Using RST for dynamic dialogue planning

T. Daradoumis

Report LSI–92–26–R

USING RST FOR DYNAMIC DIALOGUE PLANNING

Thanasis Daradoumis

ABSTRACT

This paper presents an extension of the Rhetorical Structure Theory(RST) from modeling monologues to two-party dialogues. We show how RST can be suitably applied to build an initial synoptic model of the exchange structure and then how rhetorical relations can be especially useful in modeling various types of interruptions that may dynamically occur in an exchange. Our dialogue model and theory are finally illustrated through selected examples from cooperative tutorial dialogues. The result of our work is the construction of a dynamic model of exchange, grounded on the idea of RST, that can handle interruptions to an arbitrary level of nesting in a coherent manner. We contend that our dynamic process model provides a different approach of modeling cooperative and negotiative-type of dialogues than previous models; controls when the completion of the exchange has been reached; facilitates natural passing to the next exchange after the current one is completed; allows smooth resumption of its interruption; and most importantly accounts for the global coherence of the dialogue by knowing the role and function that each (*static* or *interrupting*) move plays in the dialogue at any point.

1. INTRODUCTION

This paper discusses a way of extending linguistic theories about exchange structure and exchange dynamics [Martin 1991] by means of an extension of the Rhetorical Structure Theory(RST) [Mann and Thompson 1987].

RST has been mainly used for analyzing and generating texts(monologues) in a coherent manner. However, no real attempt has been made so far to apply RST in modeling dialogues coherently. This is because things in conversation look differently from the ones in monologues, the latter being termed as a *synoptic(static) perspective* to language, where the overall structure of the text is decided sometimes either all at once or, in some other cases, incrementally but without taking interactions with the user into account.

On the other hand, in a conversational environment, the structure of the text is not decided all at once. Rather, at various points of the text-event participants are faced with sets of alternatives from which they choose. Thus, features of the text are selected interactively; the text itself is viewed as a process and thus takes a *dynamic perspective.*

In other words, the identifying characteristic of a *dynamic model* is that it generates structures as one works one's way through it. Because of this, it provides a good means of modelling the way in which interactions *negotiate* their way through a discourse structure; it shows them as *making choices as they go*. A few attempts have been made so far in formulating a dynamic potential for interrupting moves. These involved making use of flowchart models (such as suggested by [Fawcett et al. 1998] or transition networks (as outlined in [Winograd 1983]). In this work, we discuss a different approach and notation which makes choices available in terms of what has already been syntagmatically determined.

In particular, we claim that discourse(rhetorical) relations are both a useful tool for presenting a dynamic model of discourse and a suitable and important means for building a rich (and sophisticated) dialogue structure - a conceptual model of the dialogue as it progresses, by *relating subsequent utterances to the evolving model* in appropriate ways.

[Moore 1989] also makes this claim by noticing that the construction of such a large structure may be useful in determining when topic shifts have occurred, and aid in making effective use of the previous dialogue when producing new explanations. However, she didn't provide a way of building such a sophisticated dialogue structure. The reason for that was that it was not immediately obvious how all follow-up utterances (and, especially in our case, all types of interruptions that may occur at any point within an interaction) are related to the previous dialogue.

To that end, we present here a first attempt to determine what these relations are and how a dialogue planning module can recognize them in order to build a richer representation of the dialogue structure.

We start from the observation that choices, that a participant can have at any point in the dialogue, are consistently shown to depend on what has already been generated. Since discourses are typically contributed to by more than one interactant, they typically include many points at which a potential contributor can fail, quite legitimately, to fulfil the predictions set up by the earlier contribution(move) of another interactant. It is exactly these points that constitute the various types of interruptions that may occur in a dialogue. As a consequence, it is the types of interruptions and the failed presuppositions(predictions) set up by the earlier contribution(move) that we should precisely determine in order to account for the determination of both the most appropriate relation that will relate the interrupting move to the evolving model of the dialogue and the most suitable answer to the interruption.

