agents using the Matched-Pair-Inference-Explicit strategy in the Matched-Pair version of the task, as illustrated in the dialogue in 117. In this situation, inferences are made explicit while they are currently salient because they are meant to be used as premises for the inference of a matched-pair.

As shown in figure 8.17 the strategy is **not** beneficial. The only AWM setting where the distributions are significantly different is at AWM of 11 (KS = .2, p < .05).

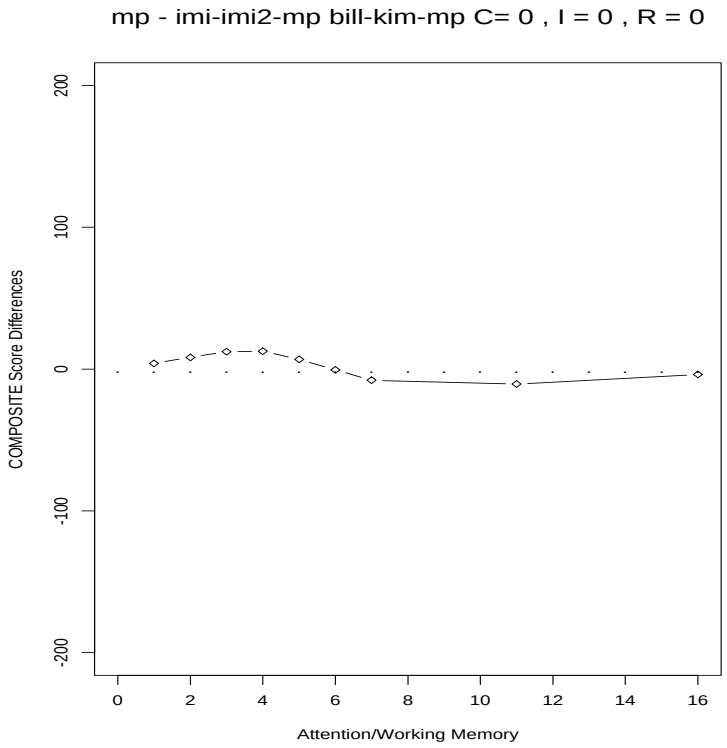


Figure 8.17: Matched-Pair-Inference-Explicit not beneficial for Matched-Pair,Strategy 1 is two Matched-Pair-Inference-Explicit agents and Strategy 2 is two All-Implicit agents, Task Definition = MP, commcost = 0, infcost = 0, retcost = 0

These results show that making an inference explicit need not be beneficial, even in the context where it can be used as a premise for another inference. In contrast, the next section will show that the Matched-Pair-Inference-Explicit strategy is beneficial for the Matched-Pair-Two-Room version of the task.

8.4.4.3 Matched Pairs with Non-Salient Antecedents

This section presents results on evaluating the MATCHED-PAIR-INFERENCE-EXPLICIT strategy in the MATCHED-PAIR-TWO-ROOM version of the matched pair task. In this situation, the matched pair inference premise is made explicit in a situation in which it is not currently salient. This situation/strategy combination parallels the subset of inference explicit IRUs in the corpus where the antecedents are not currently salient. These were discussed in sections 8.3.1 and 8.3.2.

The difference plot in figure 8.18 for matched pair points shows that this strategy is effective in increasing agents' ability to make matched pair inferences. As the figure shows, agents using this strategy get increasingly more points from matched pairs as AWM increases. The differences in the distributions are significant at AWM of 3,4,5,6,11, and 16 (KS(3,4,5,6,11,16) > 0.21, p < .05).

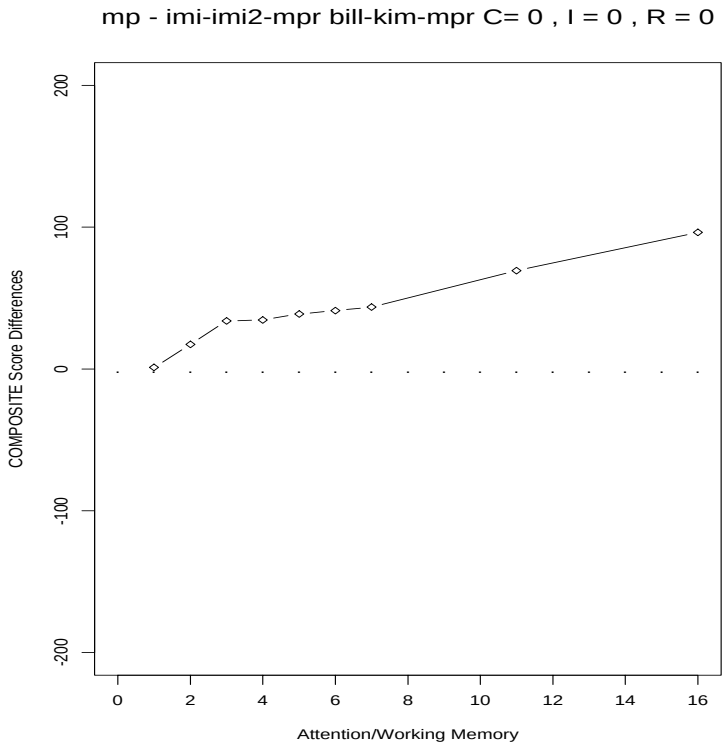


Figure 8.18: Matched-Pair-Inference-Explicit beneficial for Matched-Pair-Two-Room. Strategy 1 is two Matched-Pair-Inference-Explicit agents and Strategy 2 is two All-Implicit agents, Task = Matched-Pair-Two-Room, commcost = 0, infcost = 0, retcost = 0

Thus the Matched-Pair-Inference-Explicit strategy is beneficial when inferential complexity is increased, and when agents are data-limited on the premises that they can use for inferences. Making an inference explicit is most beneficial in this situation when the inference, which was made in an earlier discourse segment, is made salient in the current segment to support other inferences. If

the inference is currently salient, the strategy has no beneficial effect in terms of helping agents to make inferences.

8.4.5 Summary: Consequence in Design-World

The Design-World experiments discussed here tested two processing aspects of Consequence IRUs. The Close-Consequence strategy experiments discussed in section 8.4.1 showed that even when agents are logically omniscient that the Close-Consequence strategy can be beneficial in helping agents avoid mistakes. This is basically a result that it is helpful to rehearse facts that keep one from making mistakes. When agents are inference-limited, making inferences explicit are especially beneficial, as we would expect. Surprising however, section 8.4.3 showed that inference-limited agents can perform as well as logically omniscient agents if they have the right conversational partner and hence shows the power of strategies that include IRUs.

The Matched-Pair-Inference-Explicit strategy makes inferences explicit that are then used as the premises for matched pair inferences. This strategy was tested in two situations, one in which the inference has been made and is currently salient, and one in which the inference has been made but is not currently salient. The strategy was beneficial only when the inference was not currently salient, thus we expect that in these cases making the inference explicit functions much like other cases of making premises salient at the point that they will be used by the other agent in reasoning, i.e. in a similar way to an Attention strategy. This strategy can be compared with the Explicit-Warrant strategy discussed in chapter 7 which also makes premises explicit when they are to be used. In the Explicit-Warrant strategy however the premises are used in deliberation.

The results are interesting with respect to the distinction between data-limited and inference-limited processing discussed earlier (Norman and Bobrow, 1975). While various cognitive processes may be data-limited (attention) or inference-limited (means-end reasoning and reasoning about the effects of actions), the data here demonstrate that no matter how much attentional capacity an agent has, if s/he doesn't make the correct inferences, then s/he cannot achieve a high level of performance, at least not without the right conversational partner. However as we have also seen, even agents that are logically omniscient will have less than perfect performance when they are attention limited. IRUs help agents improve their performance in both cases.

8.5 Consequence:Summary

This chapter argued that inference in discourse is constrained by the DISCOURSE INFERENCE CONSTRAINT, i.e. propositions must be currently salient to be premises for inferences. This is supported

by the distributional analysis showing that Inference-Explicit IRUs, unlike Paraphrases, typically have salient antecedents.

The Design-World simulations in section 8.4.2 showed that even when agents are logically omniscient, making inferences explicit can help when tasks are fault-intolerant or have higher inferential complexity. The simulations also supported the discourse inference constraint by showing that IRUs can be beneficial in these situations at lower AWM settings when agents are attention limited.

Furthermore, Design-World experiments were conducted to test the differences between Inference-Explicit IRUs with salient or remote antecedents. The Matched-Pair-Inference-Explicit strategy makes inferences explicit either when their antecedents are salient or when their antecedents are remote, depending on the version of the matched pair task. The strategy was shown to be beneficial when the antecedents for the inference are remote. In this situation the IRU can help agents make other inferences that are dependent on that inference as a premise. The Explicit-Warrant strategy discussed in chapter 7 has a similar beneficial effect when the task requires beliefs about warrants to be shared.

The distributional analysis discussed the fact that in many cases where inferences are made explicit it would appear that the inference was already made. In these cases, Inference-Explicit IRUs may also function similarly to Attitude or Attention IRUs.

Chapter 9

Conclusions and Future Work

This thesis investigates the interaction of resource bounds with speakers' language behavior through an examination of INFORMATIONALLY REDUNDANT UTTERANCES, IRUs. IRUs are apparently inefficient since their information content is already believed by all conversants in a dialogue. I argue however that their function can be explained by a model of dialogue in which conversants are autonomous and resource-limited. The resource limits that I've investigated are limited attentional capacity and limited reasoning capacity.

In chapter 1, I proposed a theory of the function of IRUs based on categorizing IRUs by three communicative functions: Attitude, Consequence and Attention. Then I claimed that Consequence and Attention IRUs are related to limited inferential and limited attentional capacity and that Attitude IRUs are also partially attributable to these limits. Attitude IRUs are the externalization of attitudes toward propositions conveyed, e.g. ACCEPTANCE or REJECTION of an assertion. Conversants externalize their attitudes for at least two reasons: (1) their autonomy means that acceptance cannot be automatically assumed and (2) their resource bounds introduce uncertainty as to what they are currently attending to and what they have inferred from what has been said. The remainder of the thesis supports and elaborates the theory and provides empirical support for the claims.

This chapter summarizes the methodological and theoretical contributions and then proposes future work. Section 9.1 summarizes the methodological contributions and section 9.2 summarizes the theoretical contributions. Section 9.3 discuss specific results from the distributional analysis and section 9.4 reviews the results from the Design World simulations. Section 9.5 will suggest potential computational applications of the results presented here. Finally, section 9.6 will discuss the limitations of the results, possible extensions to the results and other future work.

9.1 Methodological Contributions

Theoretical research on discourse models is plagued by the gap between surface form and discourse function. The problem is that the function of an utterance is determined by many factors such as the context of the utterance, its syntactic and prosodic realization, cultural conventions and the shared beliefs of the conversants. The discourse analyst typically has only partial access to these factors and is faced with the difficulty of determining their interaction.

If the discourse analyst attempts to ascribe intentions to the speaker, often many ascriptions are possible. These ascriptions can at best be well founded hypotheses about the speaker's mental state. Furthermore, utterances can realize multiple intentions and it is usually not clear which if any of the possible hypothesized intentions were the primary motivation of the speaker. The analyst who relies on intuitions about intention may be surprised to discover that the intuition changes on another occasion of analysis.

