The Structure of Multiple-Headed Negotiations

by

Carolyn Penstein-Rosé

June 1995

Report CMU-PHIL-65

Philosophy
Methodology
Logic

Pittsburgh, Pennsylvania 15213-3890

The Structure of Multiple Headed Negotiations

Carolyn Penstein Rosé

Computational Linguistics Program

Carnegie Mellon University

cprose@lcl.cmu.edu

Abstract

In this paper I demonstrate that theories which postulate a strict tree structure of discourse on either the intentional or attentional level are not totally adequate for handling spontaneous dialogues. I use as evidence analyses of the structure of multiple headed negotiations. These are negotiation dialogues in which multiple suggestions are being negotiated at once. The question I explore in this paper is how to characterize the structure of a discourse in which multiple non-embedded adjacency pairs are initiated within a single turn since it is unclear in these cases how response parts are matched up with their corresponding initiating parts if discourse is tree structured. I argue that each non-embedded initiating part exists on a separate thread of the discourse so that an initiating part can be matched with its corresponding response part on its own thread. I develop my theory for the structure of these dialogues in the spirit of (GS86) and (Loc94). I support my claim that discourse is structured as a tapestry of interwoven threads by demonstrating that discourse level parallel structure makes it necessary to model attentional state as a graph structured stack rather than as a simple stack in order to make correct predictions about which discourse entities are readily available for pronominal reference, particularly with deictic pronouns like "that" and "those". I demonstrate that the graph structured stack model is a generalization of the simple stack model and therefore covers all of the phenomena covered by the simple stack model in addition to the data discussed in this paper which is problematic for the simple stack model. An implementation of this theory is discussed in (RELED95).

1. INTRODUCTION

Conversation Analysts have traditionally characterized the structure of spoken discourse, in particular the relationship between speech acts, in terms of adjacency pairs (SS73; SSJ74; Gof76; Cou77) where first pair parts are matched with an expected second pair part. For example, questions are paired with their answers, and greetings are often paired with similar greetings, and suggestions are paired with acceptances or rejections. In this paper, I analyze the structure of multiple headed negotiations. These are negotiation dialogues in which multiple suggestions are being negotiated at once. The question I explore in this paper is how to characterize the structure of a discourse in which multiple non-embedded first pair parts are presented within a single turn[1] since it is unclear in that case how second pair parts are matched up with their corresponding first pair parts if discourse is tree structured. I will demonstrate that theories which postulate a strict tree structure of discourse on either the intentional or attentional level are not totally adequate for handling spontaneous dialogues.

The corpus which my model is based on is composed of 50 recorded and transcribed dialogues in which speakers attempt to schedule a meeting together. Although both speakers share the intention of scheduling a meeting together, the speakers differ with respect to which days and which times are preferable for them to meet according to their respective schedules. So in these dialogues, speakers discuss their potential intentions to meet at particular times on particular days. It is these potential intentions which they do not necessarily share.

The question is how the structure of the turn in which multiple topics are presented affects the possibilities for the structure of subsequent turns which address the same topics. In particular, I will explore how this structure constrains possibilities for discourse pronominal reference. The role which discourse structure plays in constraining resolution of pronominal reference has been explored in (Pol88; Fox87; Ash93; GS85). I will develop my theory of discourse structure in the spirit of (GS86; Loc94) which has played an influential role in the analysis of discourse entity saliency and in the development of dialogue processing systems.

The examples below, all of which were extracted from my corpus with the exception of (IV) and (V) which were constructed for the purpose of illustrating the difference between presenting multiple suggestions

[1] A turn is defined as the largest span of continuous talk by a single speaker.

in a single clause verses presenting them each in separate clauses, demonstrate how speakers present multiple suggestions in parallel within a single turn. Each parallel suggestion expresses a potential intention to meet at a particular time.

In example (I), s1 suggests two different Tuesdays in parallel and s2 rejects both of them. In order for the reference to "those" in unit (4) to be resolved, both suggested meeting times must be available for pronominal reference.

(I)

s1: (1) What about Tuesday afternoons?

(2) The twenty third I'm free from three to five.

(3) And on the thirtieth I'm free after noon.

s2: (4) Okay, both of those are bad for me.

In example (II), s1 again suggests two meeting times, the fourth and the sixth. In order for the reference to "that" to be resolved in unit (3), the meeting time on the sixth must be relatively more salient than the meeting time on the fourth. On the other hand, we know that both meeting times must be salient since the reference to "those" in (6) can be resolved. The reference to "that" in unit (5) is ambiguous. It could potentially refer either to the meeting on Friday or both the meeting times.

(II)

s1: (1) Okay, so the fourth is a potential.

(2) Or how about Friday the sixth in the morning?

(3) Do you have any time free that day?

s2: (4) No.

(5) that won't work.

(6) Both of those days are bad.

In example (III), s1 suggests that next week is a good time to meet and then proceeds to make two

parallel suggestions which make this suggestion more specific. Then s2 rejects the first of these suggestions and accepts the other.

(III)

s1: (1) Next week is pretty good for me.

(2) I think next Wednesday afternoon would be best,

(3) or Thursday after two.

s2: (4) Wednesday doesn't seem to be too good.

(5) Thursday is.

(6) Anywhere from two thirty to five on Thursday is fine.

In example (IV), s1 makes three parallel suggestions, each in different units. In unit 4, in order to resolve the reference to "that", the suggested meeting time on Wednesday must be relatively more salient than the other two.

(IV)

s1: (1) I am free Monday.

(2) Tuesday is good as well.

(3) I could also make it Wednesday.

s2: (4) That day won't work

(5) because I have class on Wednesday.

(6) But either Monday or Tuesday would work.

In contrast to example (IV), the unnatural example (V) indicates that when the suggested times are presented in a list in the same unit, it isn't the case as it is when they are presented in separate units that the last suggested time is relatively more salient than the other two.

(V) ??

s1: (1) I am free Monday, Tuesday, and Wednesday.

s2: (2) That day won't work

(3) because I have class on Wednesday.

(4) But either Monday or Tuesday would work.

In the following sections, I will use this data to argue that a simple stack model of attentional state, where it accounts for many interesting discourse phenomena, is not adequate for handling the case where there is discourse level parallelism as in these examples. I will argue that a graph structured stack model makes more correct predictions, handling all of the examples above.

1.1. The Grosz and Sidner Account

Grosz and Sidner's model is composed of three separate but interacting components: the linguistic structure, the intentional structure, and the attentional state. The linguistic structure is the structure of the sequence of sentences which make up the discourse. It is on this level that the discourse is divided into segments. Each discourse segment is associated with a discourse segment purpose. The relationships between these discourse segment purposes make up the intentional level. There are two relationships discussed in their 1986 paper, namely *dominates* and *satisfaction-precedes*. A discourse segment purpose *dominates* a second one when satisfaction of the second one contributes to the satisfaction of the first one. A discourse segment purpose *satisfaction-precedes* another discourse segment purpose if the first purpose must be satisfied before the second one can be attempted to be satisfied. The attentional state is modeled with a stack where each discourse segment is associated with a focus space[2] which is pushed on the stack when that discourse segment is processed. The attentional stack determines which entities can be referred to pronominally, generally those entities on or near the top of the focus stack.

For each discourse segment, a focus space is pushed onto a focus stack with all of the entities which are evoked in that segment. The dominance relations in the intentional structure determine which operations are performed on the attentional stack when a discourse segment is processed. If discourse segment purpose

[2]The focus space contains all of the referents evoked within the corresponding discourse segment.

A is dominated by discourse segment purpose B then focus space A can be pushed on top of focus space B if focus space B is on top of the stack when focus space A is pushed. Otherwise, if discourse segment purpose A is not dominated by discourse segment purpose B, focus space B must be popped from the stack when focus space A is pushed. When a focus space is pushed onto the attentional stack, it is pushed on top of the focus space associated with the discourse segment purpose its corresponding discourse segment purpose contributes to. All of the focus spaces in between the new one and the one which it will be pushed on top of are popped off prior to pushing the new focus space. Pronominal references within a discourse segment are computed after the segment's focus space has been pushed on to the stack.

In a recent extension to Grosz and Sidner's original theory, described in (Loc94), each discourse segment purpose represents an intention that the speakers form a shared plan (GK93). These discourse segment purposes are expressed in terms of the intention operator *Int.That*, one of four intention operators discussed in (GK93) based on the theory of intentional behavior and practical reasoning discussed in (BIP88). *Int.That* represents the intention that some proposition hold. Another intention operator discussed in (GK93) is *Pot.Int.That*. This represents an agent's potential intention that some proposition hold. Potential intentions are used to account for an agent's process of weighing different means for accomplishing an action he is committed to performing (BIP88). These potential intentions are not discourse segment purposes in Lochbaum's theory since they cannot form the basis for a share plan having not been decided upon yet. It is not until they have been decided upon that they become *Int.That*'s which can then become discourse segment purposes.

For example, deliberation over how to accomplish a shared plan can be represented as an expression of multiple *Pot.Int.That*'s, each corresponding to different alternatives. For each factor distinguishing these alternatives, the potential intentions are all discussed inside of a single discourse segment whose purpose is to explore the options so that the decision can be made. Lochbaum discusses an example of a dialogue containing examples of such deliberation in (Loc94), included here in Figure 1. In each of the three subdialogues, a different type of information is discussed. In the first one, two alternative ways of opening a savings account are introduced. In the second one, the interest rates of each are compared. In the third, required initial deposits of each are discussed. Both alternatives are discussed in each subdialogue.

(1) Customer: I'd like to open a savings account.

(2) What types do you offer?
(3) Teller: Passbook and investment.

(4) Customer: What's the interest rate on your passbook account?
(5) Teller: 2.5 %.
(6) Customer: And the rate for the investment account?
(7) Teller: 3.0 %
(8) Customer: Okay.

(9) How big are the initial deposits for the two accounts?
(10) Teller: $1000 for the passbook and $5000 for the investment.
(11) Customer: Okay.

(12) Whom do I see to open a passbook account?