In order to determine the type of interruption occurred, we have classified interruptions into two broad classes, according to [Martin 1991]: the class of *tracking interruptions* that focus on the experiential meaning coded in the previous turn and its function is to clarify it by elaborating it in specific ways; and the class of *challenging interruptions* that focus on the interpersonal meaning of the previous move and negotiate degrees of attitude, modality(probability, usuality) and modulation(obligation, inclination). An initial definition of the notions *experiential* and *interpersonal* is provided which results in using the set of *ideational relations* to model tracking and the set of *interpersonal relations* to model challenging interruptions, these two sets of relations taken from the network of discourse relations developed by [Maier and Hovy 1990]. Then, our principal idea of modelling interruptions according to an extension of RST is presented.

Furthermore, representing explicitly the pragmatic presuppositions underlying a participant's utterance(move), we are able to identify which presupposition failed (and caused the interruption), which in turn will allow us not only to generate the most suitable answer to that interruption, but also, and most importantly, determine the most appropriate relation that will coherently link the interruption to the evolving model of dialogue in an appropriate way.

In order to precisely account for interrupting moves, their role and their position in the exchange structure, an initial *synoptic* representation of the exchange structure is presented, again in terms of the RST theory, and then it is shown how interruptions are accommodated and fit naturally into it. In particular, in modelling exchanges, *constituency* relations will be shown using a specific notation, and *dependency* relations caused by

interruptions will be shown using rhetorical relations. Also, rhetorical(logico-semantic) relations are used to model a move consisting of several speech acts or adjacent moves that contribute in achieving the same discourse goal, called *move complex* [Ventola 1987, 1988]. This gives rise to exchange structures that include both synoptically and dynamically generated moves.

Furthermore, our exchange structure allows us to handle interruptions to an arbitrary depth of nesting without loosing control or being bogged down because of too many interruptions or destroying the coherence of the dialogue.

Finally, our dialogue model and theory are illustrated through selected examples from the type of negotiative cooperative tutorial dialogues our domain deals with. In order to build the dialogue model, we present, as a starting point, a lesson plan generated by a Tutoring Strategist. Following the task-oriented structure of the lesson plan, we identify the levels which the conversation consists of and relate them to the corresponding levels of the intentional structure of the lesson plan. The lesson plan actually, as it is presented, presents a synoptic representation of the dialogue, since only the tutor's interactions are shown. Relations that link related dialogues, related subparts(subdialogues) of a dialogue, and related exchanges of a subdialogue are shown and justified. Thus, a rhetorical analysis of the dialogue is done and is shown from the dialogue down to the move level.

Concluding, the main points that our research focuses and makes our new dialogue model differ from the current ones are:

1. Extension of linguistic theories on exchange structure and exchange dynamics.
2. Extension of the RST theory from modelling monologues to two-party dialogues; representation of both a synoptic and dynamic exchange model.
3. Building of a rich(sophisticated) dialogue structure.
4. Explain the role that each part plays in the dialogue.
5. Account for a global coherence and better representation of dialogue structure by identifying and providing a number of contributing structures: pedagogical, instructional, intentional, and rhetorical structure.
6. Coherent treatment of interruptions with respect to the evolving model of dialogue to an arbitrary depth of nesting.
7. Building of a computationally efficient dynamic dialogue model.

2. CURRENT NON-RST APPROACHES TO MODELLING DIALOGUES

Rhetorical Structure Theory(RST) recently developed by [Mann and Thompson 1987] has been extensively applied both to the text analysis and generation. In particular, Mann and Thompson in analyzing texts, have identified a number of rhetorical relations that can be used to coherently relate the various (small and large) units composing a text. Furthermore, in the area of text planning and generation, a number of researchers used rhetorical relations to generate coherent texts [Hovy 1988, Moore and Paris 1990].

However, all of the work concerning the use and application of RST and rhetorical relations has been mostly concentrated on the analysis and generation of multisentential texts. For such monologues, where no or not extensive interaction occurred between system and user, it was fairly easy for a text planner to identify and make use of such relations in order to construct coherent text. Monologues are regarded as the *synoptic* perspective of language.

On the other hand, when features of the text are selected interactively, models take a dynamic perspective, by viewing the text as a process rather than a product. For conversation, however, different approaches have been used so far. Most of the dialogue models have been solely based on a hierarchical structure of the dialogue consisting of several levels. [Cawsey 1989], for example, builds a hierarchical model of dialogue structure by using a set of dialogue planning rules to structure transactions, exchanges and moves that constitute the levels of the dialogue structure. However, it is not certain that it could provide full account for the points 3 to 6 presented in section 1 above.