In order to avoid some of the most common methodological problems, this thesis uses two empirical methods, discussed in detail in chapters 4 and 5.

The first empirical method is distributional analysis, a quantitative method based on coding utterances according to a range of distributional factors and then using these factors in multi-variate analysis. Factors are chosen which are (1) objectively defined, (2) posited to be a surface correlate of or a constraint on the hypothesized discourse functions. These factors measure characteristics of both utterances and contexts.

Here 210 IRUs were selected from a naturally occurring corpus of problem solving dialogues. The distributional analysis suggests hypotheses about function and provides empirical support for a model for the function of IRUs. Each IRU was coded according to 13 coding factors. These coding factors were used both to determine classes of IRUs and subclasses within these (taxonomy). They were also used to argue that these classes had particular discourse functions. Types of IRUs defined by one type of coding factor such as sequential location are examined in many different contexts before positing a discourse function. In many cases statistically significant correlations were found between forms and contexts.

In this case the distributional analysis was done completely by hand. In the future, with the advent of new online corpora, this type of analysis should be much easier to perform.

The theory of IRU function suggested by the distributional analysis included processing claims which are difficult to support by a distributional analysis alone. Thus the second method involved

testing the parameters of the model via a computational simulation environment called Design-World.

Design-World is first of all an implementation of the key aspects of the theory. The following section discusses the theoretical model that Design-World was based on. Implementations usually demonstrate the plausibility of a theory, but Design-World is parametrizable so that complex interactions between different parameters of the model can be tested. The parameters included the task definition, agents' cognitive limitations and discourse strategies. The use of simulations allowed me to support the major claim of the thesis about the relationship of resource limits to agents' language behavior.

9.2 Theoretical Contributions

The primary contribution of the thesis is a theory of the function of IRUs in dialogue. IRUs have not been systematically investigated in previous work and many accounts assume that they are infelicitous and thus require a special mode of analysis to interpret. The major claim of the thesis is that IRUs make sense if we adopt a model of discourse **processing** that reflects conversants' resource bounds.

Research in the Gricean tradition would lead one to expect that IRUs would be infelicitous or would be interpreted on the basis of recognizing their infelicity as discussed in section 2.2. I have shown that it is possible to interpret IRUs with standard interpretation mechanisms as long as we acknowledge agents' resource bounds. The claim that IRUs support the DISCOURSE INFERENCE CONSTRAINT is not identical to Gazdar's claim that tautologies are felicitous because they add inferences to the discourse model. All utterances add inferences to the discourse model, so IRUs are not unique on that account, and testing whether inferences are added does not determine when IRUs will be felicitous as opposed to infelicitous. The only type of IRU that does **not** occur in the corpus are cases of a speaker repeating his or her own proposition, in the same form, with the same prosody, adjacent to the original assertion. Any variation is acceptable. Even cases of such repetition, which probably occur in other contexts, are interpretable without resorting to Gricean reasoning by assuming that the speaker 'really means' a proposition that is repeated while still salient.

The claimed relationship between IRUs and resource bounds required positing and supporting an account of the underlying resource-limited cognitive processes and representations used in discourse processing. While much current work focuses on the relationship between resource-bounds and reasoning about action (Dean and Boddy, 1988; Bratman, Israel, and Pollack, 1988; Pollack and

Ringuette, 1990), to my knowledge there has been no other work on the relationship of resource bounds to language behavior. The effects of utterances on another agent's mental state are context dependent, uncertain, unobservable and highly dependent on the other agent's resource limitations.

The theoretical framework for discourse processing is developed in chapter 3, and extended and supported throughout the rest of the thesis. Chapter 3 (1) specifies a psychologically plausible model of attention/working memory; (2) specifies a model of belief deliberation; (3) integrates the two models by proposing the DISCOURSE INFERENCE CONSTRAINT, i.e. that inference and deliberation only operate on salient beliefs; (4) extends the model of belief deliberation to a model of deliberating about mutual beliefs (mutual suppositions); and (5) uses this as the theoretical framework for the underlying cognitive processes that motivates the proposed functions of IRUs.

The model of Attention/Working Memory (AWM) produces both recency and frequency effects for propositions asserted in discourse with very simple mechanisms (section 3.2). AWM is based on a proposal by Landauer (1975), but Landauer's model was extended and specified to support complex behavior in discourse. Specific extensions are:

1. The cognitive objects stored in AWM are propositions;
2. No explicit links are maintained between propositions;
3. Belief change is a result of storing new beliefs rather than deleting or negating old beliefs;
4. Both internal cognitive processes and external processes store propositions in AWM with the effect that both retrieval and reasoning duplicate items and store them in a new configuration in memory.

Assumption (2) means that associations between beliefs are only a result of the proximity resulting from the fact that two beliefs were processed near one another in time. Other associations between beliefs are the result of specific strategic retrieval processes. The effect is that AWM models the PRINCIPLE OF POSITIVE UNDERMINING so that beliefs persist even when their supports are degraded (Harman, 1986; Tversky and Kahneman, 1982; Ross and Anderson, 1982; Galliers, 1990). Assumption (4) produces associative effects via a simple model which does not require positing different types of memory for different types of beliefs, i.e. semantic vs. episodic memory.

In the Design-World simulations, AWM produces the desired effects of limited attentional capacity. It is used there to show that attention-limited agents behave differently and benefit from different communication strategies than agents who are not resource limited.

In chapter 3, I propose a formulation of Grosz and Sidner's stack model of Attentional State in discourse in terms of AWM. The motivation is to show that a simple reformulation of the stack

model makes it much more cognitively plausible. The reformulation of PUSH and POP as the effect of storage and retrieval operations in AWM makes predictions about what kinds of discourse behavior increase or decrease processing load and makes sense of cases of Attention IRUs where items on the top of the stack as the result of a POP are not treated as salient by the conversants.

Section 3.3 presents a model of belief and intention deliberation. Deliberation provides a basis for the ACCEPTANCE and REJECTION of assertions and proposals and different types of language behavior affect the degree to which a belief is 'epistemically entrenched' (Gardenfors, 1990).

Deliberation and inference processes are posited to depend on what is currently salient. This assumption links the AWM model with deliberation, and is encapsulated in the DISCOURSE INFERENCE CONSTRAINT. Communication is a strategic process targeted at deliberation (Galliers, 1990) and this strategic process must take agents' resource limits into account.

In section 3.4, the model of belief deliberation is extended to model deliberation about what is mutually believed. Since the effects of utterances are not directly observable, mutual beliefs are recast as MUTUAL SUPPOSITIONS. Endorsement on supporting beliefs reflect the fact that inferences about what other agents accept are defeasible, and are more or less supported by evidence made salient in the discourse at the time of the assertion or proposal.

In chapter 6, the distributional analysis of Attitude IRUs is used to investigate how ACCEPTANCE and REJECTION are communicated. The analysis of Attitude IRUs provides support for and develops the account of MUTUAL SUPPOSITION based on an analysis of the variation in the surface form of Attitude IRUs. The analysis of Attitude IRUs that convey acceptance shows that the difference in form affects the defeasibility of the mutual supposition. The fact that one type of Rejection IRUs reject by an implicature shows that assertions are only added to the context as HYPOTHESES until after the hearer has an opportunity to reject.

Then section 6.5 shows how the theory of Attitude IRUs can be used directly to model the process by which agents construct COLLABORATIVE PLANS. Conditions on when an agent should convey ACCEPTANCE or REJECTION is presented in the COLLABORATIVE PLANNING PRINCIPLES, which proscribe what is collaborative communicative behavior.

The resulting account is similar to that of Power (1974), Grosz and Sidner (1990) and Cohen and Levesque (1991). The difference is that I focus on the dialogue process by which agents communicate that they ACCEPT and REJECT other agent's proposals or assertions. I assume that agents retain their autonomy throughout a dialogue and that each step of a COLLABORATIVE PLAN is negotiated. Agents' autonomy is tied to the theory of belief and intention deliberation discussed in chapter 3. The role of intention deliberation is included in the definition of a COLLABORATIVE PLAN by specifying that

both agents must assure themselves of the utility of the resulting plan. In Grosz and Sidner (1990) and Cohen and Levesque (1991), once an agent agrees to a SharedPlan or a Joint Intention, it has agreed to accept the initiating agent's proposals for each step of the plan. Thus there is no deliberation of each proposal.

In chapters 7 and 8, attention-limited belief deliberation is proposed as the mechanism underlying Deliberation IRUs and Affirmation IRUs. These IRUs make premises salient that are the basis of an inference or that are used to decide whether to accept or reject proposals and assertions.

Chapter 7 also discusses Open Segment and Close Segment IRUs. There I investigate the evidence that these are used as conventional signals about discourse structure. I show that another hypothesis is also plausible, namely that they are motivated by the same mechanisms as Deliberation IRUs. Chapter 8 discusses the class of IRUs that make inferences explicit. I show that these can be motivated by an account of limited inference, where the IRU ensures that the inference is made. However some cases of Inference-Explicit IRUs also seem to be motivated by the need to make an inference salient that is relevant to the current purpose, despite the fact that the inference has already been made.

The support for the theory from the distributional analysis will be discussed in section 9.3. Specific results from the Design-World simulations will be discussed in section 9.4. Then section 9.6 will discuss future work.

9.3 Distributional Analysis Summary

The functional analysis of IRUs was based on the distributional analysis. A summary of the specific results for each class of IRUs is discussed in the following sections.

9.3.1 Attitude Distributional Analysis

Figure 9.1 summarizes the distributional correlates of Attitude IRUs. On the taxonomic level, the distributional analysis is used to argue that Attitude IRUs are characterized by the two INFORMATION STATUS parameters of Adjacent and Other: these define the ATTITUDE LOCUS.

The distributional analysis shows that both ACCEPTANCE and REJECTION are typically displayed in the ATTITUDE LOCUS, and this provides support for the COLLABORATIVE PRINCIPLE. the COLLABORATIVE PRINCIPLE can then be used to specify the conditions under which default inferences about understanding and acceptance are licensed.