Figure 1: **Example of Deliberating Over Selecting a Plan for Action**

The stipulation that *Int.That*'s can be discourse segment purposes but *Pot.Int.That*'s cannot has a major impact on the analysis of scheduling dialogues such as the one in Figure 2 since the majority of the exchanges in scheduling dialogues are devoted to deliberating over which date and at which time to schedule a meeting. In our corpus, the only factor which distinguishes good meeting times from bad ones is whether the speakers are busy doing something else at those times or not. This would seem to leave all of the deliberation over meeting times within a single monolithic discourse segment. This leaves the vast majority of the dialogue with no segmentation. This would leave us with the question of how to account for shifts in focus which seem to occur within the deliberation segment as evidenced by the types of pronominal references which occur within it. For example, in the dialogue presented in Figure 2, how would it be possible to account for the differences in interpretation of "Monday" and "Tuesday" in (3) with "Monday" and "Tuesday" in (14)? It cannot simply be a matter of immediate focus since the week is never mentioned in (13). And there are no semantic clues in the sentences themselves to let the hearer know which week is intended. Either there is some sort of structure in this segment more fine grained than would be obtained if *Pot.Int.That*'s cannot be discourse segment purposes, or another mechanism must be proposed to account for the shift in focus which occurs within the single segment. I argue that rather than propose an additional mechanism, it is more perspicuous to lift the restriction that *Pot.Int.That*'s cannot be discourse segment purposes. In my theory a separate discourse segment is allocated for every potential plan discussed in the dialogue.

(1) S1: We need to set up a schedule for the meeting.

(2) How does your schedule look for next week?
(3) S2: Well, Monday and Tuesday both mornings are good.
(4) Wednesday afternoon is good also.
(5) S1: It looks like it will have to be Thursday then.
(6) Or Friday would also possibly work.
(7) Do you have time between twelve and two on Thursday?
(8) Or do you think sometime Friday afternoon you could meet?
(9) S2: No.
(10) Thursday I have a class.
(11) And Friday is really tight for me.
(12) How is the next week?
(13) If all else fails there is always video conferencing.
(14) S1: Monday, Tuesday, and Wednesday I am out of town.
(15) But Thursday and Friday are both good.
(16) How about Thursday at twelve?
(17) S2: Sounds good.

(18) See you then.

Figure 2: **Example of Deliberating Over A Meeting Time**

It could potentially be argued that it does not make sense to model discourse segment purposes as potential intentions because a speaker would not say something unless that speaker was committed to it somehow and that it is necessary to use intentions instead. However, I argue that this is not a problem with our representation and that it in fact makes more sense to model the discourse segment purposes in question as potential intentions.

It is clear that in making a suggestion, a speaker is at least committed to making that suggestion. However, in Lochbaum's theory, what the speaker is committed to in engaging the other speaker in a discourse segment is something different. She writes that "The purpose of the segment is taken to be the intention that the discourse participants form that plan." (Loc94) (p. 49). Because according to (BIP88) a rational agent will not be committed to two things which are inconsistent, a rational agent will not intend to form two plans which are mutually exclusive. Since a single meeting has only one meeting time, if a speaker intends to form a plan to meet at a particular time, it does not make sense that that speaker would also

intend to form a plan to meet at a different time unless these two intentions corresponded to two separate meetings. It would make sense, however, for an agent to potentially intend to form two different alternative plans for a meeting.

It might be argued that the process of negotiating over alternative meeting times is somehow analogous to the types of knowledge precondition dialogues discussed in (Loc94). This process is most similar to her subdialogues in which the purpose is to achieve a satisfactory description of a plan. This satisfactory description would contain in its completion a time and date on which to meet in the case of scheduling dialogues. It cannot be the case that the speaker intends that the time and date in question be suggested since simply suggesting a meeting time and date does not directly contribute to building a plan for a meeting. So the speaker must have some sort of intention about the outcome of suggesting a particular time and date. However, it cannot be the case that for each potential meeting time and date the speaker intends that the corresponding time and date be acceptable instantiations in order to achieve a satisfactory description of a plan for a meeting since this would not allow the speaker to engage in subdialogues about mutually exclusive alternatives in parallel since that would make it necessary for the speaker to have inconsistent intentions. One might suggest as the only alternative left that the purpose is to come to some agreement on the status of the time and date as a possible acceptable description for the meeting plan. But intending to establish some unspecified status of a suggestion has the same problems as intending to make a suggestion.

Formulating each parallel negotiation segment as having the purpose of a potential intention that the discourse participants form a plan to meet on the specified time and date does not have these problems. Since it is acceptable for a speaker to have mutually exclusive potential intentions, this avoids the problem of the speaker intending that the suggested meeting time and date be acceptable. And since the role of potential intentions in the deliberation process is already built in to Grosz and Kraus's SharedPlans framework, it would avoid the problem of the speaker simply intending to suggest a particular time and date. The question which remains is whether it makes sense for a potential intention being sufficient motivation for a speaker to actually say something. I argue that it is.

In (GK93), it is theorized that an agent weighs a number of potential intentions internally in the process of deliberating over multiple alternatives before making a commitment. In order to weight these alternatives,

an agent must have some basis by which to evaluate them. In some cases, an agent does not have enough information to make the decision. One such case is when the speaker has part of the information required in order to determine if a particular instantiation is appropriate and the other speaker has the other part of the information. In my theory of the negotiation process in these circumstances, the agent can invite the other agent to participate as an equal party in her internal deliberation process by making it explicit by verbalizing it and allowing the other agent to respond. This is the process which I am modeling when I say that discourse segment purposes within these types of negotiation segments are potential intentions. Thinking of the negotiation process this way makes using potential intentions as discourse segment purposes make sense since it gives the speaker a motivation for verbalizing something which she is not specifically committed to.

1.2. An Alternative Account

A number of different potential intentions are discussed in Figure 2. In (2), the potential intention to meet sometime next week is expressed. In (3), the more specific potential intention to meet either Monday or Tuesday in the morning is expressed. In (4), the potential intention to meet on Wednesday is expressed. Each time a speaker expresses an attitude towards a potential plan, a potential intention is expressed. Other potential intentions are expressed throughout this dialogue.

By allocating a separate discourse segment to each potential plan, a situation arises where discourse segments seem to overlap one another in some cases. For example, if the discussion about the potential meeting on Thursday is in a separate segment from the potential meeting on Friday, then either the two segments are interleaved such that (5), (7), and (10) are in one segment and (6), (8), and (11) are in another, or all six of these sentences are in separate segments altogether. Which of these options are selected would make definite predictions about which entities are available for pronominal reference at different points between sentence (5) and sentence (11).

A similar issue is addressed in (Web88) where the following example is discussed:

> House A is in Palo Alto, House B in Portola Valley. Both were built in 1950, and both have three bedrooms. House A has two baths, and B, 4. House B also has a kidney shaped pool. House

A is on a quarter acre, with a lovely garden, while house B is on 4 acres of steep wooded slope, with a view of the mountains. The owner of house A is asking $425K. The owner of house B is asking $600K. That's all I know about house A. This I heard from a friend, who saw house B before it came on the market. (p 115)

In this example, because the reference of "that" in the second to last sentence seems unnatural to resolve to the description of house A and the interpretation of "this" as referring to the description of house B seems very strange, Webber concludes that in discourses where there is discourse-level parallel structure as in the example above, the individual parallel components cease to be available for pronominal reference[3]. I argue that it is not the parallel structure which makes these references difficult to resolve but rather a combination of two factors. The first factor is that because the description of house B was most recently modified, it is more in focus than the description of house A at the end of the two descriptions. So if a pronominal reference is made to one or the other description directly after the description of B is modified, it will most naturally refer to the description of B. Secondly, when "this" is used along side of "that" it is communicated that what is referred to with "this" is closer to the speaker than what is referred to with "that". So it would seem very odd to use "this" to shift the focus. Because the focus was shifted to the description of house A in the previous sentence, it seems odd to use "this" to shift the focus back to the description of house B. If this is the reason why those two references are unnatural, I would predict that reversing the last two sentences would make the two references seem more natural and would indicate that it is not the case that the individual descriptions are not available for pronominal reference after the two descriptions are given in parallel. And this turns out to be the case.

House A is in Palo Alto, House B in Portola Valley. Both were built in 1950, and both have three bedrooms. House A has two baths, and B, 4. House B also has a kidney shaped pool. House A is on a quarter acre, with a lovely garden, while house B is on 4 acres of steep wooded slope, with a view of the mountains. The owner of house A is asking $425K. The owner of house B is

[3] Some readers have indicated that "this" seems to refer to the whole description including both house A and house B. This is an alternative reading. The point I am addressing here, however, is simply the ability of "this" in this context to refer to the description of house B.

asking $600K. This I heard from a friend, who saw house B before it came on the market. And that's all I know about house A.

This indicates that each parallel description is maintained throughout the description, supporting the view that it is two discourse segments interleaved[4] rather than several fragmentary discourse segments in a sequence. I will argue based on a corpus analysis of 50 dialogues that the model in which the discourse segments are interleaved makes more correct predictions about possibilities for discourse pronominal reference than the model where a totally separate discourse segment is created for each non-contiguous portion of the discourse.

This issue raises interesting questions about the representation of attentional state. Discourse segment purposes neither dominate nor are dominated by parallel discourse segment purposes by definition. Therefore, according to Grosz and Sidner's theory, focus spaces associated with one parallel thread must all be popped from the top of the stack when focus spaces associated with another parallel thread are pushed on. This predicts that pronominal references can never be made which include entities from focus spaces associated with more than one parallel thread at the same time. I will demonstrate that this prediction is false.

[4] i.e. I am saying that the description of house A is in a separate segment from the description of house B.

2. METHODOLOGY

In this section I will discuss my methodology for analyzing the dialogues used as data for the theory developed in this paper. I will explain how I derive the representation for the attentional state from this representation. In the next section I will use this coding scheme to demonstrate that incorrect predictions are made if attentional state is modeled as a simple stack structure and that correct predictions are made if it is modeled as a graph structured stack.

I first introduced this methodology in (Ros95). The coding scheme presented here is an extension of the coding scheme presented in that paper. My goal for this coding scheme is that it should be reliable in the sense that the analyses which I produce with this coding scheme should be reproducible. My coding scheme is based on (Edm81) in which Edmondson develops a formalism for analyzing adversarial conversations in scenarios such as "Speaker1 wants Speaker2 to do X" or "Speaker1 did X and X is bad for Speaker2". What distinguishes Edmondson's formalism from those of others who have studied sequences of speech acts and how speech acts relate to one another in discourse is that he has separated the notion of illocutions from that of interactional structure which others have left conflated (Sea92; Kla73; LF77; SS73; SC75). In his formulation, the illocution is an action which communicates the beliefs, desires, and intentions communicated by an utterance without reference to what function this action plays in the discourse. The interactional component represents the relationship between this action and the surrounding discourse, how this sentence works together with surrounding sentences in order to produce conversational outcomes, i.e. what is accomplished through a unit of discourse.