Research on discourse understanding has been mostly influenced by the work of [Grosz and Sidner 1986] who argue that recognizing the structural relationships among the intentions underlying a discourse is necessary to identify discourse structure. Following this line, [Lambert and Carberry 1991] build a tripartite plan-based model of dialogue that captures the relationship between discourse, problem-solving, and domain actions on adjacent levels. Similarly, the discourse model of [Litman and Allen 1990] can handle two types of interruptions, *clarifying* and *correction*, to an arbitrary depth of nesting.

Nevertheless, none of the above dialogue models account for the relationships that hold among actions (such as, exchanges) that appear to be on the same level of the conversation structure. Moreover, none of them considered rhetorical relations as a sufficient tool for modelling and building a rich (sophisticated) dialogue structure.

[Moore 1989], in her thesis work, is able to use a rhetorical relation to relate a system's answer to a user's specific follow-up question to the system's previous utterance. This was

possible, because system's answer was a natural continuation of its previous utterance, where both of them had been embodied as one unit of text in a plan operator which made use of that relation to link the two parts of the text.

However, she is unable to always determine an appropriate rhetorical relation that would relate any kind of student's follow-up question to the previous dialogue. This fact then makes it impossible to build a rich dialogue history or a more sophisticated dialogue structure that accounts for discourse relations that coherently link any type of interruption with the preceding dialogue. As a consequence, the most important and challenging question is how one can relate dynamically generated interruptions of any type in the middle of a text, i.e., at any point in the dialogue.

To that end, we are based on an initial theory on a dynamic exchange model [Martin 1991], in which interruptions are classified into two large classes, *tracking* and *challenging*, depending on whether they focus on the *experiential* or *interpersonal* meaning of a previous utterance respectively. We then extend this theory in three ways: First, by providing a more precise definition of what the notions experiential and interpersonal mean. Thus, by knowing what we really mean by those terms, the determination of the most appropriate relation to model a given interruption is facilitated, since the classification of interruptions into tracking and challenging corresponds to the relations classification into ideational and interpersonal respectively [Maier and Hovy 1990].

Second, given the above, we claim that the set of ideational and interpersonal relations is sufficient enough to model any type of interruption that falls into the two classes mentioned above. Furthermore, the main idea of the RST theory (i.e., its notion of nuclearity and supporting satellites) applies neatly to both a synoptic model of the exchange and a dynamic one by coherently relating the interrupting sequence of utterances with the previous dialogue in appropriate ways.

Third, rhetorical relations that model both synoptic and interrupting exchanges can be used as planning operators (as they have been used in previous RST-based text planning methods) to control and organize the conversation with the user, thus providing a computationally efficient approach in modelling the dialogue dynamically.

Finally, our model captures the goal-oriented nature of discourse since we distinguish three levels of intentional structure (by differentiating among pedagogical, teaching, and discourse goals/actions) that satisfy the two intentional relations: *dominance* and *satisfaction-precedence*.

In the remainder of this paper, we will first present our synoptic model of exchange, based on [Martin 1991]. Then, we will further extend it by introducing a different notation

other than the Berry's labels[Berry 1981] that represent the moves in an exchange, as well as discuss its structure and representation according to an extension of the RST theory.

Furthermore, we incorporate our exchange model into our three-upper and two-lower level model of dialogue, relating the levels of the curriculum planning (intentional structure) to the corresponding dialogue upper levels(dialogue, subdialogue, exchange). Then, given a specific lesson plan, the rhetorical relations that are used to model each dialogue level are determined. We motivate our choice of rhetorical relations and describe how the different kinds of goals and relations are used to build incrementally our dialogue model as a discourse proceeds. So far, this represents a synoptic model of dialogue.

Next, dynamically generated interruptions are incorporated into the model and a discussion is presented concerning their type and fitting into the dialogue in a coherent manner by means of rhetorical relations. Thus, a further extension of the RST theory is suggested. We also account for an explicit representation of the presuppositions as well as for presupposition failure in the preceding move that caused the interruption.

Our approach is illustrated through selected examples (with and with no interruptions) by also presenting a full RST analysis of the dialogues at all levels.

Finally, we outline the main points of our current research on modelling negotiation dialogues and show two principal broad directions of our future work.