Acceptance and rejection can both be conveyed by Attitude IRUs, but they differ in terms of

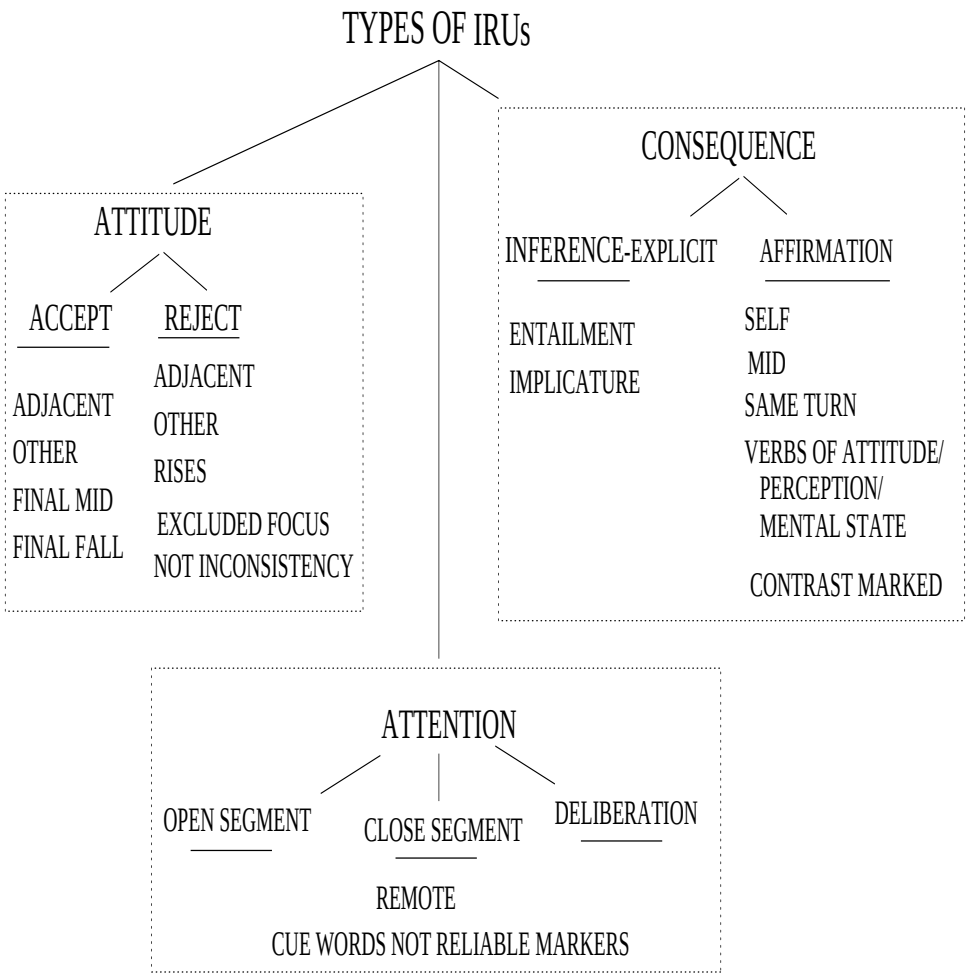


Figure 9.1: Overview of the Taxonomy of IRUs

prosodic features. Attitude Accept IRUs tend to be prosodically marked with a phrase final Mid tone which is a surface cue that the speaker is treating the content of the IRU as HEARER OLD information. This means that the Attitude Accept function is in all likelihood recognizable from surface prosodic cues. The theoretical ramification is that it may not be necessary for the hearer to recognize redundancy by any logical process.

There are two types of Attitude Reject IRUs. One type are realized with phrase final rises and apparently simply query the previous assertion. The second type reject the previous assertion by an implicature, and to my knowledge have never been noted before. The implicature arises when the EXCLUSION OF FOCUS CONDITION is met, i.e. the speaker excludes the information focus of the assertion in the IRU. The fact that rejection can be conveyed by implicature provides support for the theory of mutual supposition proposed in section 3.4 and elaborated in section 6.3, because

an assertion can be defeated by an implicature only if the assertion is added to the context as an HYPOTHESIS until after the opportunity for acceptance or rejection has passed.

9.3.2 Attention Distributional Analysis

Figure 9.1 summarizes the results of the distributional analysis for Attention IRUs. Attention IRUs are characterized by the single distributional information status parameter of Remote. If the antecedent for an IRU is remote, then the proposition realized by the antecedent is unlikely to be salient. Thus the function of the IRU can be to make that proposition salient.

Attention IRUs have few distributional correlates except that they are not marked as HEARER-OLD information by phrase final prosody. However, more specific subclasses of Attention IRUs are defined by other distributional correlates. In section 4.5, I described the criteria for deciding that an Attention IRU is an Open Segment or a Close Segment IRU. To my knowledge, Open Segment IRUs have never been noted before. The observation that they exist and the characterization of their function resulted from the distributional analysis of IRUs.

Close Segment IRUs had been noted before, e.g. the observation that summaries often occur at the end of discourse segments (Whittaker and Stenton, 1988). The observation that IRUs are often used to support Deliberation had not been made, although both work in RST and Cohen's work on argumentation had noted that propositions that the hearer already believed or would find plausible make good justifications (Mann and Thompson, 1987; Cohen, 1981).

9.3.3 Consequence Distributional Analysis

Figure 9.1 summarizes the distributional analysis of Consequence IRUs. There are two taxonomic classes of Consequence IRUs: Inference-Explicit and Affirmation.

Inference-Explicit IRUs are identified by the single distributional factor that the HEARER OLD relation between the IRU and its antecedent(s) is either Implicature or Entailment. The distribution of Inference-Explicit IRUs provides evidence for the hypothesized DISCOURSE INFERENCE CONSTRAINT introduced in chapters 1 and 3. The argument is based on the fact Paraphrases and Inference-Explicit IRUs both realize entailed information, but Inference-Explicit IRUs are much more likely to have salient antedents. Thus whenever we have evidence that inferences are being drawn, the antecedents for those inferences are currently salient.

Inference-Explicit IRUs are not marked as HEARER OLD information by phrase-final prosody. Thus

either the speaker believes they are new information, i.e. the inference has not been made, or that they should be treated as new information.

Affirmation IRUs contrast rhetorically with an adjacent utterance and co-occur with the cue phrase *but*. Contrast is supported because the IRU and an adjacent utterance would lead to conflicting defaults, or because there is a focus/open proposition structure based on a salient set. Affirmation IRUs typically have salient antecedents and are often prosodically marked as hearer old information by a final Mid tone. Their cooccurrence with a verb of perception or mental state such as *see, hear, remember* contributes to a theory of belief and intention deliberation because they indicate typical sources of evidence and counter-evidence for propositions being deliberated.

9.4 Design-World Simulations Summary

The second empirical method was to test hypotheses about the function of IRUs with Design-World simulations.

One general result is that unbounded resources can be a disadvantage. When agents are parametrized to have the capability of searching all of memory as they carry out the task, the increased retrieval and inference costs may far outweigh any potential increase in performance. In addition, in certain situations, accessing all of memory can be detrimental because agents don't distinguish recent information from information that is out of date. There seems to be a cognitive benefit of forgetting facts that are no longer valid that interacts with the general cognitive detriment of forgetting.

A key result is that there is a complex interaction between strategies, task parameters and agent properties. This interaction supports the argument that IRUs are not a component of a generally cooperative strategy: a beneficial strategy must be targeted at the particulars of the situation. For example, Chapter 7 discussed strategies that help agents who have low attentional capacity in section 7.6.3, while section 7.6.2 presented results from a strategy that helps agents with high attentional capacity avoid mistakes. To my knowledge, there has been no previous work which has been able to demonstrate that particular discourse strategies are beneficial in some situations and not in others; this area deserves further exploration.

Another interesting general result across all the strategies is that IRUs can be detrimental because they may displace propositions from memory that are potentially more useful than the proposition that the IRU realizes. Thus despite any potential rehearsal benefits of 'important' or potentially valuable information, strategies in which IRUs are not targetted at the current purpose are not beneficial. Even in cases where IRUs can be beneficial, the benefit must outweigh the potential detriment of displacing facts that could be useful now or at a later point in the task.

The following sections will discuss the simulation results for each of Attitude, Consequence and Attention. In each case a strategy was designed based on the distributional analysis and this strategy was compared with the All-Implicit (Baseline) strategy.

9.4.1 Attitude Simulations Summary

Strategy1	Agent Params	Task Params	Costs comm,inf,ret	Beneficial?
Explicit-Accept	All-Inference	Standard	0,0,0	AWM > 6
Explicit-Accept	All-Inference	Standard	1,1,.0001	AWM > 7
Explicit-Accept	All-Inference	Standard	1,1,.01	NOT SIG
Explicit-Accept	All-Inference	Standard	10,1,0	NOT, AWM < 7
Explicit-Accept	All-Inference	Zero NM Bel	0,0,0	NOT SIG
Explicit-Accept	All-Inference	Zero Invalids	0,0,0	AWM > 4

Figure 9.2: Summary of Attitude Strategy Simulations: Strategy 2 is always All-Implicit, Evaluation with COMPOSITE COST. If strategy 1 is beneficial, the AWM range is given. If strategy 1 is not Beneficial, then it is DETRIMENTAL and the AWM range is given, NOT SIG is no differences found.

The results from Design-World for Attitude IRUs presented in section 6.6 are summarized in figure 9.2. In general, Attitude IRUs are beneficial for rehearsal unless communication is too expensive or the cost of retrieval is too high. The figure shows that Explicit-Acceptance has a rehearsal benefit in the standard situation when retrieval is free or a thousandth the cost of communication and inference. When the cost of retrieval is .01, there are no significant differences between Explicit-Acceptance and All-Implicit.

This result only holds at higher AWM because it is at higher AWM that agents have sufficient capacity to retrieve large numbers of beliefs and thsu can confuse outdated beliefs with recent beliefs, making more mistakes as a result. Thus the result in this case is due to the specific assumptions of the belief change mechanism. Nevertheless the result shows that Attention IRUs can have rehearsal benefits whenever the information they contain or the information inferrable from their content must be remembered. Thus the simulations provide an alternate simpler explanation of the benefits of Attitude IRUs in situations where understanding and acceptance are not at issue.

The simulations also show that IRUs should be used with care since they can be detrimental if the cost of communication is high enough. When communication is very expensive (10), the Explicit-Acceptance strategy is detrimental.

In the fault-intolerant ZERO INVALIDS task in which mistakes are heavily penalized, rehearsal is

again beneficial. The Explicit-Acceptance strategy keeps agents from making mistakes at AWM values of 5 and greater.

Explicit-Acceptance has no effect in the task situation which required matching beliefs about the warrants for actions. It is not surprising that it isn't beneficial, since the information in an Explicit-Acceptance is not related to warrants for actions. However there was a possibility that this strategy would be detrimental by displacing warrants from memory. It is possible that this displacement occurs, but is offset by the concomitant rehearsal benefit discussed above.

The theoretical model of Attitude IRUs posits that they can provide 'positive' evidence of understanding (Clark and Schaefer, 1989), and license the inference of acceptance. However the simulations show that another simpler rehearsal benefit is possible and that this benefit distinguishes the use of an utterance like *okay* to indicate acceptance, from the repetitions used in the simulation.

9.4.2 Attention Simulations Summary

Strategy1	Agent Params	Task Params	Costs comm,inf,ret	Beneficial?
Open-Best	All-Inference	Standard	0,0,0	NOT SIG
Open-Best	All-Inference	Standard	1,1,.01	NOT, AWM > 3
Explicit-Warrant	All-Inference	Standard	1,1,0	NOT, AWM 3,4,5
Explicit-Warrant	All-Inference	Standard	1,1,.01	AWM > 3
Explicit-Warrant	All-Inference	Standard	10,0,0	NOT, AWM 1 . . . 3
Explicit-Warrant	All-Inference	Zero NM Bel	0,0,0	AWM 2 . . . 11
Explicit-Warrant	All-Inference	Zero Invalids	0,0,0	AWM 11,16
MP Premise	All-Inference	MatchedPair	0,0,0	AWM 3,4
MP Premise	All-Inference	Standard	0,0,0	NOT SIG
MP Premise	All-Inference	Standard	1,1,.0001	NOT, AWM > 4
MP Prem Warr	All-Inference	MPALL	1,1,.0001	AWM 3,4

Figure 9.3: Summary of Attention Strategy Simulations: Strategy 2 is always All-Implicit, Evaluation with COMPOSITE COST. If strategy 1 is beneficial, the AWM range is given. If strategy 1 is not Beneficial, then it is DETRIMENTAL and the AWM range is given, NOT SIG is no differences found.