What I am coding with my coding scheme is similar to the interactional structure in Edmondson's work. From this structure, I derive my equivalent of Grosz and Sidner's intentional structure. My analyses with this coding scheme identify moves in dialogues and their relationships with one another. With this coding scheme I have carved out a finite set of discrete relationships which I identify in the data I analyze. I have not made any assumptions about how these relationships can be ordered or how many relationships a single sentence[5] can participate in.

[5]The smallest unit of analysis in my coding scheme is actually a unit called a pseudo sentence unit which is generally a clause. More details about pseudo-sentence segmentation are given below.

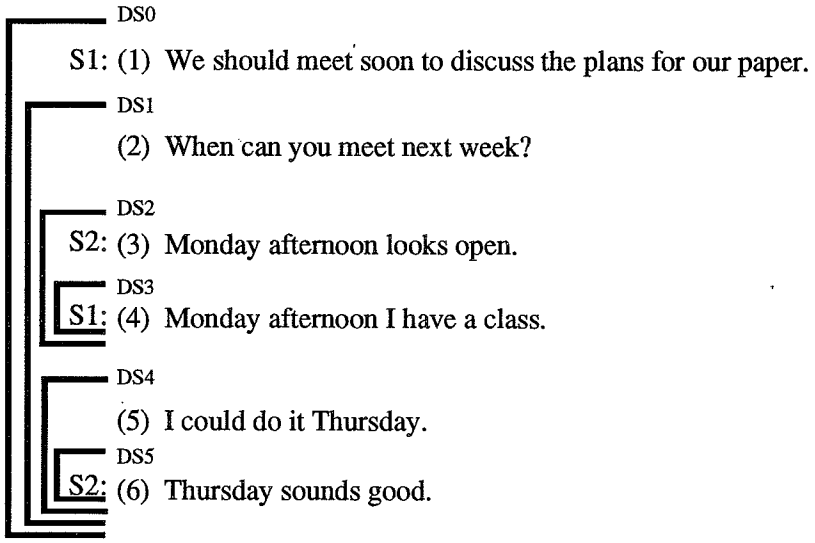


Figure 3: **Example Scheduling Dialogue**

2.1. The Coding Scheme

I will explain my coding scheme by demonstrating how to apply it to the example dialogue in Figure 3. In this figure I have noted how I would segment this discourse considering that potential intentions can act as discourse segment purposes. At the end of this section I will describe how to derive this segmentation from the coding scheme.

The discourse segment purpose of DS0 is a shared intention for the two speakers to settle on a date on which to hold their meeting. The discourse segment purpose of DS1 is S1's potential intention to make that meeting during the next week. The period of time denoted by 'next week' is only a fraction of the period of time in which the meeting could be held according to sentence (1). So this potential intention is dominated by the more general intention of finding a time for scheduling a meeting. Likewise, the potential intention to meet on Monday is more specific than meeting next week, so this requires another subordinate discourse segment. When S1 responds to S2's suggestion, a new subordinate discourse segment must again be allocated because an intention to respond to a potential intention contributes to determining the status of that potential intention (i.e. whether it is shared by the other speaker or not). S1 then makes a suggestion to meet on Thursday. This potential intention is not subordinate to the potential intention of meeting on Monday. It is dominated by the potential intention to meet next week. And because it is more specific to

meet on Thursday than to meet next week, a new subordinate discourse segment is allocated inside of DS1. Finally, sentence (6) is placed in DS5 which is subordinate to DS4 since sentences (5) and (6) were expressed by different speakers. This is simply an intuitive justification for this segmentation. After the coding scheme is introduced, concrete rules for deriving the segmentation from the coding scheme will be described.

My coding scheme was inspired by Edmondson's (81) dissertation research discussed above. In Edmondson's coding scheme, utterances were related to one another through conversational moves. Through these moves, the utterances worked together to produce conversational outcomes. These conversational outcomes were said to be produced through the interactional structure, but they were never formalized, nor was the process of deriving them from the interactional structure. In my coding scheme, I have defined what these outcomes are and how they are derived in order to make it possible to determine whether any given analysis for a dialogue is correct for a given interpretation. Also, rather than assign each sentence a label indicating how it fits into the interactional structure as Edmondson did, I assign these labels to relationships between conversational outcomes. There are two types of conversational outcomes, local outcomes and non-local outcomes. Local outcomes are produced by single pseudo-sentence units. And non-local outcomes are produced as these local outcomes interact with previous outcomes. The relationships between outcomes are conversational moves. The set of relationships between outcomes makes up the interactional structure. These relationships make it possible for a sentence to play multiple roles in the interactional structure.

OUTCOME: 3-b
SPEAKER: speaker1
START TIME: next week, Monday afternoon
DURATION: two hours
TOPICS: plan for the paper
ATTITUDE: positive

Figure 4: **Sample Outcome**

Because scheduling dialogues such as the one in Figure 3 are what I am applying my coding scheme to, I consider conversational outcomes to be partial or full plans for meetings which are either positive, indicating that the speaker or speakers are in favor of this plan, or negative, indicating that the speaker or speakers

are not in favor of this plan. See figure 4 for a sample outcome expressing a positive attitude towards a meeting next week on Monday afternoon for two hours to discuss plans for the paper. At any given point in the dialogue, the coder can check the validity of the analysis by making sure the outcomes which are open correspond to the smallest full set of mutually exclusive full or partial plans open for negotiation at that point in the dialogue. Outcomes are represented in a structure with slots for the label of the outcome, the speaker, the start time of the meeting, the duration of the meeting, the location of the meeting, the topics to be covered at the meeting, activities to take place at the meeting, conditions placed on the plan for the specified meeting (e.g. "that depends upon how I'm feeling that day"), and the speaker's expressed attitude towards the plan.

It should be noted that this coding scheme is not limited to analyzing scheduling dialogues. It could be used to analyze negotiations over other sorts of alternatives like alternative recipes[6] for accomplishing a goal or alternative variable bindings once a recipe has been selected. It is only a matter of making the conversational outcomes represent what is being negotiated over. It can also be applied to examples where one or two speakers are comparing alternatives even when it is not a negotiation. For example, look again at Webber's example from (Web88) where house A and house B are being compared.

> House A is in Palo Alto, House B in Portola Valley. Both were built in 1950, and both have three bedrooms. House A has two baths, and B, 4. House B also has a kidney shaped pool. House A is on a quarter acre, with a lovely garden, while house B is on 4 acres of steep wooded slope, with a view of the mountains. The owner of house A is asking $425K. The owner of house B is asking $600K. That's all I know about house A. This I heard from a friend, who saw house B before it came on the market. (p. 115)

In this case the outcomes would be specifications for houses with slots for features such as city, date built, number of bedrooms, etc. Each further specification of one or both houses would be a conversational move, causing a local outcome to interact with a previous outcome, producing an updated description of one or both houses. Although the coding scheme can be applied to different types of negotiations and some

[6] A recipe is a generic plan for accomplishing a goal which can be instantiated for a particular instance of a problem.

other types of dialogues, it is not clear how it could be applied to dialogues which are not negotiations in general since it is made up of moves which one would expect in a negotiation but not necessarily in discourse in general. Therefore, it would not be adequate for analyzing all types of non-negotiation portions of dialogues. How to apply the model of attentional state proposed in this paper to non-negotiation dialogues with parallel structure is a direction for future research. Without being able to apply the coding scheme in a straightforward way to these types of discourse, it is unspecified how to derived the segmentation.

The input to the coding process is a transcription of a recorded scheduling dialogue with the following features. Each turn is labeled with the initials of the speaker and the number of the turn. The transcription of each turn includes words, human noises, non-human noises, pauses and false starts. Punctuation based on intonation, not based on any grammatical rules, is marked as well.

The procedure followed by coders contains four steps. The first step is to segment the dialogue into pseudo-sentence units, described below, and number the units. Then for each unit, coders mark which of the possible local outcomes, if any, is produced from this unit in isolation. Next they label this unit with the conversation act which this unit expresses, described below. Finally, they derive the interactional structure produced by this unit in relation to its context. These last three steps are repeated for each unit.

The unit of analysis in this coding scheme is called a pseudo-sentence. Because I am working with spoken discourse, it is difficult to segment the discourse into units by purely syntactic means. Although pseudo-sentences are not always syntactic sentences strictly speaking, pseudo-sentence breaks roughly correspond to clause boundaries. The main exception is conditional clauses. For example, although "I can meet you Wednesday if you're out of your class before 2:00." is two clauses, it is only one pseudo-sentence unit since neither clause is true in isolation. False starts are included in the pseudo-sentence unit which they precede. Space limitations prevent discussing this in more detail here.

Each pseudo-sentence can potentially be labeled with a conversation act[7], be assigned relations in the interactional structure, and produce conversational outcomes. Each pseudo-sentence with propositional content expressing an attitude towards some partial plan or aspect of some partial plan for scheduling a meeting produces a local conversational outcome. For example, the pseudo-sentence "The twenty eighth is

[7]I am calling illocutions in Edmondson's sense conversation acts.

bad for me" produces a negative local conversational outcome about the partial plan for meeting on the twenty eighth. Pseudo-sentence units which produce local conversational outcomes also express conversation acts. It is the conversation acts which produce the local outcomes. The conversation act which produces a negative local outcome is called *prohibit* whereas the one which produces a positive one is called *license*. Utterances such as "No" which have no propositional content produce no local outcome. They are purely interactional and produce an outcome only through the interactional structure.

Input/Output

Specify/Specified
Restrict/Restricted
Satisfy/Satisfied
Reject

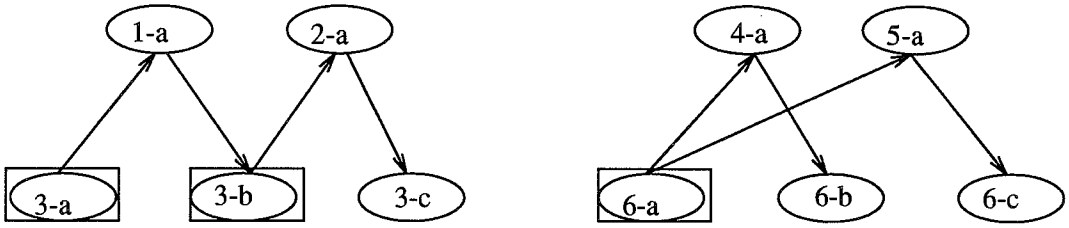
Figure 5: **Interactional Relations**

A local outcome may have implications regarding the status of one or more previously produced conversational outcome(s), giving rise to the interactional structure. The full set of moves I have included in my coding scheme are included in Figure 5 and explained below. I will be making reference to the example in Figure 3.