3. USING RST TO REPRESENT A SYNOPTIC MODEL OF EXCHANGE

In our theory of dialogue structure, we consider five main levels: *Dialogue, subdialogue, exchange, intervention(or move)*, and *speech act.* In this structure, a top level can be also considered and is called the *conversation* level. It is the level that stands above all the five levels mentioned before. We will present our full dialogue model and discuss its levels in some more detail in the next session.

We now turn and focus on the exchange level which is probably the most important of all, since it is the level where interactions occur, the point in the dialogue where text generation actually begins, and interruptions are handled. Thus, it deserves a particular attention as well as a neat and precise representation and development of its structure.

An **exchange**, as a structure, consists of interventions(moves) and can be initiated from both parties in the dialogue. In other words, it gains its name from the fact that it is the unit in which some comodity - either information(knowledge) or action - passes from one participant to the other.

Exchanges consisting up to six moves will be proposed for the cooperative tutorial-type of dialogues we are dealing with here. Our work is mainly based on linguistic work on exchange structure as described in [Martin 1991] who also took an initial work of [Berry 1981] as a point of departure.

3.1 The Synoptic Representation of the Exchange Structure

Berry's discussion focuses on exchanges which fit a *polite consensus-collaborative* model, where it is assumed that participants provide the move expected of them on the basis of the previous move. Her model assumes two participants referred to as primary and secondary knowers and actors.

The *primary knower (PK)* is defined as the person "who is supposed to know the information" which is the basis of the exchange and the *primary actor (PA)* as the person who "is actually going to carry out the action". The *secondary knower (SK)* or *actor (SA)* is the person who is supposed to "receive the information or action".

Exchanges can be initiated by either party. The general formula that provides a synoptic representation of "polite consensus" exchanges is the following:

((dX1) ^ X2) ^ X1 ^ (aX1) ^ (X2f) ^ (X1f)

where () indicate optional elements and ^ linear order. X1 is called the **obligatory move** [Martin 1991] in the sense that it must be provided (explicitly or implicitly) so that the exchange be successfully completed. The rest of the moves are called **optional moves.**

Following Martin, we distinguish two general types of exchanges: the ones that describe **Knowledge(Information)** and the ones that describe **Action**. Thus, "X" in the formula can be replaced by "K" or "A" depending on the type of the exchange.

Both types of exchange can begin in any of three ways, and may or may not follow up the obligatory move. Thus, we further classify the exchanges into totally six classes: Give-information/action, solicit-information/action, demand-information and propose-own(offering)-action. Some initial examples illustrate both exchange types below, and then explanation of the notation used is provided.

INFORMATION(KNOWLEDGE)	**ACTION**
1a) Give-information	1b) Give-action
K1 The concept C is defined as follows...	**A1** Have an example
K2f OK	**A2f** Thanks
K1f Good	**A1f** You're welcome
2a) Solicit-information	2b) Solicit-action
K2 Do you have any question about C?	**A2** Will you give me an example?
K1 No, I understand it.	**A1** Of course
K2f OK	**A2f**
K1f	**A1f**
3a) Demand-Information	3b) Propose-own(offer)-action
dK1 What is this?	**dA1** Would you like to do an exercise?
K2 This is a function declaration.	**A2** Yes, please.
K1 Right	**A1** OK, here you go.
aK1 This is how a function is declared...	
K2f OK	**A2f** Thanks
K1f	**A1f** You're welcome.

The labels that represent the moves of the exchanges reflect the following generalizations [Berry 1981]:

dX1: The PK/PA **delays** his/her admission that he/she has the information (or his/her willingness to carry out the action) in order to find out whether the other participant(SK/SA) also knows the information (or wants the action to be done). dK1 is a frequent move in tutorial dialogues, where the tutor attempts to elicit(demand) student's state of knowledge by "delaying" stating his/her own knowledge. dX1, thus, represents the *delaying* of X1 (see examples 3a/3b).

X2: It is the move where the SK indicates his own state of knowledge to the proposition before PK does. For instance, in 2a, SK indicates his/her ignorance by attempting to elicit the information; in 3a, SK indicates his/her knowledge by replying correctly to PK's elicitation.