Figure 9.3 summarizes the results of the simulations of Attention strategies in Design World. Overall the results showed that Attention IRUs can be beneficial when the content of the IRU consists of information that is targeted at specific inferences that the agent is performing at that time, when inferential complexity is increased, and when retrieval is not free.

The experiments on Open Segment and Deliberation IRUs test the hypothesized DISCOURSE INFERENCE CONSTRAINT and were presented in section 7.6. The Open Segment strategy was Open-Best, which makes premises salient that might be beneficial in means-end reasoning for the put-act intention associated with a discourse segment. This strategy was not beneficial in the standard task situation, despite the fact that it included information which might have been useful. This was because the information displaced other information which would have been just as useful.

I conjecture that if strategies were designed in which the production and content of Open Segment IRUs was more precisely controlled, they would be beneficial. Potention ways of producing this precise control will be discussed in section 9.6.

The Matched-Pair-Premise strategy is very similar to the Open-Best strategy, but it was tested in the MATCHED PAIR task which increases inferential complexity over the standard task (See section 7.6.3). The Matched-Pair-Premise strategy increased agents' ability to make matched-pair inferences and was beneficial. As shown in figure 9.3 the Matched-Pair-Premise strategy is **only** beneficial for the inferentially more complex MATCHED-PAIR version of the task. In the standard task, there are no significant differences if all processing is free. Even worse, if processing has associated costs, then the strategy is detrimental for AWM > 4.

The Deliberation strategy that was tested is called the Explicit-Warrant strategy and was presented in section 7.6.2. In this strategy agents include the warrant for their proposal along with every proposal. This strategy was shown to be beneficial for the Zero-NonMatching-Beliefs task because it makes the information available that is needed to achieve matching beliefs for the warrant for proposals. The strategy is beneficial, even when retrieval is free, as might be expected. In addition, this strategy reduces retrieval costs and thus provides a benefit whenever retrieval is not free. However, if communication is much more expensive than retrieval, then this strategy is not beneficial.

9.4.3 Consequence Simulations Summary

The Design-World simulations discussed in section 8.4.1 and summarized in figure 9.4 show that even when agents are logically omniscient, strategies that make inferences explicit can be beneficial. There are two situations that show this. In the first situation we showed that the Close-Consequence strategy meant that agents are less likely to forget that they have done particular acts so they are less likely to have invalid steps in their plans.

In the second situation, making inferences explicit is beneficial because the task has a higher level of inferential complexity. In this situation, although agents can make all the inferences from their

Strategy1	Agent Params	Task Params	Costs comm,inf,ret	Beneficial?
Close-Conseq	All-Inference	Standard	0,0,0	AWM 11,6;NOT AWM 3,4
Close-Conseq	All-Inference	Standard	1,1,.01	NOt, AWM 1-5
Close-Conseq	All-Inference	Standard	10,0,0	NOT, AWM > 0
Close-Conseq	All-Inference	Zero Invalids	0,0,0	AWM > 3
Close-Conseq	All-Inference	Zero NM Bel	0,0,0	NOT, AWM 7,8
Close-Conseq	NO Inf, NO Inf	Standard	0,0,0	AWM > 0
Close-Conseq	NO Inf, NO Inf	Standard	1,1,.01	AWM > 0
Close-Conseq	NO Inf, NO Inf	Standard	1,1,.0001	AWM > 0
Close-Conseq	NO Inf, All-Inf	Standard	1,1,.0001	AWM > 7; NS
Close-Conseq	Half/Half	Standard	1,1,.01	NOT SIG
MP Inf-Explic	All-Inference	MatchPair	0,0,0	NOT SIG
MP Inf-Explic	All-Inference	MatchPair-2R	0,0,0	AWM 3,4,5,6,11,16

Figure 9.4: Summary of Consequence Strategy Simulations: Strategy 2 is always All-Implicit, Evaluation with COMPOSITE COST. If strategy 1 is beneficial, the AWM range is given. If strategy 1 is not Beneficial, then it is DETRIMENTAL and the AWM range is given, NOT SIG is no differences found.

salient beliefs, the necessary beliefs might not be salient. Making the inference explicit makes it one of the salient beliefs. Thus this simulation support the DISCOURSE INFERENCE CONSTRAINT by showing that IRUs are beneficial when agents are attention limited.

In addition, figure 9.4 shows that when agents are inference limited that making inferences explicit is beneficial. This is exactly what we would expect. More significantly, if their conversational partner uses a strategy of making inferences explicit, No-Inference agents can do just as well as logically omniscient agents. This shows that in situations in which inference is highly unconstrained, a strategy such as Close-Consequence could be very beneficial by indicating exactly which inference should be made.

Finally, the results include testing the difference between Inference-Explicit IRUs with salient or remote antecedents by comparing the MATCHED-PAIR task with the MATCHED-PAIR-TWO-ROOM task. The Matched-Pair-Inference-Explicit strategy makes an Inference-Explicit either when its antecedents are salient or when its antecedents are remote, depending on the version of the MATCHED PAIR task. The strategy was shown to be beneficial when the antecedents for the inference are remote, i.e. the MATCHED-PAIR-TWO-ROOM task. In this situation the Inference-Explicit IRU can help agents make other inferences that are dependent on that inference as a premise. The Explicit-Warrant strategy discussed in chapter 7 has a similar beneficial effect when the task requires beliefs

about warrants to be shared. Thus, this result provides support for the claim that INFERENCE-EXPLICIT IRUs that are realized remotely from their antecedents function in much the same way as Attention IRUs.

The following section discusses possible applications of the results and then section 9.6 will discuss the limitations of the current results and future work.

9.5 Possible Applications of the Results and the Methods

In this section I will discuss four application areas of the results and the theory. In many cases, I am actually discussing applications of the platform and the methods.

9.5.1 Expert Advice Systems

The dialogues use as the corpus for the distributional analysis were financial advice dialogues in which the talk show host was the expert advice giver and the callers sought advice on how to invest their money. Thus the analysis of discourse strategies are probably strategies that are best suited to expert advice systems. Various strategies tested here could be designed into systems like this that rely on natural language generation, e.g. the Explicit-Warrant strategy in which the expert provides a WARRANT for the advice could be included with some idea of the situations in which it might be most useful. One situation identified here is where it is important for both agents to agree on the reasons behind a course of action. In other situations other parameters might also be relevant.

9.5.2 Intelligent Tutoring Systems

Another potential computational application of the results here are to systems for explanation and learning. Generally the advice or tutoring given by these systems is supported by an underlying reasoner in which every step of the reasoning process is available. A major problem with these systems is determining what information to say and what information to leave out, but this is exactly one of the dimensions explored in this work.

In chapter 8, I investigated when making inferences explicit is beneficial. While this only provides a single data point, it would be possible to use and extend the results here. For example, a way to measure and vary the complexity of the domain could be devised and additional Design-World experiments could provide data on which strategies are most effective.

9.5.3 Spoken Language Interfaces

Another potential application area is in the design of dialogue strategies for systems where there is uncertainty in communication such as current speech recognition systems. Current systems have two modes: full paraphrase and no feedback. It seems that the amount of feedback required to avoid errors that are difficult to recover from depends, in all likelihood, on an interaction between the complexity of the task, the degree to which it is fault-tolerant, and the degree of uncertainty in communication. Another set of Design-World experiments could be defined which varied these parameters and investigated when different types of feedback were most useful.

9.5.4 Protocols for Distributed Agents

A final potential application area is in the design of communication protocols for multi-processor environments with heterogenous capabilities. The experiments discussed here investigated the benefits of strategies for inference-limited and attention-limited agents. Agents with other capabilities could also be designed, and the methods used here could demonstrate when certain communication protocols are useful and when they are not.

9.6 Limits of the Results and Future Work

There are a number of specific limitations of the results with respect to the theory argued for here. First, a number of theoretical claims related to Attitude were not tested. In particular, in section 6.2, I argued that Attitude IRUs mainly serve to demonstrate understanding and acceptance. One way to test this hypothesis is to make communication uncertain in Design-World or to have implicit rejection be a possibility. Under these conditions, demonstrating understanding and acceptance would detect errors in communication at the point where they happen, potentially avoiding costly repair and replanning (Brennan and Hulteen, 1993). This hypothesis has not been tested because an evaluation of it depends on a way to introduce errors that might actually occur and the specification of reasonable recovery strategies, neither of which are the topic of this thesis. Carletta provides a theory of recovery strategies in dialogue that could be used to extend the work presented here (Carletta, 1992), but testing hypotheses of how Attitude-Explicit strategies are related to uncertainty in communication and the possibility of rejection must be left to future work.

In addition, I have not tested the differences between different types of Attitude IRUs or discovered

which factors may determine how acceptance is communicated, e.g. by saying nothing, producing a backchannel such as *uh huh*, repeating, paraphrasing or making an inference explicit.

Another limitation of the results here is that it was not possible to distinguish among the various hypotheses presented in chapter 7 about the mechanisms underlying the functionality of Attention IRUs. There I simply argued that the DISCOURSE INFERENCE HYPOTHESIS was a viable hypothesis that could explain all types of Attention IRUs. However the discourse structure and control structures in Design-World do not support testing the COORDINATION HYPOTHESIS and the RETRIEVAL CUE HYPOTHESIS. In order to test these hypotheses one of the agents would have to be ignorant of the structure of the task and there should be some ambiguity about which facts or inference rules are retrieved upon understanding an utterance. Since Design-World has only one possible task structure, there is no ambiguity as to what should be done at a particular point in interaction.

Another limitation is that Design-World only supports testing simple hypotheses about the relationship of inference limitations to discourse strategies. In section 5.9.2 I discussed some simple ways to vary inferential capacity and then showed that there were discourse strategies that were beneficial for inference limited agents. A more complex system of inference in which it was possible to systematically vary the likelihood that an inference would be drawn with a certain amount of resource would allow me to extend the results here and explore additional hypotheses about when it is beneficial to make an inference explicit.

Earlier I mentioned that one of the more interesting results of the simulation is that there is a complex three way interaction between agents cognitive limitations, discourse strategies and the definition of the task. Much more work could be done to explore aspects of this interaction. What has been shown is that IRUs are generally beneficial in tasks where a higher degree of agreement or perfection are required, when there is limited inference, or when there is a high cost to retrieval or retrieval is unreliable. Even with the current set of results, it would be possible to examine the tradeoffs between different costs and benefits in a more systematic fashion, and thus provide better insights into the types of interactions between discourse strategies, agents' cognitive limitations and task parameters.