Some Preliminary Explanations and Definitions Before explaining each of the conversational moves represented in my coding scheme and presenting explicit examples, I would like to introduce some preliminary definitions and explain the graphical representation of the interactional structure. This is meant to give the reader a taste of what to expect in the coming sections so the presentation of the coding scheme will be easier to follow. In Figure 6 there are two example interactional structures. The ovals represent outcomes. Each oval is labeled with a number indicating the pseudo-sentence which the outcome resulted from and a letter which indicates the order in which the outcomes were produced. Outcomes labeled with letter 'a' are local outcomes resulting directly from the propositional content of the corresponding pseudo-sentence unit. For example, outcome 3-a is the local conversational outcome produced by pseudo-sentence unit 3. All the other outcomes result from the interaction of an outcome with previous outcomes through conversational moves. Boxed outcomes are ones which are no longer active. Inactive outcomes are not available for interaction with

S1: (1) I could meet you on Tuesday.
(2) Wednesday is also good for me.
S2: (3) Both of those days are possibilities in the morning.

S1: (4) Can you meet on Thursday?
(5) Or what about Friday?
(6) Mornings are generally preferable for me.



2-a dominates 3-c
1-a dominates 3-b
1-a dominates 3-c

4-a dominates 6-b
5-a dominates 6-c

Figure 6: **Dominance Example**

subsequent outcomes. Arrows are normally labeled with the move through which an outcome interacts with a previous outcome.

Each conversational move involves three outcomes. A current outcome interacts with a previous outcome producing a new outcome. For example, outcome 3-a interacts with outcome 1-a producing outcome 3-b. The previous outcome *dominates* the new outcome. So in this example, outcome 1-a dominates outcome 3-b. And 2-a dominates 3-c. I consider that outcomes which dominate the current outcome also dominate the new outcome. So outcome 1-a also dominates outcome 3-c. Also, dominates is transitive so that if outcome a dominates outcome b and outcome b dominates outcome c then outcome a dominates outcome c. These dominates relations will become important when the correspondence between interactional structure and discourse segmentation is discussed.

Each outcome is a *description* which can *apply* to some set of objects of negotiation - in this case potential meetings. For example, if the start time for the outcome is "Next week, Tuesday" and no other slots are

filled, then this description can apply to any meeting which takes place next Tuesday. In general, an outcome applies to all objects of negotiation which can be described with all of the positive and negative features in the outcome. And it makes *available* all objects of negotiation open for negotiation which it applies to and does not eliminate with negative restrictions. For example, if the start time was "Next week, Tuesday, not at 3:00", then it applies to the same set of meetings that one which had as a start time simply "Next week, Tuesday", except that it *eliminates* all of those meetings which start at 3:00 on that day. Two outcomes are *mutually exclusive* if the sets of objects of negotiation which they apply to are mutually exclusive. And two outcomes *overlap* if the sets of objects of negotiation which they apply to intersect. An outcome *further specifies* another outcome if they overlap and both are active (i.e. still available for interaction). An outcome is more *specific* than another outcome if both are active and the set of objects of negotiation which it applies to is a proper subset of the set of objects of negotiation which the other outcome applies to.

When a current local outcome must interact with more than one previous conversational outcome, it can do so in two different ways. These two different cases are illustrated in the two sample dialogues above Figure 6. The reader should not be concerned at this point if it is not obvious how the analysis was derived from these examples. That will become clear as the coding scheme is described in the following section. The two interactional structures in Figure 6 represent the two alternative approaches. In the first example (on the left), outcome 3-a interacts with outcome 1-a producing outcome 3-b and then 3-b interacts with 2-a producing 3-c. This has the affect of tying 1-a and 2-a together since they both dominate outcome 3-c. The other approach has the affect of keeping the two previous outcomes separate since outcome 4-a dominates 6-b but not 6-c, and likewise outcome 5-a dominates 6-c but not 6-b. In the example on the right, outcome 6-a interacts with outcome 4-a producing outcome 6-b and then 6-a interacts with 5-a producing 6-c. The second strategy is the default strategy, but if this strategy will result in two outcomes which overlap, then the first strategy is taken instead. If the second strategy was used with the dialogue in (1)-(3), two outcomes would result which would each apply to meetings on Tuesday and Wednesday. Since these outcomes would overlap, the first strategy is used which ties the two outcomes together, avoiding the presence of outcomes which overlap.

The Specify Move A *Specify* is a move the purpose of which is to further specify a previous outcome. It accepts part of what has been suggested, but it adds a new restriction. The resulting outcome is identical to the previous outcome which is made more specific except for having been modified in the specified way. Two examples of this move are displayed in Figure 7. Note that the bold ovals represent the result of speaker2's utterances, and the other ovals represent ones which result from speaker1's utterances. In the first *Specify* move, the suggestion for meeting to discuss the paper is made more specific by suggesting to meet sometime next week. A further *Specify* move is performed when speaker2 makes this suggestion even more specific by suggesting a specific day next week in which to meet to discuss the paper.

S1: (1) We should meet soon to discuss the plans for the paper.
(2) When can you meet next week?
S2: (3) Monday afternoon looks open.

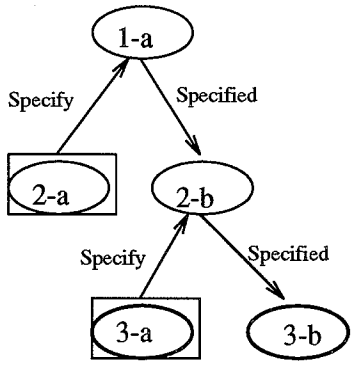


Figure 7: **The Specify Move**

Figure 8 contains specifications of all of the outcomes represented in Figure 7. Outcome 1-a was produced by the first pseudo-sentence unit in the example scheduling dialogue in Figure 3. This pseudo-sentence expresses a positive attitude towards the plan for a meeting with two features: first, the start time is "soon"; and second, the topic of discussion is "plans for the paper". There are no previous outcomes for outcome 1-a to interact with, so this pseudo-sentence has only one associated outcome.

Outcome 2-a results from the second pseudo-sentence which expresses a positive attitude towards a meeting with one feature, namely that the start time should be "next week". This outcome is compatible with outcome 1-a since the set of start times in "the next week" are a subset of the start times "soon". So a *Specify* move is performed by outcome 2-a interacting with outcome 1-a in order to produce outcome 2-b

OUTCOME: 1-a
SPEAKER: speaker1
START TIME: soon
TOPICS: plan for the paper
ATTITUDE: positive

OUTCOME: 2-a
SPEAKER: speaker1
START TIME: next week
ATTITUDE: positive

OUTCOME: 2-b
SPEAKER: speaker2
START TIME: soon, next week
TOPICS: plan for the paper
ATTITUDE: positive

OUTCOME: 3-a
SPEAKER: speaker2
START TIME: Monday afternoon
ATTITUDE: positive

OUTCOME: 3-b
SPEAKER: speaker2
START TIME: soon, next week,
monday afternoon
TOPICS: plan for the paper
ATTITUDE: positive

Figure 8: **Outcomes 1-a, 2-a, 2-b, 3-a, and 3-b**

which specifies a positive attitude towards a meeting whose start time is "soon, next week", and whose topic of conversation is "plans for the paper". Whenever an outcome W produced by pseudo-sentence X interacts with a previous outcome Y with a *Specify* move to produce outcome Z, outcome W is no longer available for future interaction with outcomes produced by pseudo-sentences subsequent to pseudo-sentence X. This is illustrated in Figure 9. The same applies when a *Satisfy* move takes place, discussed below. This is called **The New Specification Principle.**

Outcome 3-a results from the third pseudo-sentence unit. It expresses a positive attitude towards a meeting with start time on "Monday afternoon". Outcome 3-a further specifies both 1-a and 2-b. **The Feature Applicability Principle** determines which of the set of outcomes which are further specified by a current outcome will participate in a **Specify** move with it. It states that if the current outcome is more

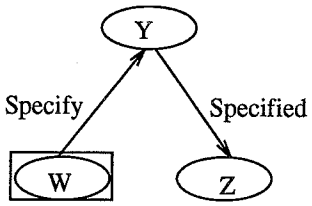


Figure 9: **The New Specification Principle**

specific than a previous outcome, then it will not interact with any previous outcomes dominating that outcome. So although outcome 3-a seems to be able to make a *Specify* move with both 1-a and 2-b, it only does so with 2-b since 1-a dominates 2-b and is less specific. So a *Specify* move is performed by outcome 3-a interacting with outcome 2-b producing outcome 3-b which specifies a positive attitude towards a meeting whose start time is "soon, next week, Monday afternoon" and whose topic of conversation is "the plans for the paper".

The Restrict Move A *Restrict* is a move which attempts to modify a previous outcome by removing some part or all of what was specified there. If the speaker produces a pseudo-sentence which expresses a *prohibit* and therefore attempts to retract part of an open partial plan, this would act as a *Restrict* in relation to that outcome. The resulting outcome is identical to the outcome which is restricted except for the the addition of a restriction which expresses the negation of the specified part. See Figure 10 and Figure 11. All outcomes which dominate the previous outcome will also get the addition of the restriction which expresses the negation of the specified part.

Pseudo-sentence unit 4 in Figure 10 expresses a negative attitude toward a meeting with one feature, namely that the start time is on "Monday afternoon". **The Feature Applicability Condition** comes into play in order to select outcome 3-b as the only outcome which 4-a interacts with. So outcome 4-a interacts with outcome 3-b producing outcome 4-b which specifies a description of a meeting with two features, namely the start time is "soon, next week, Monday afternoon, not Monday afternoon" and the topic of conversation is "the plan for the meeting". Note that because the start time specifies both Monday afternoon and not Monday afternoon it does not make any objects of negotiation available. Also, with *Restrict* moves in general, because the resulting outcome specifies the set of meetings specified by the description in the previous

S1: (1) We should meet soon to discuss the plans for the paper.
(2) When can you meet next week?
S2: (3) Monday afternoon looks open.
S1: (4) Monday afternoon I have a class.