X1: As mentioned above, it is an obligatory element in every successfully completed exchange. For instance, in 1a, PK states the proposition outright; in 2a, PK states in reply to SK's initiation; and in 3a, PK confirms SK's completion of the proposition.

aX1: It is a particular move of the PK, especially used in tutorial dialogues, where the tutor consistently replays answers to dK1 moves mostly by way of providing an opportunity for reformulation into slightly more technical or abstract language where this is appropriate. So, it principally serves as an elaboration to the SK's answer(K2) and as an extension to PK's confirmation(K1), as in the example 3a. It is called *accept* move.

X2f: It is a follow-up move that the SK may perform and indicate his/her attitude to the preceding move, but does not have to. This may consist of a reinforcement of expectations underlying a K2 move (as in 2a), or represent an indication of the SK's state of knowledge in regard to the proposition given in K1 (as in 1a). "f" stands for "follow-up".

X1f: It is another follow-up move performed by the PK.

3.2 Extensions to the model by means of RST

Our purpose is to consider the theory of RST as a way of modelling and realizing the classes of exchanges identified above as dialogue planning rules(operators) whose function is to control and organize the dialogue with the user(student).

We take, as our starting point, the *notion of nuclearity* in RST which represents the most significant part of a text unit and apply this idea on the exchange unit.

RST theory claims that if a particular nucleus is removed(not presented), then the significance of the material in its satellites will not be apparent; in other words, if we delete all units that function as *nuclei* anywhere in a text, the result will be incoherent and the central message difficult or impossible to comprehend. On the other hand, if units that only function as *satellites* are deleted, we should still have a coherent text that captures the central message, like a synopsis, of the original text.

If we transfer the principal idea of nuclearity to the concept of the exchange, in the exchange level of the dialogue structure, and apply it on the exchange model described above, we note the following observations:

Every exchange refers to a specific discourse goal that is to be accomplished (e.g., "give information about concept C", "solicit information", "propose-own-action", etc). A discourse goal is accomplished only when the exchange, that describes it, has been successfully completed. In our model, an exchange is successfully completed only when the X1 move is provided and supported. X1 is an obligatory element in every successfully completed exchange. If it is not explicitly presented or implicitly inferred, then it is not possible to account for the exchange's completion (and consequently for the accomplishment of the current discourse goal). Thus, the *X1 move* (or slot) in the exchange structure can perfectly play the role of the *nucleus*, in the same sense that nucleus has been defined in the RST theory.

The *optional moves*, then, can play the role of the *satellites* whose function is to contribute or provide support to the success(or completion) of the nucleus (i.e. the exchange). Satellite presence is optional depending on the type of the exchange and attitude of the participant. Thus, for the *give-information* exchange, if the first satellite(follow-up move) is not provided, i.e. the participant's attitude was implicitly positive(supportive) to the K1 move, we can have a perfectly coherent text, represented by the K1 move, which captures and completes the message and discourse intentions of the exchange.

In other types of exchanges, like the *solicit-information* exchange, in order to provide a K1 move(nucleus) we first need to have a K2 move(satellite). However, the need of the K1 is absolutely necessary to account for the successful completion of the exchange. If K1 is not given as expected, then the exchange is considered as suspended or aborted. After the exchange's completion, if we remove the satellite(K2), the exchange still conveys the completion of its goal for which it has been generated(initiated).

For *demand-information* exchanges, in order to provide a K1 move, we first need to delay K1 by a dK1 move and then have the answer(K2 move) that K1 will evaluate.

To conclude, our exchange structure can be modelled as an RST operator, where the obligatory move(X1) will constitute the (fixed) nucleus while the optional moves dX1, X2, aX21, X2f, and X1f the less-fixed satellites, the presence of the latter depending on the type of the exchange and the attitude of the participant. Moreover, the realization of all classes of exchanges is going to be uniform since the nucleus of any exchange is always the same move X1 that plays the same role in each exchange.

Note that although we use constituency (non-RST) relations to represent and realize our synoptic model of exchange, its representation conforms with the whole idea of RST where the obligatory move, that contains the point of the exchange - i.e., it carries (or indicates the completion of) the illocutionary intent of the exchange - is the nucleus, while the optional moves that function to increase the probability of (i.e., contribute to) the success of the obligatory move(nucleus), are equivalent to satellites.