9.6.1 Limits of the Strategies Tested

The primary limitation of the results is that only discourse strategies related to IRUs were tested. Some of the strategies that include IRUs may actually represent more general strategies. For example, the Matched-Pair-Premise strategy includes information at the beginning of the segment that the speaker believes will be useful for achieving the intention of a segment. This information

doesn't have to be already known and it would be possible to design and test a general strategy such as this, based on the hypothesis that in some circumstances it might be more useful to start out a segment by saying the facts that are considered relevant rather than letting those facts come out over the course of the discussion.

A similar argument could be constructed for Deliberation and Affirmation IRUs. Both of these classes are related to argument structure and providing evidence for conclusions or justifications for beliefs or courses of action. The use of IRUs for this purpose was exemplified by the Explicit-Warrant strategy. IRUs may make especially good warrants, non-IRUs can be warrants as well. Thus another set of experiments could be designed which tested the benefits of always including warrants and attempted to determine under which situations one should do so.

Another range of experiments would involve testing agents with different capabilities. While some results were based on dialogues where one agent used one strategy and the other agent used another, no situations were tested in which one agent had unlimited attentional capacity and the other agent was extremely attention limited or one agent had unlimited inferential capacity and the other agent was inference limited. I would also expect these factors to interact with which agent has most of the information about the task (Walker and Whittaker, 1990). In many expert systems, there are differences between agents about what knowledge they have. It would be extremely simple to perform these tests with the current system, but this was not the focus of the work presented here.

Another limitation is that the strategies tested were always compared with the All-Implicit strategy, but there may be interesting differences and interesting interactions between IRU strategies as well. A way of testing this will be discussed in the next section.

Finally, it would be possible to extend the system so that the effect of text structure and different thematisations could be explored. Currently the agents communicate with a propositional language and there is no representation of information such as INFORMATION FOCUS or CENTER. For example, a sample hypothesis would be that the encoding or retrieval cue for propositions is either the INFORMATION FOCUS or the CENTER of an utterance and that this is one way agents deal with their limited attentional capacity. Another hypothesis would be that the reason that many utterances consist in large part of HEARER OLD information because humans cannot process very much new information at a time. The HEARER OLD information in an utterance is the 'glue' that makes it possible to process the new information. Some versions of these hypotheses might be testable in a suitably modified version of the current system in which processing could be made easier by differential marking of HEARER OLD and NEW information.

9.6.2 IRUs in Plan-Based Generation and Recognition Systems

In this work, agents communicate in an artificial language and each utterance intention is explicitly communicated. IRUs are the content of these utterances and an IRU can be the content of the utterance intentions of ACCEPT, OPEN, CLOSE or SAY. In each of these cases the content of the utterance is asserted in memory and the utterance intention is used by the listener to decide what to do next. But how could we recognize the utterance intention for an IRU if it wasn't communicated directly?

This is a type of plan-inference or plan-recognition, studied by many researchers in discourse (Allen, 1979; Sidner and Israel, 1981; Sidner, 1985; Litman and Allen, 1990; Carberry, 1989). In Chapter 2, I mentioned that the axioms and plan-recognition heuristics in these plan-based systems disallow the occurrence of IRUs or suggest that IRUs be interpreted as indirect speech acts (Perrault, 1990). Instead I have argued that the interpretation of IRUs does not require resorting to an inference process based on their non-informativeness.

There are however several different ways in which the treatment of IRUs might be incorporated into speech-act based accounts. One way is to introduce a new speech act for each type of IRU. In order to generate IRUs, the system would reason about generating these new speech act types. Then, in the plan recognition system, IRUs would not be recognized as INFORM speech acts but rather as one of the newly introduced speech act types. This treatment depends on recognizing IRUs as redundant in order to eliminate INFORM as a possible recognized plan. This could possibly be done by prosody, but in other cases could involving checking a large number of possible inferences.

Another way in which IRUs could be incorporated into these systems is to change the level of granularity of the primitive act definition. Thus, rather than the uttering of a single proposition constituting an act, a schema for multiple propositions might be reasoned about, in which an IRU was part of the schema and communicated because of the relation that it bears to other propositions that are conveyed at the same time.[1] The details of this treatment is beyond the scope of this thesis, but the incorporation of IRUs in messages in the Design-World system presented in chapter 5, and used throughout the thesis actually approximates the multiple proposition schema briefly sketched here.

A final way in which at least the Attention IRUs could be incorporated into a speech-act based system is by changing the definition of the felicity conditions for an INFORM to reflect the distinction between SALIENT and HEARER OLD propositions advocated here. The new felicity conditions would mean that INFORM utterances are felicitous if the proposition realized by that utterance is not

[1] Cawsey (P.C. Spring92) includes redundant propositions in explanations when the explanation would seem incoherent without them (Cawsey, 1992).

currently salient. The plan-generation heuristics then can generate an INFORM and the plan-inference heuristics can recognize an utterance as an INFORM as long as their content is not currently salient. However, testing the details of this account must be left to future work.

9.6.3 Adaptability, Learning, and Optimization of Strategies

One limitation of the simulation design is that Design-World agents are parametrized as to discourse strategy and each discourse strategy must be designed into the system. Furthermore, an agent with a strategy uses that strategy persistently without reflection. In other words, an agent always does the same discourse level action for each discourse sequence defined situation. Finally, agents carry out these strategy-defined behaviors without using any knowledge of their conversational partners.

There are at least three ways in which these limitations could be addressed. First, it would be possible to have strategies based on an agent's own reflections. Second, it would be possible to have strategies based on a model of the conversational partner. Third, it might be possible to automatically evolve complex strategies based on a set of primitive strategies through the use of techniques such as genetic algorithms. Each of these extensions will be discussed in the following sections.

9.6.3.1 Reflective Strategies

In the current version of Design-World, agents are parameterized for strategy in a rather rigid way. For example, if an agent's strategy is to include extra information in a proposal, the agent will **always** include that information in a proposal. However the results discussed in sections 7.6.3 and 8.4.4.3 show that this strategy is most beneficial when the extra information that is included is not currently salient or would be costly to retrieve. This suggests that better versions of some of the strategies could be designed in which Attention IRUs are only produced when their content is not currently salient.

This leaves open the question of how agents determine when a proposition is not currently salient for another agent. One heuristic way to do this is to have an agent use its own mind as an approximation of the other agent. An agent knows which propositions are salient for it, and an agent can keep track of the amount of processing it took it to retrieve a fact or deduce a conclusion. This information could be used as an an estimate of the amount of processing it would require for the other agent to produce the same result (Joshi, 1982; Joshi, Webber, and Weischedel, 1984). If the cost of communicating the fact or the inference is less than this processing estimate, then the agent would include the IRU.

One aspect of this form of reflective strategy is already incorporated into the current system. Agents select from among the facts currently in use in their own means-end reasoning process to decide which facts to include as IRUs in the OPEN-BEST and MATCHED-PAIR based strategies.[2] This selection process could be based instead on a model of what information would be most useful to the other conversant. This is the topic of the next section.

9.6.3.2 Using a Model of the Conversational Partner

Design-World agents do not currently maintain information about what other agents know, or what their capabilities are, so that the agent strategies do not rely on such information. There are several types of information about the other agent or the history of the interaction that agents could use to adapt their strategies to their conversational partners. First, agents might keep track of what information they believe their partners know and/or have currently salient. Second, if agents kept information about discourse strategies, they could recognize which discourse strategy the other agent is using and modify their own strategy accordingly. Third, an agent might notice when another agent makes mistakes say by proposing an action whose preconditions are not met, and adjust its own strategy accordingly. Finally, without tracking any particular aspect of the partner's behavior, an agent might randomly try variations in strategies and keep track of the scores achieved with those strategies, adjusting strategies to those that seem to work best with that partner.

Another aspect of modeling the conversational partner might be to design strategies that take the timing in interaction into account. Timing information may be used by agents to determine when to elaborate on what they have previously asserted, for example by paraphrasing what they have said or by making an Inference-Explicit.

It would be possible to set up Design-World experiments that would test whether it is useful to model the other conversant. For example, one could compare a strategy that relies on one's own reasoning as the basis for selecting the content of IRUs, such as the strategies discussed above, with strategies that require maintaining a model of the user and making predictions based on reasoning about this model.

[2]Because score propositions are the only type of warrant propositions used in deliberation, no selection process is needed to determine what to include in the EXPLICIT-WARRANT strategy. If the system were extended so that other kinds of warrants or supports were possible, then a similar strategy of selection based on ones own reasoning would be possible.

9.6.3.3 Optimization and the Development of Complex Strategies

All of the current Design-World strategies consist of changing the way that one discourse level action is performed, e.g. whether an opening is done implicitly or explicitly or whether a warrant is included in a proposal. These choices are always binary, a strategy either always includes an explicit opening or it never includes one. Section 9.6.3.1 suggested one way of varying the binary aspect of these strategies based on the agents' reflection about their own reasoning processes. Another possibility would be to take a godlike view of agents' behavior and attempt to optimize it by using a technique such as genetic algorithms that randomly create new strategies as combinations of old ones and use the performance measure in Design-World to decide which strategies will be kept through successive generations (Goldberg, 1989). The remainder of this section discusses in more detail how this might be done.

Genetic algorithms (GAs) are a stochastic learning and optimization technique that relies on two things: (1) it must be possible to code the parameters of the optimization problem as as a vector string, (2) a payoff function must be defined for the performance of the algorithm and this is associated with each parameter vector string. In Design-World, the parameters of the algorithm are the agent parameters. These include what strategies are accessible to an agent, the agent's inferential capacity and the agent's attentional capacity. The payoff function is the performance measure for Design-World. The fact that performance is dependent on assumptions about processing costs will be discussed further below.

Application of GAs involves 3 steps :

1. Defining an original population
2. Testing the fitness of this population
3. Reproducing a new generation modeling a notion of 'survival of the fittest'. Then use this new generation to repeat steps 2 and 3.

An original population is defined that consists of a random selection of vector strings representing a large enough subset of total population. The problem is then attempted with the members of this original population and population members are replicated in the next generation in proportion to the degree of success that they had. New vector strings can be introduced into the population by 'mating' two existing vector strings, with subsets of parameters randomly traded between the mating strings. After a number of generations, the optimal parameter string should dominate the population. Below I will consider the specifics of a Design-World experiment using this method.

Parameter Name	Parameter Values
Attitude-Explicit	0 , 1
Inference-Explicit	0 , 1
Open-Best	0 , 1
Explicit-Warrant	0 , 1
Inference Capacity	No, Half, All
Attentional Capacity	1 . . . 16

Figure 9.5: Possible Genes for a Design-World Genetic Algorithm Experiment

Parameter	Agt-1 Pars	Agt-2 Pars	Agt-3 Pars	Agt-4 Pars
Attitude-Explicit	0	1	1	0
Inference-Explicit	0	1	0	0
Open-Best	0	0	0	0
Explicit-Warrant	0	0	0	0
Inference Capacity	All	All	Half	No
Attentional Capacity	7	3	11	16

Figure 9.6: Possible Original Population for a Design-World Genetic Algorithm Experiment

Imagine that the parameters of the algorithm were the strategies of the agents that we have seen. We could set up a genetic algorithm experiment in which the probability of using a strategy was represented by parameter values that would take on discrete settings between 0 and 1. For simplicity, let's consider 4 different strategies as represented in the table below, and assume that they are either 'on' or 'off' as in the current implementation. We let inferential capacity take the three discrete settings we used in chapter 8 and attentional capacity range from 1 to 16 as in all the experiments presented.