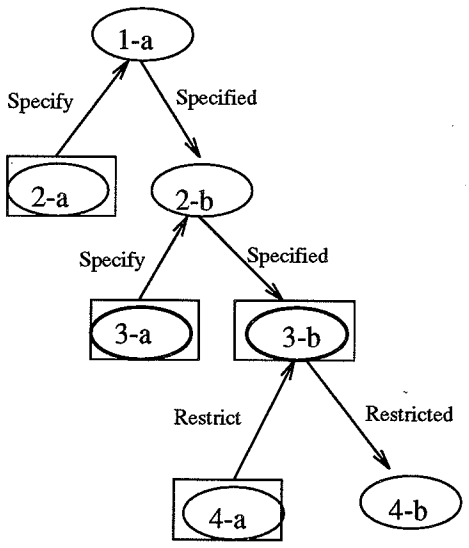


Figure 10: **The Restrict Move**

outcome which have not been eliminated by the current outcome, there is no reason for the previous outcome or the current outcome to be available for future interaction. This is called **The Restriction Principle** and it is illustrated in Figure 12. It states that whenever an outcome W produced by pseudo-sentence X interacts with a previous outcome Y with a *Restrict* move to produce outcome Z, outcomes W and Y are no longer available for future interaction with outcomes produced by pseudo-sentences subsequent to pseudo-sentence X.

In pseudo-sentence unit 5 in Figure 13, a new suggestion is made. Pseudo-sentence unit 5 produces outcome 5-a which is a description for a meeting with start time "Thursday". At this point, the available outcomes are 1-a, 2-b, and 4-b. Outcome 5-a is not a further specification of outcome 4-b since "Thursday" and "Monday afternoon" are totally disjoint periods of time. However, "Thursday" does further specify "next week". By **The Feature Applicability Principle**, outcome 2-b is selected over 1-a. So outcome 5-a interacts through a *Specify* move with outcome 2-b to produce outcome 5-b which is a description for a meeting with start time "soon, next week, Thursday" and topic of conversation "the plans for the paper".

OUTCOME: 4-a
SPEAKER: speaker1
START TIME: Monday afternoon
ATTITUDE: negative

OUTCOME: 4-b
SPEAKER: speaker1
START TIME: soon, next week, Monday afternoon, not Monday afternoon
TOPICS: plan for the paper
ATTITUDE: positive

Figure 11: **Outcomes 4-a and 4-b**

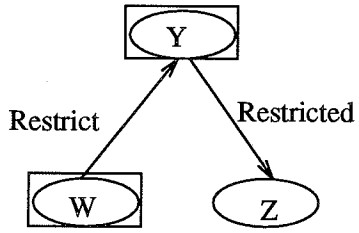


Figure 12: **The Restriction Principle**

By **The New Specification Principle**, outcome 5-a is no longer available for interaction with outcomes produced by pseudo-sentence units subsequent to 5. An additional effect that outcome 4-b is no longer available for future interaction also occurs. This is by **The Inactivity Priciple** which specifies that when an interaction with a previous outcome occurs, all available outcomes not dominating the current outcome which do not make any objects of negotiation available are no longer active. An outcome which has been deactivated through this principle can be reactivated if the restriction which lead to any one of the negative restrictions is challenged later. This is called **The Reactivity Principle.**

The Satisfy Move A *Satisfy* is an action which brings about a possible closing point in part of a negotiation, or a transition relevance point (SSJ74). Essentially, something which acts as a *Satisfy* in relation to another outcome accepts it without modification. If a suggestion for Monday is open and someone says, "Monday I'm free" This would act as a *Satisfy.* The resulting outcome is the same as the current unit's local outcome with addition of the restrictions from the previous outcome which was satisfied by it. See Figure

S1: (1) We should meet soon to discuss the plans for the paper.
(2) When can you meet next week?
S2: (3) Monday afternoon looks open.
S1: (4) Monday afternoon I have a class.
(5) I could do it Thursday.

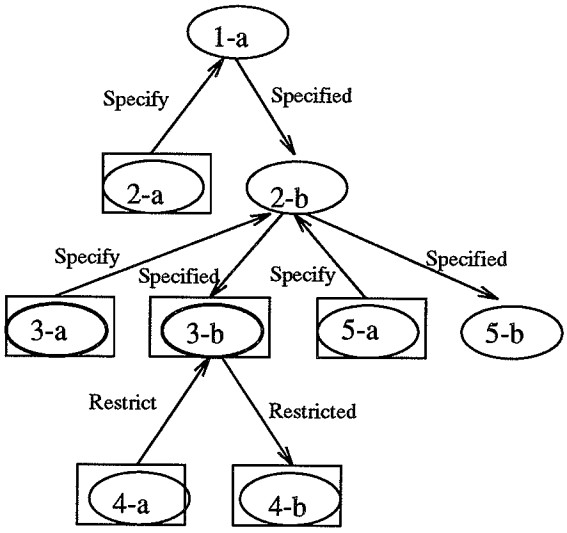


Figure 13: **A Parallel Specify Move**

15 and Figure 16.

Pseudo-sentence unit 6 in Figure 15 produces outcome 6-a which expresses a positive attitude towards a specification of a meeting on Thursday. Previous outcomes available for interaction at this point are 1-a, 2-b, and 5-b. By **The Feature Applicability Principle**, outcome 5-b is selected. So outcome 6-a interacts with outcome 5-b through a *Satisfy* move to produce outcome 6-b which expresses a positive attitude towards a meeting with start time "soon, next week, Thursday" and with topic of conversation "the plans for the paper". By **New Specification Principle**, outcome 6-a is no longer open for future interaction.

OUTCOME: 5-a
SPEAKER: speaker1
START TIME: Thursday
ATTITUDE: positive

OUTCOME: 5-b
SPEAKER: speaker1
START TIME: soon, next week, Thursday
TOPICS: plan for the paper
ATTITUDE: positive

Figure 14: **Outcomes 5-a and 5-b**

S1: (1) We should meet soon to discuss the plans for the paper.
(2) When can you meet next week?
S2: (3) Monday afternoon looks open.
S1: (4) Monday afternoon I have a class.
(5) I could do it Thursday.
S2: (6) Thursday sounds good.

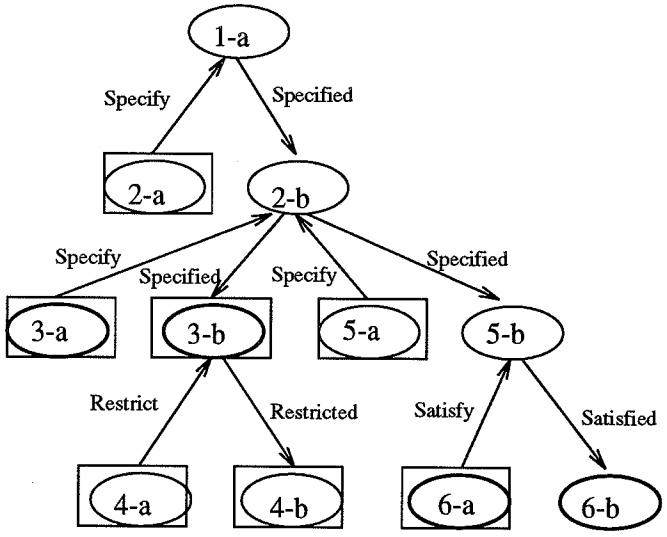


Figure 15: **The Satisfy Move**

OUTCOME: 6-a
SPEAKER: speaker2
START TIME: Thursday
ATTITUDE: positive

OUTCOME: 6-b
SPEAKER: speaker2
START TIME: soon, next week, Thursday
TOPICS: plan for the paper
ATTITUDE: positive

Figure 16: **Outcomes 6-a and 6-b**

The Reject Move A final move is *Reject*. A *Reject* is a move which challenges the appropriateness of a previous move. It produces no outcome since if it is not challenged (i.e. retracted), it makes the outcome which it is in relation to as if it had never occurred. An example of a *Reject* is something like, "Who asked you for a suggestion?" or "How could you even think of suggesting a meeting on my birthday?". I have never observed this move to have occurred in my corpus.

2.2. The Intentional Structure and the Attentional State

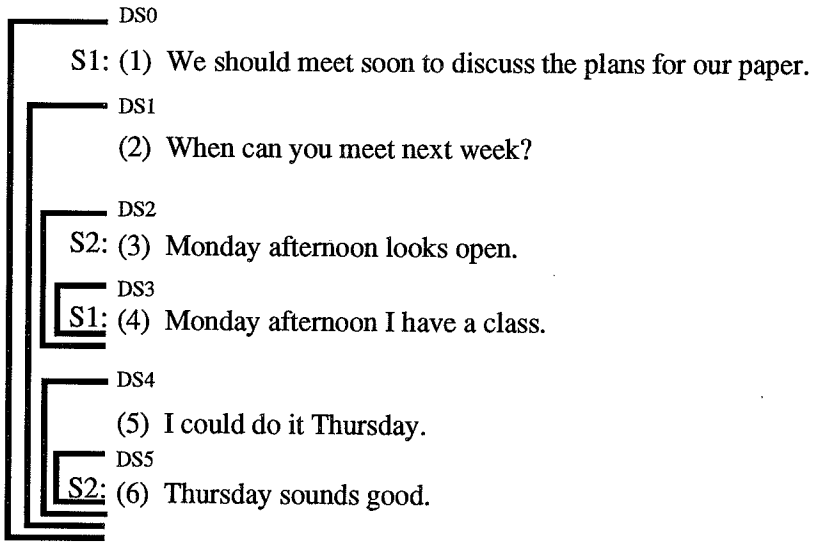


Figure 17: **Example Scheduling Dialogue**

In this section, I will review the principles introduced above. I will then describe how to derive the intentional structure from the interactional structure. Finally, I will discuss how this intentional structure relates to the representation of attentional state in the example in Figure 3 repeated here as Figure 17.

- **The New Specification Principle**: Whenever an outcome W produced by pseudo-sentence X interacts with a previous outcome Y with a *Specify* or a *Satisfy* move to produce outcome Z, outcome W is no longer available for future interaction with outcomes produced by pseudo-sentences subsequent to pseudo-sentence X.