3.3 Advantages and Disadvantages of the Synoptic Model of Exchange

The synoptic representation provides a global view of the exchange which is marked by the discourse intentions underlying the participant who initiates the exchange, and is delimited by the moves that are expected to fill in the slots determined for each type of exchange. Thus, given a discourse goal, we know when and how an initiated exhange starts, proceeds, and completes, its completion corresponding to (i.e. signaling) the accomplishment of the discourse goal it refers to. As a consequence, we can have control of what is generated (i.e., in which move we are) at any point in the exchange.

In fact, this control is not going to be achieved by just using labels that represent the constituent moves of the exchange. As [O' Donnell 1990] has pointed out, there are three problems with these labels: (a) confusion between "functions" and "places" in structure, (b) lack of generalisability, and (c) opaqueness of labels.

In particular, Berry calls these labels both "functions" and "places in structure". However, a "place in structure" marks the point at which various points become available, while "function" refers to the role an item plays in relation to other items in the same structure.

We thus need to separate the concepts of *function* and *places in structure*. For the first, since no rhetorical relations can be used to refer to the role a move plays in relation to other moves in the same structure, some other means (special-type relations) are needed to explicitly define the exchange context and realize the exchange.

Finally, having the model articulated from a synoptic perspective, choices, that are produced opportunistically, cannot be consistently shown to depend on what has already

been generated. A dynamic perspective of the model has to be taken. Then, in a dynamic model, the focus will not be any more on "what type of exhange can we produce", but on " what move can we produce next". We have to look at the development of the exchange move by move, not at the overall shape of the exchange. We have to identify the options open to the participants at each decision point (the *places in structure*) - the dynamic potential.

We focus on the *exchange context(function)*, *places in structure*, and its realization in the next section, while the dynamic representation of our exchange model is discussed in section 6.

4. EXCHANGE CONTEXT, PLACES IN STRUCTURE, AND REALIZATION

4.1 Exchange Context and Places in Structure

As we mentioned before, we realize exchanges as dialogue planning rules (RST operators) that control and organize the dialogue with the user(student). These operators consist of a nucleus(the obligatory move X1) and a number of satellites(the otional moves). The relations, however, between nucleus and satellites are not considered to be rhetorical relations (except the relation between K1 and aK1 move). What kinf of relations are then between the moves that constitute an exchange?

By considering the exchange context that represents the various points(moves) of the exchange structure [O' Donnell 1990], the reasons why these points are identified at a certain place of the exchange and the function(role) that each point plays in relation to other points in the exchange, we identify three special types of relations between moves in the exchange that we call *proposition-delay, proposition-completion* and *proposition-support.*

Proposition delay is the relation between dK1 and K1 moves. *Proposition-completion* is the relation between dK1 and K2 moves, and K2 and K1 moves(in a solicit-information exchange). *Proposition-support* is the relation between K1 and K2f/K1f moves, and K2 and K1 moves(in a demand-information exchange). On the other hand, between K1 and aK1 moves there is the rhetorical relation *elaboration*, derived because of the nature and purpose of the aK1 move.

Then, these relations can be used to precisely describe the state of any move in the exchange and set out explicitly the contextual features which condition the availability of move options. For example, consider the *demand-information exchange*. A dialogue planning rule(RST operator) is chosen which satisfies the discourse goal "demand-information". Then, this operator causes a dK1 move to be produced while at the same time builds the overall shape(context) of the exchange; that is, by the means of relations identified above, it relates the dK1 move to K2 by a *proposition-complete* relation, dK1 to K1 by a *proposition-delay* relation, and K2 to K1 by a *proposition-support* relation. Optional moves, such as aK1, and K2f, K1f can be also related to K1 by *elaboration* and *proposition-support* relations respectively. Thus, an exchange context is incrementally built.

Then, the function(role) that each move plays in relation to other moves in the exchange can be determined and constrain the options available at that point for the kind of move that should follow. In particular, considering dK1 move, the role that dK1 move plays in

relation to the immediatelly following move(K2) and the nucleus(K1) of the exchange is determined by the relations: *proposition-completion* and *proposition-delay* respectively. Thus, dK1's role is to delay K1 and should be completed by K2 (i.e., K2's role is to complete the proposition conveyed by dK1). The two relations above are satisfied only when both K2 and K1 are *inform* moves, in other words, K2 should be a move that completes the proposition(dK1), and K1 a move that expresses the knowledge delayed by dK1. Furthermore, if K1 is such an inform move, that conveys agreement with the completed proposiition(K2), then it also satisfies the *proposition-support* relation between K2 and K1, since K1's role is to support K2.