An original population of agents randomly selected as a combination of these parameters could be as shown in figure 9.6.

The next step is to test the fitness of this population of agents. Since fitness has to do with the scores that the agents achieve in dialogue, each agent carries out N dialogues with each of the other agents. In previous experiments N was 100. Each agent in a dialogue gets the score that is the joint score for the dialogue. Assuming that each agent participates in the same number of dialogues, we can simply sum performance over all the dialogues with all the other agents to get an individual agent's fitness. Then we sum performance over all the agents to get a total fitness measure.

Reproduction is based on the notion of 'survival of the fittest'. An agent's fitness divided by the population's overall fitness gives the likelihood that that agent will be replicated in the next generation. For example, if agent A represents 50% of the total fitness and agent B represents 25% of the total fitness, then a copy of A is twice as likely to appear in the next generation. We then randomly select a subset of the replicated population and 'mate' agents in the subset by 'cross-over', which consists of randomly exchanging some parameter settings between agents. This introduces new types of agents into the population. It is also possible to have mutations as part of reproduction, which would involve randomly changing a parameter setting at some low frequency.

After some number of generations, the idea is that we would get a stable population of agent types that don't change from one generation to another.

One remaining issue with conducting such an experiment in Design-World is the fact that the fitness of an agent has been shown to depend on features of the situation (environment) such as the cost of communication, inference and retrieval and the definition of the task. In order to use GAs, these environmental features would have to be fixed for the duration of some number of dialogue runs. Fitness over a range of situations might be desirable, so we might want to define performance as performance over a range of situations.

The benefits of the GA method are that if a novel situation arises, we can set the situation parameters, generate a population and then see if any of the currently defined strategies work well in that situation.

The experiment described here is similar to the Tit-For-Tat experiments (Axelrod, 1984). One issue with the design is the fact that each agent's fitness is determined by summing their performance with all the other agents in the population. This procedure loses information about effective combinations of strategies, but one of the findings of the experiments so far is that particular combinations of agents perform better than other combinations. Fitness is dependent on ones conversational partner. Perhaps over successive runs pairs of agents might form stable populations precisely because their interaction is beneficial. Another possibility would be that the population members are 'teams' of agents, i.e. the pairs of agents who do the task together.

9.7 Summary

This thesis focused on explaining the phenomenon of IRUs, a type of action in dialogue which appears patently inefficient. I've argued that the occurrence of IRUs can be explained by positing that cognitive agents are resource bounded actors.

The approach that I've taken is to argue for an account in which the 'meaning' of an IRU is emergent from underlying processing constraints. The account here makes minimal assumptions about the use of convention in determining the meaning of IRUs and avoids positing that the interpretation of IRUs relies on reflexive Gricean reasoning about violations of the MAXIM OF QUANTITY. While it is possible that Gricean reasoning is used in the interpretation of IRUs, I've shown that it is also possible to interpret IRUs without resorting to this type of reasoning.

The effects of limited processing on agents' behavior is currently an active area of research (Konolige, 1986; Pollack and Ringuette, 1990; Bratman, Israel, and Pollack, 1988), but models of limited processing have not previously been used to explain agents' discourse behavior. I've shown that discourse strategies that include IRUs are beneficial under specific assumptions about agents resource limitations. Thus, the main contribution of the thesis is in linking discourse behavior to underlying cognitive limitations.

Bibliography

James F. Allen and C. Raymond Perrault. 1980. Analyzing intention in utterances. *Artificial Intelligence*, 15:143–178.

James F. Allen. 1979. A plan-based approach to speech act recognition. Technical report, University of Toronto.

James F. Allen. 1983. Recognizing intentions from natural language utterances. In M. Brady and R.C. Berwick, editors, *Computational Models of Discourse*. MIT Press.

Jens Allwood, Lars-Gunnar Andersson, and Osten Dahl. 1977. *Logic in Linguistics*. Cambridge University Press.

Jens S. Allwood. 1992. On the semantics and pragmatics of linguistic feedback. *Journal of Semantics*, 9:1–26.

J. R. Anderson and G. H. Bower. 1973. *Human Associative Memory*. V.H. Winston and Sons.

J. R. Anderson. 1974. Verbatim and propositional representation of sentences in immediate and long term memory. *Journal of Verbal Learning and Verbal Behavior*, 13:149–162.

Jean-Claude Anscombre and Oswald Ducrot. 1983. *L'argumentation dans la langue*. Pierre Mardaga, Bruxelles.

J.L. Austin. 1965. *How to do things with words*. Oxford University Press.

Robert Axelrod. 1984. *The Evolution of Cooperation*. Basic Books.

Alan Baddeley. 1986. *Working Memory*. Oxford University Press.

John Barwise. 1988a. *The situation in Logic*. CSLI.

Jon Barwise. 1988b. The situation in logic-iv: On the model theory of common knowledge. Technical Report No. 122, CSLI.

Betty Birner. 1992. *Inversion in English*. Ph.D. thesis, Northwestern University.

Charles P. Bloom, Charles R. Fletcher, Paul Van Den Broek, Laura Reitz, and Brian P. Shapiro. 1990. An on-line assessment of causal reasoning during comprehension. *Journal of Memory and Cognition*.

Sara S. Bly. 1988. A use of drawing surfaces in different collaborative settings. In *Proceedings of the Conference on CSCW*.

J. D. Bransford, J. R. Barclay, and J. J. Franks. 1972. Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, 3:193–209.

Michael Bratman, David Israel, and Martha Pollack. 1988. Plans and resource bounded practical reasoning. *Computational Intelligence*, 4:349–355.

Susan E. Brennan and Eric A. Hulteen. 1993. Interaction and feedback in a spoken language system. In *AAAI Workshop on Human-Computer Collaboration*, pages 1–6.

Susan E. Brennan. 1990. *Seeking and Providing Evidence for Mutual Understanding.* Ph.D. thesis, Stanford University Psychology Dept. Unpublished Manuscript.

Derek G. Bridge. 1991. *Computing Presuppositions in an incremental natural language processing system.* Ph.D. thesis, Cambridge University. Technical report number 237, Cambridge Computer Lab.

Penelope Brown and Steve Levinson. 1987. *Politeness: Some universals in language usage.* Cambridge University Press.

Bertram Bruce. 1975. Belief systems and language understanding. Technical Report AI-21, Bolt, Berenak and Newman.

S. Carberry. 1989. Plan recognition and its use in understanding dialogue. In A. Kobsa and W. Wahlster, editors, *User Models in Dialogue Systems*, pages 133–162. Springer Verlag, Berlin.

Jean C. Carletta. 1992. *Risk Taking and Recovery in Task-Oriented Dialogue.* Ph.D. thesis, Edinburgh University.

Alison Cawsey, Julia Galliers, Steven Reece, and Karen Sparck Jones. 1992. Automating the librarian: A fundamental approach using belief revision. Technical Report 243, Cambridge Computer Laboratory.

A. Cawsey. 1992. Planning interactive explanations. *International Journal of Man-Machine Studies.* (in press).

Wallace L. Chafe. 1976. Givenness, contrastiveness, definiteness, subjects, topics and points of view. In Charles N. Li, editor, *Subject and Topic*, pages 27–55. Academic Press.

Herbert H. Clark and Susan E. Brennan. 1990. Grounding in communication. In L. B. Resnick, J. Levine, and S. D. Bahrend, editors, *Perspectives on socially shared cognition.* APA.

Herbert H. Clark and T. B. Carlson. 1982. Speech acts and hearer's beliefs. In Neil V. Smith, editor, *Mutual Knowledge*, pages 1–37. Academic Press, New York, New York.

Herbert H. Clark and Catherine R. Marshall. 1981. Definite reference and mutual knowledge. In Joshi, Webber, and Sag, editors, *Elements of Discourse Understanding*, pages 10–63. CUP, Cambridge.

Herbert H. Clark and Edward F. Schaefer. 1987. Collaborating on contributions to conversations. *Language and Cognitive Processes*, 2:19–41.

Herbert H. Clark and Edward F. Schaefer. 1989. Contributing to discourse. *Cognitive Science*, 13:259–294.

Herbert H. Clark and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition*, 22:1–39.

P.R. Cohen and H.J. Levesque. 1985. Speech acts and rationality. In *Proceedings of the twenty-third Annual Meeting of the Association of Computational Linguistics*, pages 49–60.

Phillip R. Cohen and Hector Levesque. 1990. Rational interaction as the basis for communication. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press.*

Phil Cohen and Hector Levesque. 1991. Teamwork. *Nous*, 25:11–24.

Phillip R. Cohen. 1978. On knowing what to say: Planning speech acts. Technical Report 118, University of Toronto; Department of Computer Science.

Robin Cohen. 1981. Investigation of processing strategies for the structural analysis of arguments. In *Association of Computational Linguistics.*

P. R. Cohen. 1985. *Heuristic Reasoning about Uncertainty: an Artificial Intelligence Approach.* Pitman, Boston.

Robin Cohen. 1987. Analyzing the structure of argumentative discourse. *Computational Linguistics*, 13:11–24.

A. M. Collins and M. R. Quillian. 1969. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8:240–247.

Cruttenden. 1986. *Intonation.* Cambridge University Press.

Anne Cutler and D. J. Foss. 1977. On the role of sentence stress in sentence processing. *Language and Speech*, 29:1–10.

Anne Cutler. 1976. Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception and Psychophysics*, 20:55–60.

Dean and Boddy. 1988. An analysis of time dependent planning. In *AAAI88*.

Barbara Di Eugenio. 1993. *Understanding Natural Language Instructions: a Computational Approach to Purpose Clauses*. Ph.D. thesis, University of Pennsylvania.

Jon Doyle. 1992. Rationality and its roles in reasoning. *Computational Intelligence*, November.

Ronald Fagin and Joseph Halpern. 1985. Belief, awareness and limited reasoning: Preliminary report. In *Proc. International Joint Conference on Artificial Intelligence, Los Angeles*, pages 491–501.

Timothy W. Finin, Aravind K. Joshi, and Bonnie Lynn Webber. 1986. Natural language interactions with artificial experts. *Proceedings of the IEEE*, 74(7):921–938.

Charles R. Fletcher, John E. Hummel, and Chad J. Marsolek. 1990. Causality and the allocation of attention during comprehension. *Journal of Experimental Psychology*.

B.A. Fox. 1987. Interactional reconstruction in real-time language processing. *Cognitive Science*, 11:365–388.

Julia R. Galliers. 1990. Belief revision and a theory of communication. Technical Report 193, University of Cambridge, Computer Laboratory, New Museums Site, Pembroke St. Cambridge England CB2 3QG.

Julia R. Galliers. 1991a. Autonomous belief revision and communication. In P. Gardenfors, editor, *Belief Revision*, pages 220 – 246. Cambridge University Press.