- **The Feature Applicability Condition**: An outcome interacts with the set of previous outcomes which it further specifies such that if the current outcome is more specific than a previous outcome, then it will not interact with any previous outcomes dominating that outcome.

- **The Restriction Principle**: Whenever an outcome W produced by pseudo-sentence X interacts with a previous outcome Y with a *Restrict* move to produce outcome Z, outcomes W and Y are no longer available for future interaction with outcomes produced by pseudo-sentences subsequent to pseudo-sentence X.

- **The Inactivity Priciple**: Whenever an interaction with a previous outcome occurs, all available outcomes not dominating the current outcome which do not make any objects of negotiation available are no longer active.

- **The Reactivity Principle**: An outcome which has been deactivated through **The Inactivity Principle** can be reactivated if the restriction which lead to any one of the negative restrictions is challenged.

The principles above guide the analysis process. After processing each pseudo-sentence unit, each new outcome which is available for interaction with subsequent outcomes is translated into a potential intention of the speaker associated with the outcome to have a meeting which fits the description specified in that outcome. If outcome A dominates outcome B, then the potential intention derived from outcome A dominates the potential intention derived from outcome B. Potential intention A satisfaction-precedes intention B only if the discourse segment in which potential intention A is expressed satisfaction precedes the discourse segment

in which potential intention B is expressed.

I will now describe how the segmentation in Figure 17 can be derived from the interactional structure given the previous discussion. Notice that after pseudo-sentence 1 in Figure 17 is expressed, there is one available outcome, namely 1-a. Speaker1's potential intention to have a meeting which fits the description in outcome 1-a is the DSP of DS0. After pseudo-sentence 2 is processed, there is one new outcome available, namely outcome 2-b. It is dominated by outcome 1-a. So speaker1's potential intention to have a meeting which fits the description in outcome 2-b is the DSP of DS1. Because it is dominated by the DSP of DS0, DS1 is embedded in DS0. By the same token, DS2 is embedded within DS1, where the DSP of DS2 is speaker2's potential intention to have a meeting which fits the description in outcome 3-b. And again, DS3 is embedded in DS2 because the DSP of DS3 is dominated by the DSP of DS2. DS4, on the other hand, is embedded in DS1 because outcome 5-b from which the DSP of DS4 is derived is dominated by the DSP of DS1. Finally, since outcome 6-b is dominated by outcome 5-b, DS5 is embedded in DS4. Given the original simple stack model of attentional state, my coding scheme and the resulting discourse segmentation make predictions about which patterns of referring expressions we should be able to observe in discourse, particularly where parallel structure on the intentional level is present.

3. PARALLELISM: PROBLEMS AND SOLUTIONS

In this section I will present an analysis of each of the examples presented in the introduction of this paper and will discuss the problem which arises in resolving pronominal references if a simple stack model of attentional state is assumed. I will argue that a graph structured stack model of attentional state makes more correct predictions about pronominal reference in this data which is problematic for the simple stack model.

In the graph structured stack model of attentional state, focus spaces for parallel intentions could both be pushed onto the stack as separate tops. The most recently modified top is relatively more salient than the other tops. But the entities in all of the tops are accessible.

3.1. Example 1

The analysis of Example 1 can be found in Figure 18. Unit 1 produces one local outcome, outcome 1-a.

(I)
s1: (1) What about Tuesday afternoons?
(2) The twenty third I'm free from three to five.
(3) And on the thirtieth I'm free after noon.

s2: (4) Okay, both of those are bad for me.

DSP to Outcome Correspondences

DSP0	1-a
DSP1	2-b
DSP2	3-b
DSP3	4-c

Dominates Relationships

DSP0 dominates DSP1
DSP0 dominates DSP2
DSP1 dominates DSP3
DSP2 dominates DSP3

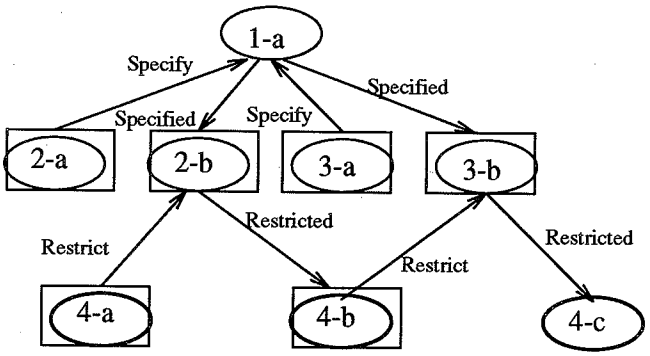
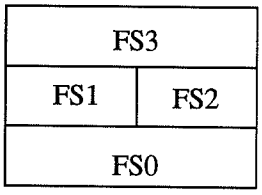


Figure 18: **Example 1**

DSP0 represents speaker1's potential intention to have a meeting which fits the description in outcome 1-a. This is indicated in Figure 18. Unit 2 further specifies the description in unit 1-a. The result of this *Specify* move is outcome 2-b which corresponds to DSP1. Because outcome 2-b is dominated by outcome 1-a, DSP0 dominates DSP1. Unit 3 also further specifies the description in outcome 1-a, but not outcome 2-b since the description of a meeting on the twenty third is mutually exclusive to the description of a meeting on the thirtieth. The result of this *Specify* move is outcome 3-b which corresponds to DSP2. Because outcome 3-b is



After unit 4

Figure 19: **Graph Structured Stack for Example 1**

dominated by outcome 1-a, DSP0 dominates DSP2. Unit 4 produces outcome 4-a as its local conversational outcome. This outcome restricts both outcome 2-b and outcome 3-b since it says something about the status of both of those descriptions. Outcome 4-a first participates in a *Restrict* move with outcome 2-b[8], producing outcome 4-b which then participates in a *Restrict* move with outcome 3-b producing outcome 4-c which is dominated both by outcome 2-b and by outcome 3-b. So DSP3 which corresponds to outcome 4-c is dominated both by DSP1 and by DSP2.

If a simple stack model of attentional state is assumed, and it is also assumed that dominance relationships between DSPs determine which push and pop operations are performed on the stack, then this graph structured intentional structure would require the stack to push a segment on to another segment which had already been popped from the stack. FS0, the focus space associated with DS0, will be pushed on the stack first. Next FS1 is pushed. Because DSP0 dominates DSP1, FS1 can be pushed on top of FS0. Next, FS2 is pushed. DSP2 is dominated by DSP0 but not by DSP1, so FS1 must be popped from the stack before FS2 is pushed. The problem occurs when FS3 is pushed. DSP3 is dominated by both DSP1 and by DSP2. But FS3 cannot be pushed on FS1 since it has already been popped. One could argue that this is not a problem because FS1 can just be pushed on top of FS2 which is still on the stack. But because FS1 which contains the discourse entity referring to the twenty third is not on the stack, it becomes impossible to resolve the reference to "both of those" in unit 4. One could claim that this problem could be resolved by not dividing this discourse into segments at all but rather by considering that it takes place within a single discourse segment, or that each turn is a single discourse segment. Notice that this entails that both suggestions must be represented as a single intention in the intentional structure since every DSP is in a one to one correspondence with a focus space on the attentional level. In this case, both Tuesdays are available for pronominal reference in the following turn. This approach will be addressed again below.

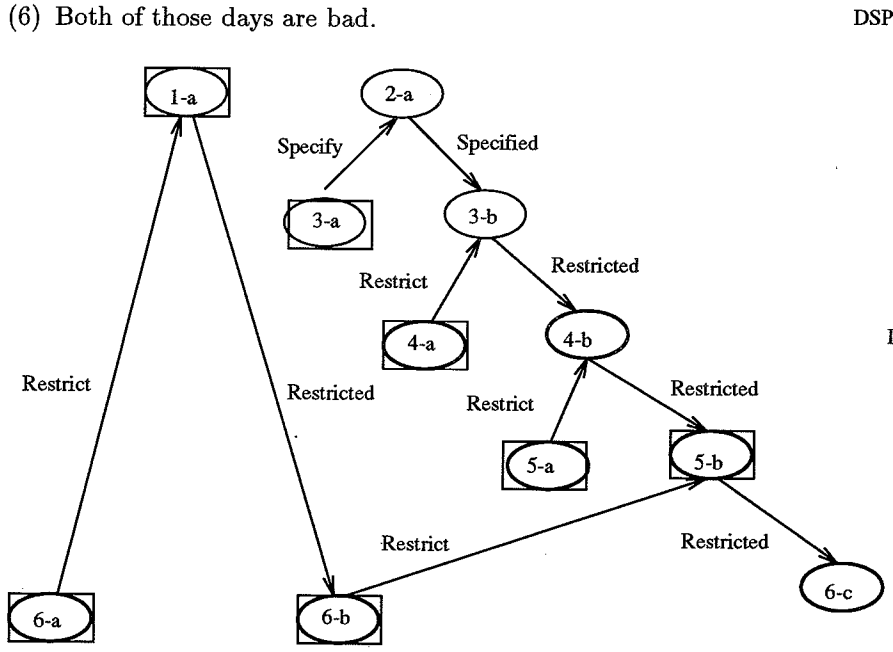
An alternative solution is to model attentional state as a graph structured stack. See Figure 19 for an illustration of the graph structured stack for this example. If both FS1 and FS2 can be at the top of the stack at the same time, there is no problem in resolving the reference to "those" in unit (4).

[8] It could have started with outcome 3-b with the same result.

3.2. Example 2

(II)
s1: (1) Okay, so the fourth is a potential.
(2) Or how about Friday the sixth in the morning?
(3) Do you have any time free that day?

s2: (4) No.
(5) that won't work.
(6) Both of those days are bad.