Thus, moves whose role is to complete a proposition we refer to as *inform* moves, where moves whose role is to support (i.e., agree with the just completed proposition) are a special kind of inform move, that we call *support* moves and may take forms such as "good", "that's right", "ok" (confirming), or "oh" (accepting). In general, initiating moves can be *elicit* (as PK or SK) or *inform*, while expected responding moves, for the synoptic model, can be *inform* or *support*. Thus, "places in structure" can be accounted for as well.

To conclude, "exchange context" determines the state of the exchange at any point; in the synoptic model, the possible states are *completed* and *supported*. "Places in structure" identify the move options open to the participants at each decision point - the dynamic potential; however, for the synoptic model described here, the available move options are *inform* for *proposition completion* and *support* for *proposition support*.

4.2 Exchange Realization

In order to formalize and realize exchanges as dialogue planning rules (RST operators) we use a language similar to that of [Cohen and Levesque 1985, 1990] to represent beliefs, goals, and intentions of both participants(tutor and student). Speakers' role is denoted by PK(primary knower), PA(primary actor), SK(secondary knower), and SA(secondary actor). Nucleus and satellite slots represent the (illocutionary) effects and goals(prelocutionary effects) that a speaker's move may have on the hearer's beliefs. We also note that these operators are not used to generate any text. Rather, its function is to organize the conversation with the user, and control the content planning operators by calling them when content is to be planned and produced(*action* slot). An example of a representation of such an exchange is presented below:

Give-information exchange:

Action: Call Discourse-Plan-D2(Explain(?PK, ?SK, ?concept, ?prop))

Nucleus:

effects: equal(?prop, INFORM)
done(?PK, TURN)
completed(?PK, ?prop)
believe(?SK, want(?PK, believe(?SK, ?prop, ?concept)))

goals: know(?SK, ?prop, ?concept)
BMB(?PK, ?SK, ?know(?SK, ?prop, ?concept))
supported(?SK, ?prop)

Satellite-1:

effects: equal(?prop, SUPPORT)
done(?SK, TURN)
supported(?SK, ?prop)
believe(?PK, know(?SK, ?prop, ?concept))

Satellite-2:

effects: equal(?prop, SUPPORT)
done(?PK, TURN)
supported(?PK, ?prop)
believe(?SK, believe(?PK, know(?SK, ?prop, ?concept)))

This dialogue plan operator shows how the nucleus carries (and indicates completion) of the illocutionary intent(effects and goals) of the exchange, whereas the other components (satellite 1 and 2) raise the probability of success of the nucleus by supporting(i.e., comntibuting to) it as expected. The exchange is said to be successfully completed when all of its goals become true.

It is obvious to see how the goals in the nucleus can be made to hold, i.e., are supported by the satellites: The PK's goal *know(?SK, ?prop, ?concept)* holds since *believe(?PK, know(?SK, ?prop, ?concept))* in satellite-1 is true. *BMB(?PK, ?SK, ?know(?SK, ?prop, ?concept))* holds because of the combination of the true effects *believe(?PK, know(?SK, ?prop, ?concept))* and *believe(?SK, believe(?PK, know(?SK, ?prop, ?concept)))* of both satellites. Finally, it is also evident that the *supported(?SK, ?prop)* goal of the nucleus is true.

Discourse-Plan-D2 is a content planning operator in our library of plans that is called to plan and generate the proposition(discourse goal) that refers to the inform move K1. We also note that the *effects* part of the *nucleus* slot in the exchange represents the state

resulted after the K1 inform move is generated. This state is characterized as the *completed* state, which describes the role of the K1 move in the "give-information" exchange (i.e., K1 is the move that completes the proposition in such an exchange).

This state then constrains the available options for following moves that lead toward the successful completion of the exchange. This constraint is further explicitly represented in the *goals* part of the nucleus and it indicates that a *support* move or no move must follow so that the exhange be considered as successfully completed.

Thus, our dialogue planning operators provide a rich formal representation of the exchange structure, in which *contextual states* (such as *completed*) are seen as preconditions for the availability of various move options, the latter also represented explicitly in the plan operators. In other words, if we are at the point of the nucleus(in this case, at the K1 inform move), *effects* represent the **current state of the exchange**, while *goals* the **desired state**.