Julia R. Galliers. 1991b. Cooperative interaction as strategic belief revision. In M.S. Deen, editor, *Cooperating Knowledge Based Systems*, pages 148 – 163. Springer Verlag.

Peter Gardenfors. 1988. *Knowledge in flux : modeling the dynamics of epistemic states*. MIT Press.

Peter Gardenfors. 1990. The dynamics of belief systems: Foundations vs. coherence theories. *Revue International De Philosophie*.

Gerald Gazdar. 1979. *Pragmatics : implicature, presupposition and logical form*. Academic Press.

David Goldberg. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison–Wesley.

H. P. Grice. 1967. *William James Lectures*.

J. Groenendijk and M. Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy*, 14(1):39–100.

Barbara J. Grosz and Candace L. Sidner. 1986. Attentions, intentions and the structure of discourse. *Computational Linguistics*, 12:175–204.

Barbara J. Grosz and Candace L. Sidner. 1990. Plans for discourse. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press.*

Barbara J. Grosz. 1977. The representation and use of focus in dialogue understanding. Technical Report 151, SRI International, 333 Ravenswood Ave, Menlo Park, Ca. 94025.

Curry I. Guinn. 1993. A computational model of dialogue initiative in collaborative discourse. In *AAAA Fall Symposium on Human Computer Collaboration.*

M.A.K. Halliday and Ruquaiya Hasan. 1976. *Cohesion in English.* Longman.

Joseph Halpern and Yoram Moses. 1985. A guide to the modal logics of knowledge and belief. In *Proc. International Joint Conference on Artificial Intelligence, Los Angeles*, pages 480–490.

Hamblin. 1971. A mathematical model of dialogue. *Theoria*, 37:130–155.

Gilbert Harman. 1986. *Change of View.* MIT PRESS.

Marti Hearst. 1993. Text tiling: A quantitative approach to discourse segmentation. Technical report, University of California, Berkeley, Sequoia.

Peter A. Heeman and Graeme Hirst. 1992. Collaborating on referring expressions. Technical Report 435, University of Rochester.

Irene Heim. 1982. *The Semantics of Definite and Indefinite Noun Phrases.* Ph.D. thesis, University of Massachusets, Amherst.

S. Hellyer. 1962. Frequency of stimulus presentation and short-term decrement in recall. *Journal of Experimental Psychology*, 64:650.

Heritage and Watson. 1979. Reformulations as conversational objects. In George Psathas, editor, *Everyday language: Studies in Ethnomethodology*, pages 123–163. Irvington Publishers, New York.

D. L. Hintzmann and R. A. Block. 1971. Repetition and memory: evidence for a multiple trace hypothesis. *Journal of Experimental Psychology*, 88:297–306.

Julia Hirschberg and Diane Litman. 1987. Now lets talk about now: Identifying cue phrases intonationally. In *Proc. 25th Annual Meeting of the ACL, Stanford*, pages 163–171, Stanford University, Stanford, Ca.

Julia Hirschberg. 1985. *A Theory of Scalar Implicature*. Ph.D. thesis, University of Pennsylvania, Computer and Information Science.

Jerry R. Hobbs. 1979. Coherence and coreference. *Cognitive Science*, 3:67–90.

Jerry R. Hobbs. 1985. On the coherence and structure of discourse. Technical Report CSLI-85-37, Center for the Study of Language and Information, Ventura Hall, Stanford University, Stanford, CA 94305.

Beth A Hockey. 1991. The function of okay as a cue phrase in discourse. In *AAAI Fall symposium on Discourse Structure*, November.

Laurence R. Horn. 1972. *On the semantic properties of logical operators in English*. Ph.D. thesis, University of California at Los Angeles. Distributed by the Indiana University Linguistics Club, 1976.

Laurence R. Horn. 1986. Presupposition, theme and variations. In *Chicago Linguistic Society, 22*, pages 168–192.

Lawrence R. Horn. 1989. *A natural history of negation*. Chicago.

Laurence R. Horn. 1991. Given as new: When redundant affirmation isn't. *Journal of Pragmatics*, 15:305–328.

G. Houghton and S. Isard. 1985. Why to speak, what to say and how to say it : Modelling language production in discourse. In *Proceedings of the International Workshop on Modelling Cognition*, University of Lancaster.

P. N. Johnson-Laird. 1991. *Deduction*. Hove, U.K. ; Hillsdale, U.S. : L. Erlbaum.

Aravind K. Joshi and Scott Weinstein. 1981. Control of inference: Role of some aspects of discourse structure - centering. In *Proc. International Joint Conference on Artificial Intelligence*, pages 385–387.

Aravind K. Joshi, Bonnie Lynn Webber, and Ralph M. Weischedel. 1984. Preventing false inferences. In *COLING84: Proc. 10th International Conference on Computational Linguistics.*, pages 134–138.

Aravind K. Joshi, Bonnie Lynn Webber, and Ralph M. Weischedel. 1986. Some aspects of default reasoning in interactive discourse. Technical Report MS-CIS-86-27, University of Pennsylvania, Department of Computer and Information Science.

Aravind K. Joshi. 1964. Informationally equivalent sentences. In *International Conference on Microwaves, Circuit Theory and Information Theory.*

Aravind K. Joshi. 1978. Some extensions of a system for inference on partial information. In *Pattern Directed Inference Systems*, pages 241–257. Academic Press.

Aravind K. Joshi. 1982. Mutual beliefs in question-answer systems. In Neil V. Smith, editor, *Mutual Knowledge*, pages 181–199. Academic Press, New York, New York.

Daniel Kahneman, Paul Slovic, and Amos Tversky. 1982. *Judgment under uncertainty : heuristics and biases.* Cambridge University Press.

Lauri Karttunen and Stanley Peters. 1979. Conventional implicature. *Syntax and Semantics*, 11.

Lauri Karttunen. 1973. Presuppositions of compound sentences. *Linguistic Inquiry*, 4(2):169–193, Spring.

Lauri Karttunen. 1976. Discourse referents. In J. McCawley, editor, *Syntax and Semantics*, volume 7. Academic Press.

W. Kintsch. 1988. The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95:163–182.

Carol Kiparsky and Paul Kiparsky. 1970. Fact. In Manfred Bierwisch and Karl E. Heidolph, editors, *Progress in Linguistics.* Mouton.

Kurt Konolige. 1985. Belief and incompleteness. In Jerry R. Hobbs and Robert C. Moore, editors, *Formal Theories of the Commonsense World*, pages 359–403. Ablex Publishing.

Kurt Konolige. 1986. *A Deduction Model of Belief.* Brown University Press.

Susumu Kuno. 1987. *Functional Syntax.* Chicago University Press.

Robert Ladd. 1980. *The Structure of Intonational Meaning: Evidence from English.* University of Indiana Press. A Cornell University Dissertation.

Thomas K. Landauer. 1969. Reinforcement as consolidation. *Psychological Review*, 16:82–96.

Thomas K. Landauer. 1975. Memory without organization: Properties of a model with random storage and undirected retrieval. *Cognitive Psychology*, pages 495–531.

Hector J. Levesque, Phillip R. Cohen, and Jose H. T. Nunes. 1990. On acting together. In *AAAI90.*

Hector Levesque. 1984. A logic of implicit and explicit belief. In *AAAI84*, pages 198–202.

Stephen C. Levinson. 1979. Activity types and language. *Linguistics*, 17:365–399.

Stephen C. Levinson. 1981. Some pre-observations on the modelling of dialogue. *Discourse Processes*, 4:93–116.

Stephen C. Levinson. 1983. *Pragmatics*. Cambridge University Press.

Stephen C. Levinson. 1985. What's special about conversational inference. In *1987 Linguistics Institute Packet.*

David Lewis. 1969. *Convention*. Harvard University Press.

David Lewis. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic*, 8:339–359.

Mark Y. Liberman and Cynthia A. McLemore. 1992. The structure and intonation of business telephone openings. *The Penn Review of Linguistics*, 16:68–83.

Mark Y. Liberman. 1975. *The Intonational System of English*. Ph.D. thesis, MIT. (published by Garland Press, NY, 1979).

Diane Litman and James Allen. 1990. Recognizing and relating discourse intentions and task-oriented plans. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press.*

Diane Litman. 1985. Plan recognition and discourse analysis: An integrated approach for understanding dialogues. Technical Report 170, University of Rochester.

W.C. Mann and S.A. Thompson. 1987. Rhetorical structure theory: Description and construction of text structures. In Gerard Kempen, editor, *Natural Language Generation*, pages 83–96. Martinus Nijhoff.

John McCarthy, M. Sato, T. Hayashi, and S. Igarashi. 1978. On the model theory of knowledge. Technical Report AIM-312, CS78-657, Stanford University.

Kathleen R. McKeown. 1983. Paraphrasing questions using given and new information. *Computational Linguistics*, Jan-Mar.

Gail McKoon and Roger Ratcliff. 1992. Inference during reading. *Psychological Review*, 99(3):440–466.

Cynthia A. McLemore. 1991. *The pragmatic Interpretation of English Intonation: Sorority Speech.* Ph.D. thesis, University of Texas, Austin.

Cynthia A. McLemore. 1992. Prosodic variation across discourse types. In *Workshop on Prosody in Natural Speech.* Institute for Research in Cognitive Science, University of Pennsylvania.

Igor A. Melčuk. 1988. *Dependency Syntax: Theory and Practice.* SUNY.

G. A. Miller. 1956. The perception of speech. In M. Halle, editor, *For Roman Jakobsen*, pages 353–359. The Hague, Mouton Publishers.

Johanna D. Moore and Cécile L. Paris. 1989. Planning text for advisory dialogues. In *Proc. 27th Annual Meeting of the Association of Computational Linguistics*, Vancouver. ACL.

Johanna D. Moore and Martha E. Pollack. 1992. A problem for rst: The need for multi-level discourse analysis. *Computational Linguistics*, 18(4).

Morris and G. Hirst. 1991. Lexical cohesion computed by thesaural relations as an indicator of the structure of text. *Computational Linguistics*, 17(1):21–48.

Margaret G. Moser. 1992. *The Negation Relation: Semantic and Pragmatic Aspects of a Relational Analysis of Sentential Negation.* Ph.D. thesis, University of Pennsylvania.

Gopalan Nadathur and Aravind K. Joshi. 1983. Mutual beliefs in conversational systems: Their role in referring expressions. In *Proc. International Joint Conference on Artificial Intelligence, Austin.*

S. G. Nooteboom and J. G. Kruyt. 1987. Accents, focus distribution, and the perceived distribution of given and new information: an experiment. Technical Report 538/V, IPO, Netherlands.

Donald A. Norman and Daniel G. Bobrow. 1975. On data-limited and resource-limited processes. *Cognitive Psychology*, 7(1):44–6.

R. C. Oldfield and A. Wingfield. 1965. Response latencies in naming objects. *Quarterly Journla of Experimental Psychology*, 17:273–281.

Rohit Parikh. 1990. Recent issues in reasoning about knowledge. Technical report, Brooklyn College of CUNY.

Barbara Partee. 1982. Semantics: Psychology or mathematics? In Stanley Peters and Esa Saarinen, editors, *Processes, beliefs, and questions : essays on formal semantics of natural language and natural language processing.* Kluwer Boston.