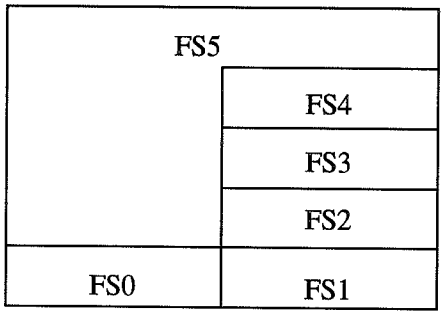
DSP to Outcome Correspondences

DSP0	1-a
DSP1	2-a
DSP2	3-b
DSP3	4-b
DSP4	5-b
DSP5	6-c

Dominance Relationships

DSP0 dominates DSP5
DSP1 dominates DSP2
DSP2 dominates DSP3
DSP3 dominates DSP4
DSP4 dominates DSP5

Figure 20: **Example 2**



After unit 6

Figure 21: **Graph Structured Stack for Example 2**

This example is similar to the previous example in that multiple suggestions are presented in a single turn. The analysis of this example along with outcome to DSP correspondences and dominance relationships

can be found in Figure 20. There is no dominance relationship between DSP0 and DSP1. So in the simple stack model, FS0 would be popped off when FS1 is pushed on. Because DSP2 is dominated by DSP1, FS2 is pushed on top of FS1. This makes the resolution of the reference of "that" in unit 3 perfectly natural. Units 4 and 5 continue to discuss the possibility of a meeting on Friday. Finally, however, DSP5 is dominated both by DSP4 and by DSP0, responding to both possible meeting times. But FS0 has already been popped from the stack by this point. So we run into the same problem as in the previous example. Again, one could argue that the problem can be solved by not segmenting this short discourse at all, or by segmenting it into only two portions, one for each turn. This leaves the problem of ranking entities within a single focus space. When "that" in unit 5 is resolved, neither date is in the immediate focus since neither were mentioned in unit 4. What makes the meeting on Friday more salient is the fact that it is the topic of the previous three units. This seems to indicate that there should be a discourse segment in which units 2-5 are included, but unit 1 is not. But if this is the case, the fourth would no longer be accessible when the reference to "those days" would need to be resolved in unit 6.

An alternative explanation is to model attentional state as a graph structured stack. See Figure 21 for an illustration of the graph structured stack for this example. After unit (2), both FS0 and FS2 are at the top of the stack, but because FS2 is more recent, it is more salient. Therefore, it is possible to explain how the reference for "that" can be made in unit 3. After unit (5), FS0 and FS4 are both at the top of the stack, so it is possible to resolved the reference to "those" in unit (6).

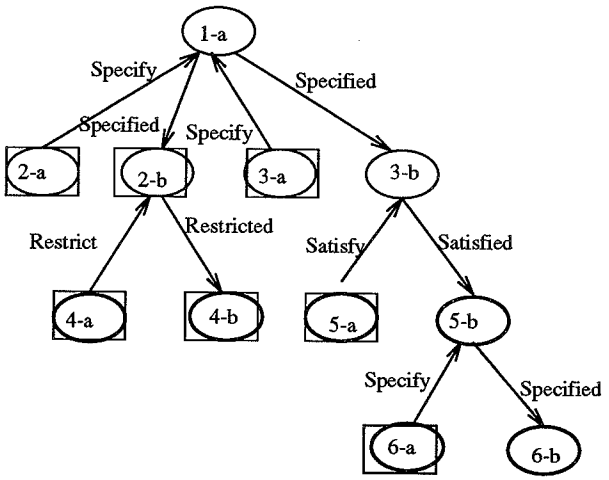
3.3. Examples 3, 4, and 5

In this section I will be discussing examples (III), (IV), and (V). The analysis of example (III) can be found in Figure 22. The corresponding graph structured stack for this example can be found in Figure 23. Likewise, the analysis for example (IV) is displayed in Figures 24 and 25. And the analysis of example (V) can be found in Figures 26 and 27.

Consider example (IV) in Figure 24 which is similar to Example (II) in Figure 20. If all of speaker1's suggestions are associated with the same focus space, there is no explanation for why the suggested meeting on Wednesday is more salient when speaker2 responds using "that". One could argue that a simple recency

(III)
s1: (1) Next week is pretty good for me.
(2) I think next Wednesday afternoon would be best,
(3) or Thursday after two.

s2: (4) Wednesday doesn't seem to be too good.
(5) Thursday is.
(6) Anywhere from two thirty to five on Thursday is fine.



DSP to Outcome Correspondences

DSP0	1-a
DSP1	2-b
DSP2	3-b
DSP3	4-b
DSP4	5-b
DSP5	6-b

Dominance Relations

DSP0 dominates DSP1
DSP0 dominates DSP2
DSP1 dominates DSP3
DSP2 dominates DSP4
DSP4 dominates DSP5

Figure 22: **Example 3**

ordering within the focus space could alleviate this difficulty. This would make the last suggested meeting time the most salient and would take care of these two counter examples. This would fail to explain, however, why the suggested meeting time on Wednesday in (V) in Figure 26 is not more salient than the others when speaker 2 responds. See Figures 25 and 27 for an illustration of the graph structured stack model for examples 4 and 5 respectively. Notice how for example 4 the suggested meeting time on Wednesday is in a separate, more salient top than the other suggestions where for example 5 all three suggested meeting times are within the same top.

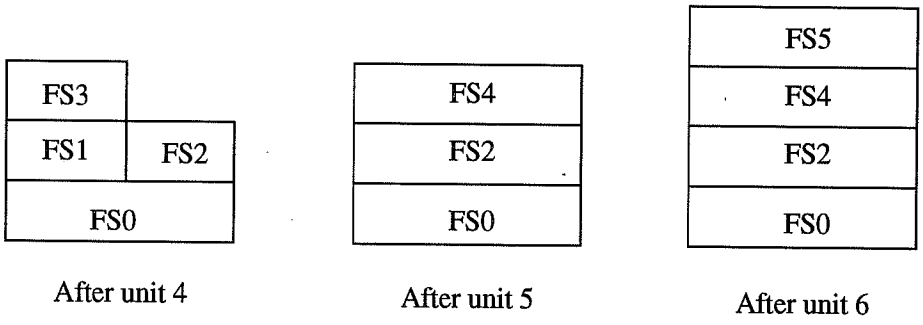
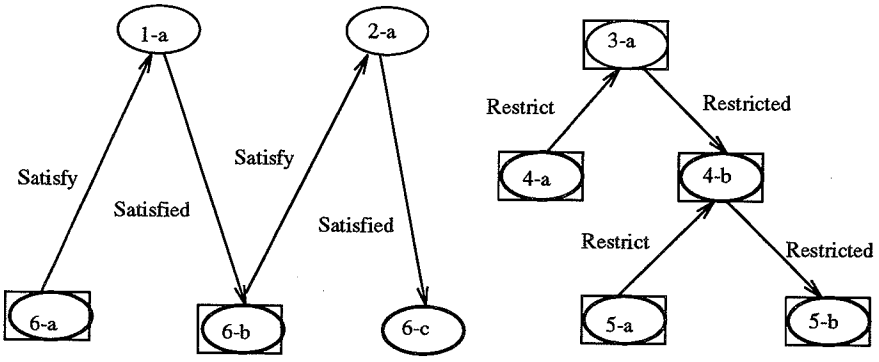


Figure 23: **Graph Structured Stack for Example 3**

(IV)
s1: (1) I am free Monday.
(2) Tuesday is good as well.
(3) I could also make it Wednesday.

s2: (4) That day won't work
(5) because I have class on Wednesday.
(6) But either Monday or Tuesday would work.



DSP to Outcome Correspondences

DSP0	1-a
DSP1	2-a
DSP2	3-a
DSP3	4-b
DSP4	5-c
DSP5	6-b

Dominates Relationships

DSP0 dominates DSP5
DSP1 dominates DSP5
DSP2 dominates DSP3
DSP3 dominates DSP4

Figure 24: **Example 4**

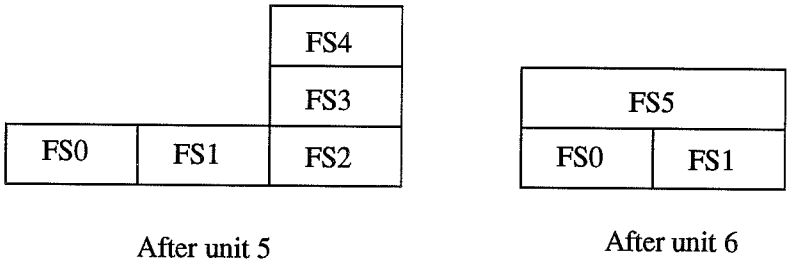
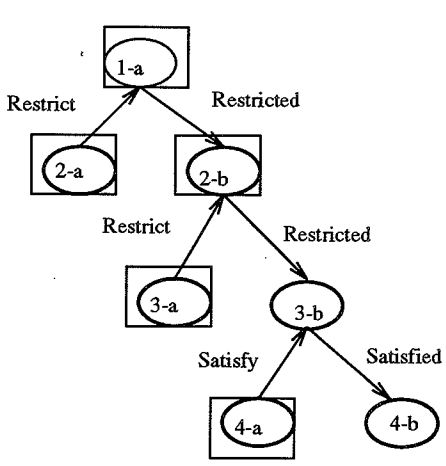


Figure 25: **Graph Structured Stack for Example 4**

Examples (III) in Figures 22 and 23 and (IV) in Figure 24 cast additional doubt on the decision to associate parallel suggestions with the same DSP, thus including the parallel descriptions within a single focus space. In these cases, speaker2's response accepts part of speaker1's set of parallel suggestions and rejects the rest. By associating the full set of parallel suggestions with the same focus space, it is unnatural to represent in the intentional structure which part of speaker1's set of suggestions were accepted by speaker2

(V) ??
s1: (1) I am free Monday, Tuesday, and Wednesday.

s2: (2) That day won't work
(3) because I have class on Wednesday.
(4) But either Monday or Tuesday would work.



DSP to Outcome correspondences

DSP0	1-a
DSP1	2-b
DSP2	3-b
DSP3	4-b

Dominates Relationships

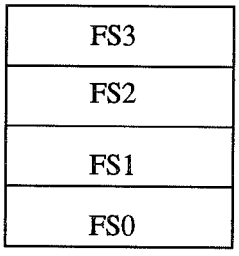
DSP0 dominates DSP1

DSP1 dominates DSP2

DSP2 dominates DSP3

Figure 26: **Example 5**

and which were rejected. If we give up the assumption that focus spaces are in a one to one correspondence with discourse segment purposes as they are represented on the intentional level, then we can avoid this problem. Then it not clear how to independently decide where discourse segment boundaries should lie and when to push and pop focus spaces. It is a much simpler solution to modify the representation of the attentional state from a simple stack to a graph structured stack.



After unit 4

Figure 27: **Graph Structured Stack for Example 5**

This leaves us with the question of how we know when to pop since it is no longer the case that all of the focus spaces in between the new one and the one which it will be pushed on top of are popped off prior to pushing the new focus space. I propose that when an outcome associated with a DSP as well as all of the outcomes dominated by that outcome are inactive according to **The Inactivity Principle**, the focus space corresponding to that DSP is popped off the stack. This predicts that if a suggestion is made and never responded to that it will continue to remain in focus until it is rejected, as in the following example:

(VI)

s1: (1) I am free Monday.

(2) Tuesday is also a good day for me.

s2: (3) That day won't work

(4) What about next Thursday?

s1: (5) Actually, now that I look at my schedule,

both of those days are bad for me.

But this is not the case in the following example which is a slight modification on the previous example:

(VII)

s1: (1) I am free Monday.

(2) Tuesday is also a good day for me.

s2: (3) That day won't work

(4) because I have class on Tuesday.

(5) What about next Thursday?

s1: (6) Thursday is bad for me.

(7) But Friday would work.

s2: (8) No.

(9) Both of those days are bad for me.

The distinction seems to lie in the amount of discourse which has elapsed between when the suggestion for meeting on Tuesday is given and when it is referenced. This distance phenomenon was noted by Walker in (Wal93) where she explored the usage of redundant information in discourse. Her model is based on the limited attentional capacity of the agents participating in the discourse. It is an open question how much discourse must elapse before the entities evoked in segments which have not been popped due to their corresponding outcome becoming inactive become unavailable for pronominal reference.

It is important to note that this graph structured stack model is not equivalent to multiple independent stacks and is therefore not turing equivalent. Note that the graph structured stack is different from multiple separate stacks in two crucial respects.

The first respect is regarding popping behavior. Consider a stack which has three elements: FSA on the bottom, FSB on top of that, and FSC on top of that. Now consider what happens when FSD which is parallel with FSC is pushed. In the multiple stack model, a new stack would be created with FSA on the bottom, FSB on top of that, and FSD on top of that. Now, if the top two elements are popped from the second stack, a situation would arise in which FSB would be popped from the second stack but not from the first stack. This is not possible with the graph structured stack, however. In the graph structured stack model, when FSD is pushed, it is pushed on parallel to FSC. Then if the top two elements along the most salient branch of the graph structured stack are popped, since the second one is the branching point focus space, all of the branches would be popped and only FSA would be left.

The second respect in which the graph structured stack model is crucially different from the multiple stack model is in its pushing behavior. The graph structured stack model is most similar to multiple stacks if you look only above the branching focus space. However, if you have multiple independent stacks, then the bottom focus spaces on each do not need to bear any relation to one another. This is not to case in the portion of the graph structured stack which we are considering here. Because the bottom focus spaces on each branch are pushed on top of the branching point focus space, the discourse segment purposes of the segments corresponding to those focus spaces must contribute to the discourse segment purpose corresponding to the

branching point focus space. You will notice in my longer paper that this only occurs when parallel proposals are made for making a previous proposal more specific. Therefore, the two discourse segment purposes corresponding to the two focusing spaces branching from the branching point will be specializations in some sense of the discourse segment purpose corresponding to the branching point focus space. Therefore, it is not possible for those two discourse segment purposes to bear no relation to one another.

Therefore, the graph structured stack is far more constrained than the multiple stack model and is therefore not turing equivalent.

3.4. Some Further Predictions

This model predicts that in the absence of any parallel intentional structures, all of the predictions of the simple stack model would hold. In the case where there is no parallel structure on the intentional level, then dominance relations among DSPs are sufficient for determining when discourse pops must occur on the attentional stack. There would be only one top at a time, and its behavior would be the same as with the simple stack. And in the case of parallel intentional structures such as those discussed above, entities from each parallel intention would be accessible but the ones from the most recently discussed parallel thread would be more salient. This would suggest that it would be easier to make pronominal references in the non-parallel case since it is much clearer which entities are available for reference. Rather than considering which are relatively more salient than the others, it is simply a matter of whether the entities in question are on the stack or not.

Non-Parallel Single "that"	**Parallel Single "that"**	**Parallel Multiple "that" or "those"**	**Non-Parallel Multiple "that" or "those"**
147	16	13	22

Table 1: **Corpus Study Results**

This is exactly what was found in a corpus study of 50 spontaneous scheduling dialogues where environments where "that" and "those" were used to refer to one or more specifications for meetings were compared. The most common occurrence of pronominal reference occurred in the non-parallel condition with "that" referring to the only entity of its kind on the stack, or the only one mentioned recently enough to be salient

enough to be referred to naturally with a pronoun according to the analyst's judgement. This occurred 147 times in the 50 dialogues. This is more than six and a half times as often as the next most common type of occurrence, supporting the prediction that the most natural pronominal reference would be one where there would be essentially no ambiguity in resolving the reference. There were 13 cases of either "that" or "those" referring to an entire group of suggested times expressed over multiple clauses as in example (I). This supports the prediction that all of the parallel suggestions must remain as salient for the purpose of pronominal reference even when they are expressed over multiple clauses. There were 16 occurrences of "that" referring to an entity in the most recently expressed parallel intention as in examples (II) and (IV). This supports the claim that the most recent specification of a meeting mentioned is more salient than the others even if the others are also salient. Also, there were no occurrences of examples such as the unnatural sounding example (V), but there were 22 cases of "that" or "those" referring to a whole group of entities expressed in a single clause. This supports the claim that when a group of specifications for meetings are in a list within the same unit, there is no difference in salience between them, indicating that the solution is not simply to rank the specification for meetings in the order in which they were presented but that in fact the difference in salience must be structural.

The model also makes predictions about dialogues which should be unacceptable. As with Grosz and Sidner's model, no entities are available for pronominal reference which have been popped from the stack. When a suggestion has been rejected and a new suggestion has been made, the outcomes corresponding to the rejected suggestion are made inactive and the corresponding focus spaces are popped. Therefore a pronoun should not be able to refer to the time which has been rejected. See (VIII) for an example:

(VIII)

s1: (1) Can you meet on Monday?

s2: (2) That day won't work

(3) What about next Thursday?

s1: (4) Actually, now that I look at my schedule,

both of those days are bad for me.

4. CONCLUSIONS AND FUTURE DIRECTIONS

While this study only explores the structure of negotiation dialogues, its results have implications for other types of discourse as well. This study indicates that it is not a structural property of discourse that only those entities evoked on the rightmost frontier of the discourse structure are available for pronominal reference. Further study is necessary in order to explore how widespread this phenomenon is in discourses other than negotiation dialogues.

In this paper I have described an approach to the analysis of negotiation discourse. I have described how to relate my analyses to a version of Grosz and Sidner's intentional level. I used this analysis to argue that it is not possible to explain the patterns of occurrence of "that" and "those" in examples (I) - (V) with Grosz and Sidner's stack model of attentional state while maintaining the one to one correspondence between discourse segment purposes and focus spaces. I then presented a graph structured stack model of attentional state which handles these patterns in the case of parallelism on the intentional level where the simple stack model fails and reduces down to the simple stack model in the absence of parallel structure in the intentional level, making it possible to cover the wide range of the phenomena that their model handles so well. I also presented a corpus study of 50 spontaneous scheduling dialogues which confirms the predictions made by the graph structured stack model of attentional state.

There are many directions for interesting future research which this paper leaves open. First, as I have already mentioned, it is an open question how much discourse must elapse before items not explicitly popped from the stack become inaccessible. There is also the question of how to apply this model of attentional state to discourses with parallelism on the intentional level which are not negotiation discourses. The corpus study presented in this paper only tracked references using "that" and "those". It would be interesting to look at patterns of other types of reference.

REFERENCES

N. Asher. *References to Abstract Objects in Discourse.* Klewer Academic Publishers, Boston, 1993.

M. E. Bratmann, D. J. Israel, and M. E. Pollack. Plans and resource-bounded practical reasoning. *Computational Intelligence*, 3:117–136, 1988.

M. Coulthard. *An Introduction to Discourse Analysis.* Longman, London, 1977.

W. Edmondson. *Spoken Discourse: A Model for Analysis.* Longman, New York, 1981.

B. A. Fox. *Discourse Structure and Anaphora: Written and Conversational English.* Cambridge University Press, 1987.

B. J. Grosz and S. Kraus. Collaborative plans for group activities. In *Proceedings of IJCAI-93*, pages 367–373, Chambery, Savoie, France, 1993.

E. Goffman. Replies and responses. *Language and Society*, 5:257–313, 1976.

B. J. Grosz and C. L. Sidner. The structures of discourse structure. Technical Note 369, SRI International, 1985.

B. Grosz and C. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12:175–204, 1986.

T. P. Klammer. Foundations for a theory of dialogue structure. *Poetics*, 9:27–64, 1973.

W. Labov and D. Fanshel. *Therapeutic Discourse.* Academic Press, New York, 1977.

K. E. Lochbaum. *Using Collaborative Plans to Model Intentional Structure.* PhD thesis, Harvard University, 1994. Technical Report TR-25-94.

L. Polanyi. A formal model of the structure of discourse. *Journal Of Pragmatics*, 12:601–638, 1988.

C. P. Rosé, B. Di Eugenio, L. S. Levin, and C. Van Ess-Dykema. Discourse processing of dialogues with multiple threads. In *Proceedings of the 33rd Annual Meeting of the ACL*, Boston, MA, 1995.

C. P. Rosé. Conversation acts, interactional structure, and conversational outcomes. In *Proceedings of the AAAI Spring Symposium on Discourse Interpretation and Generation*, 1995.

J. Sinclair and R. M. Coulthard. *Towards an Analysis of Discourse.* Oxford University Press, London, 1975.

J. R. Searle. Conversation. In H. Parret and J. Verschueren, editors, *(On) Searle on Conversation.* J. Benjamins Pub. Co., Philadelphia, 1992.

E. Schegloff and H. Sacks. Opening up closings. *Semiotica*, 8:289–328, 1973.

H. Sacks, E. Schegloff, and G. Jefferson. A simplest systematics for the organization of turn-taking for conversations. *Language*, 50:696–735, 1974.

M. A. Walker. *Information Redundancy and Resource Bounds in Dialogue.* PhD thesis, University of Pennsylvania, 1993.

B. L. Webber. Discourse deixis: Reference to discourse segments. In *Proceedings of the 26th Annual Meeting of the ACL*, Buffalo, NY, 1988.