Rebecca J. Passonneau and Diane Litman. 1993. Intention-based segmentation. In *Proceedings of the 31st Meeting of the Association of Computational Linguistics.*

Ray Perrault. 1990. An application of default logic to speech-act theory. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press.*

Janet Pierrehumbert and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press.*

Janet Pierrehumbert. 1980. *The Phonetics and Phonology of English Intonation.* Ph.D. thesis, MIT.

Kenneth L. Pike. 1945. *The Intonation of American English.* University of Michigan Press.

Livia Polanyi and Remko Scha. 1984. A syntactic approach to discourse semantics. In *COLING84.*

Livia Polanyi. 1987. The linguistic discourse model: Towards a formal theory of discourse structure. Technical Report 6409, Bolt Beranek and Newman, Inc., Cambridge, Mass.

Martha E. Pollack and Marc Ringuette. 1990. Introducing the Tileworld: Experimentally Evaluating Agent Architectures. In *AAAI90*, pages 183–189.

Martha Pollack, Julia Hirschberg, and Bonnie Webber. 1982. User participation in the reasoning process of expert systems. In *AAAI82.*

Martha Pollack. 1986. Inferring domain plans in question answering. Technical Report 403, SRI International - Artificial Intelligence Center. University of Pennsylvania Dissertation.

Martha E. Pollack. 1990. Plans as complex mental attitudes. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press.*

Richard Power. 1974. *A Computer Model of Conversation.* Ph.D. thesis, University of Edinburgh.

Richard Power. 1984. Mutual intention. *Journal for the Theory of Social Behaviour*, 14.

Ellen F. Prince. 1978. On the function of existential presupposition in discourse. In *Papers from 14th Regional Meeting.* CLS, Chicago, IL.

Ellen F. Prince. 1981a. Topicalization, focus movement and yiddish movement:a pragmatic differentiation. In D. Alford etal., editor, *Proceedings of the Seventh Annual Meeting of the Berkely Linguistics Society*, pages 249–264. BLS.

Ellen F. Prince. 1981b. Toward a taxonomy of given-new information. In *Radical Pragmatics.* Academic Press.

Ellen F. Prince. 1985. Fancy syntax and shared knowledge. *Journal of Pragmatics*, pages 65–81.

Ellen F. Prince. 1986. On the syntactic marking of the presupposed open proposition. *CLS86.*

Ellen F. Prince. 1992. The ZPG letter: Subjects, definiteness and information status. In S. Thompson and W. Mann, editors, *Discourse description: diverse analyses of a fund raising text*, pages 295–325. John Benjamins B.V.

Roger Ratcliff and Gail McKoon. 1988. A retrieval theory of priming in memory. *Psychological Review*, 95(3):385–408.

Michael Reddy. 1979. The conduit metaphor – a case of frame conflict in our language about language. In A. Ortony, editor, *Metaphor and Thought*, pages 284–324. Cambridge University Press, Cambridge.

Rachel Reichman. 1985. *Getting Computers to Talk Like You and Me*. MIT Press, Cambridge, MA.

Philip Resnik. 1993. *Selection and Information:A Class-Based Approach to Lexical Relationships*. Ph.D. thesis, University of Pennsylvania, Computer and Information Science.

Craige Roberts. 1993a. Centering and anaphora resolution in discourse representation theory. In *Institute for Research in Cognitive Science Workshop on Centering Theory in Naturally-Occurring Discourse*.

Craige Roberts. 1993b. Domain restriction in dynamic semantics. Kluwer,Dordrecht.

Mats Rooth. 1985. *Association with Focus*. Ph.D. thesis, Linguistics Dept, University of Massachusetts, Amherst.

Lee Ross and Craig A. Anderson. 1982. Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments. In *Judgment under uncertainty : heuristics and biases*. Cambridge University Press.

Jacqueline D. Sachs. 1967. *Recognition memory for syntactic and semantic aspects of connected discourse*. Ph.D. thesis, Univeristy of California Berkeley.

Harvey Sacks, Emmanuel Schegloff, and Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking in conversation. *Language*, 50:325–345.

Jerrold M. Sadock. 1977. Modus brevis: The truncated argument. In *Papers from the 13th meeting of the CLS*. Chicago Linguistic Society.

Jerrold M. Sadock. 1978. On testing for conversational implicature. In Peter Cole, editor, *Syntax and Semantics, Volume 9: Pragmatics*, pages 281–297. Academic Press.

Emanuel A. Schegloff and Harvey Sacks. 1977. Opening up closings. *Semiotica*, 8:289–327.

Emanuel A. Schegloff. 1982. Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In D. Tannen, editor, *Analyzing Discourse: Text and Talk*, pages 71–93. Georgetown University Press.

Emanuel A. Schegloff. 1987. Between micro and macro: Contexts and other connections. In *The Micro-Macro Link*. University of California Press.

Emanuel A. Schegloff. 1990. On the organization of sequences as a source of coherence in talk-in-interaction. In Bruce Dorval, editor, *Conversational Coherence and Its Development*. Ablex.

Thomas C. Schelling. 1960. *The Strategy of Conflict*. Harvard University Press.

Stephen R. Schiffer. 1972. *Meaning*. Clarendon Press.

Deborah Schiffrin. 1982. Cohesion in everyday discourse: The role of paraphrase. *Sociolinguistics Working Paper*, 97.

Deborah Schiffrin. 1987. *Discourse Markers*. Cambridge University Press.

John R. Searle. 1975. Indirect speech acts. In P. Cole and J. L. Morgan, editors, *Syntax and Semantics III: Speech Acts*, pages 59–82. Academic Press, New York.

R. N. Shepard. 1967. Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior*, 6:156–163.

P. Sibun. 1991. *The Local Organisation and Incremental Generation of Text*. Ph.D. thesis, University of Massachusetts.

Candace Sidner and David Israel. 1981. Recognizing intended meaning and speakers plans. In *Proc. International Joint Conference on Artificial Intelligence*, pages 203–208, Vancouver, BC, Canada.

Candace L. Sidner. 1979. Toward a computational theory of definite anaphora comprehension in English. Technical Report AI-TR-537, MIT.

Candace Sidner. 1985. Plan parsing for intended response recognition in discourse. *Computational Intelligence*, 1.

Candace Sidner. 1992. Using discourse to negotiate in collaborative activity: An artificial language. *AAAI Workshop on Cooperation among Heterogeneous Agents*.

Sidney Siegel. 1956. *Nonparametric Statistics for the Behavioral Sciences*. McGraw Hill.

Miriam Solomon. 1992. Scientific rationality and human reasoning. *Philosophy of Science*, September.

Karen Sparck-Jones. 1964. *Synonymy and Semantic Classification.* Ph.D. thesis, Cambridge University. Published in 1986 by Edinburgh University Press.

Dan Sperber and Deidre Wilson. 1981. Irony and the use-mention distinction. In Peter Cole, editor, *Radical Pragmatics*, pages 295–318. Academic Press, New York.

Dan Sperber and Deidre Wilson. 1986. *Relevance, Communication and Cognition.* Basil Blackwell, Oxford.

Robert C. Stalnaker. 1978. Assertion. In Peter Cole, editor, *Syntax and Semantics, Volume 9: Pragmatics*, pages 315–332. Academic Press.

Saul Sternberg. 1967. High-speed scanning in human memory. *Science*, 153:652–654.

Lucy Suchman. 1985. *Plans and Situated Actions.* Cambridge University Press.

John Tang and Larry Leifer. 1988. A framework for understanding the workspace activity of design teams. In *Proceedings of the Conference on CSCW*, pages 244–249.

Deborah Tannen. 1989. *Talking voices : repetition, dialogue, and imagery in conversational discourse.* CUP.

J. M. B. Terken and S. G. Nooteboom. 1987. Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes*, pages 145–163.

Rich Thomason. 1990. Accommodation, meaning and implicature: Interdisciplinary foundations for pragmatics. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press.*

E. Tulving. 1967. The effects of presentation and recall of material in free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 6.

Amos Tversky and Daniel Kahneman. 1982. Causal schemas in judgements under uncertainty. In *Judgment under uncertainty : heuristics and biases.* Cambridge University Press.

Moshe Y. Vardi. 1989. On the complexity of epistemic reasoning. In *Proc. 4th IEEE Symp. on Logic in Computer Science*, pages 243–252.

Marilyn A. Walker and Steve Whittaker. 1990. Mixed initiative in dialogue: An investigation into discourse segmentation. In *Proc. 28th Annual Meeting of the ACL*, pages 70–79.

Marilyn A. Walker, Masayo Iida, and Sharon Cote. 1993. Japanese discourse and the process of centering. *To Appear in Computational Linguistics, 20-2.*

Marilyn A. Walker. 1992a. A model of redundant information in dialogue: the role of resource bounds. Technical Report IRCS-92-25, University of Pennsylvania Institute for Research in Cognitive Science. Dissertation Proposal.

Marilyn A. Walker. 1992b. Redundancy in collaborative dialogue. In *Fourteenth International Conference on Computational Linguistics.*

Marilyn Walker. 1993a. Discourse strategies for attention-limited agents. Technical Report IRCS-93-21, University of Pennsylvania, Institute for Research in Cognitive Science.

Marilyn A. Walker. 1993b. Redundant affirmation, deliberation and discourse. In *Penn Review of Linguistics*, volume 17.

Marilyn A. Walker. 1993c. When given information is accented: Repetition, paraphrase and inference in dialogue. In *Presented at the LSA Annual Meeting.* Also in the Proceedings of the Institute for Cognitive Science Workshop on Prosody in Natural Speech, August 1992.

Gregory Ward and Julia Hirschberg. 1991. A pragmatic analysis of tautological utterances. *Journal of Pragmatics*, pages 338–351.

Gregory Ward. 1985. *The syntax and semantics of preposing.* Ph.D. thesis, University of Pennsylvania.

Gregory L. Ward. 1990. The discourse functions of vp preposing. *Language*, 66(4):742–763.

Bonnie Lynn Webber and Aravind Joshi. 1982. Taking the initiative in natural language database interaction: Justifying why. Technical Report 1, Department of Computer and Information Science, University of Pennsylvania.

Bonnie Lynn Webber. 1978. *A Formal Approach to Discourse Anaphora.* Ph.D. thesis, Harvard University. Garland Press.

Bonnie Lynn Webber. 1986. Two steps closer to event reference. Technical Report MS-CIS-86-74, Linc Lab 42, Department of Computer and Information Science, University of Pennsylvania.

Bonnie Webber. 1988. Discourse deixis and discourse processing. Technical Report MS-CIS-88-75, Department of Computer and Information Science, University of Pennsylvania.

Steve Whittaker and Phil Stenton. 1988. Cues and control in expert client dialogues. In *Proc. 26th Annual Meeting of the ACL, Association of Computational Linguistics*, pages 123–130.

Steve Whittaker, Erik Geelhoed, and Elizabeth Robinson. 1993. Shared workspaces: How do they work and when are they useful? *To Appear in IJMMS.*

S. Wilks. 1962. *Mathematical Statistics*. Wiley.