At the same time, the *desired state* represents the **relationship** that can hold between the current move and the subsequent one(in this case, it is the *proposition-support* relation that holds), while the *current state* represents the relationship that holds between the current move and the preceding one in the exchange.

Thus, the three relations identified above represent the relationships between the exchange states(context) and exchange move options(places in structure) at any point of the exchange. (We remind that, in our synoptic model, exchange states are: *completed* and *supported* and exchange move options are: *initiate(elicit/inform)* and *respond(inform/support)*).

Furthermore, we can see how choices(exchange move options) affect(modify) the exchange context. For example, for the plan operator(exchange) above, by choosing an initiate-inform move K1 in the beginning, that changes the state of the exchange to *completed* (but yet *not supported*). On the other hand, the support move K2f, changes the state of the exchange from *completed* to *supported.*

This is achieved, by making explicit the effect that a move may have on the exhange context, and thus the effect of that move on the availability of choices at the next point of the exchange; the latter represented in both the goal part and consequently the next satellite part in the exchange.

So far, we have treated exchange as an independent(isolated) structure that models the interaction between two participants, representing and accomplishing specific types of goals, called discourse goals. However, exchange is part of a larger structure, that we call conversation, as discourse goals are part(subgoals) of the decomposition(refinement) of higher-level goals that give rise to what we called dialogues.

Next section describes our global model of dialogue(conversation) in which our exchange model is incorporated at a specific level, called the exchange(or discourse) level. Our dialogue model naturally results from and corresponds to a global curriculum planning which is also discussed and the correspondence between task-generated levels and conversation levels is presented. Finally, discourse relations holding between individual parts of the conversation at each level are identified and prove to be an important means in determining the role that each part can play in the conversation.

5. INCORPORATING THE EXCHANGE MODEL AND DISCOURSE RELATIONS INTO A GLOBAL DIALOGUE MODEL

This section first discusses the building of a hierarchical model of dialogue consisting of six levels that correspond to tutorial task-related levels of a lesson(curriculum) plan, and is loosely based on an initial work on dialogue structure [Diaz et al. 1990]. Then, at each level of our dialogue model, we determine the discourse relations that can be used to appropriately relate the individual parts that a level consists of. We use discourse relations to model our conversation model since they provide a suitable means that account both for the building and for the global coherence of a conversation (teaching session) as well as for the role that each individual part (large or small) plays in the conversation.

5.1 Building a Hierarchical Model of Dialogue(Conversation)

A conversation actually represents a complete instantiation of a global lesson(curriculum) plan. Each part of the conversation (dialogue, subdialogue, exchange) is marked by a specific goal (pedagogical, teaching, discourse), in the sense that these goals delimit and constrain the scope of a conversational part by signaling its start(first appearance of the goal) and its termination (accomplishment of the goal). Then, if the goal cannot be accomplished, its corresponding conversational part is also considered uncompleted(temporarily suspended) or aborted in the extreme case.

We have seen that completion of an **exchange** meant accomplishment of the **discourse goal**(such as, explain concept, verify understanding) that the exhange referred to. Discourse goals in an Intelligent Tutoring System [Fernandez et al. 1988], in which our dialogue model is part of, usually constitute related steps(subgoals) of higher level goals, called **teaching goals**. In other words, information needed for teaching goals(actions) is obtained through discourse goals(actions), so discourse goals can be executed in order to accomplish teaching goals.

A teaching goal may consist of two or more discourse goal, i.e., a teaching goal may take two or more exhanges to be achieved. Since teaching goals constitute a separate level in the tutorial task-related structure, we also define a higher level that represents a structure consisting of two or more related exchanges, that we call **subdialogue**.

Furthermore, in the tutorial task-related structure, teaching goals constitute related steps(subgoals) of top-levels goals, called **pedagogical goals**. In other words, pedagogical goals are achieved through teaching goals, so teaching goals should be executed in order to eventually perform(accomplish) pedagogical goals.

As before, a pedagogical goal may consist of two or more teaching goals; that is, a pedagogical goal may take two or more subdialogues to be achieved. Thus, pedagogical goals, which represent the highest level of the tutorial task-related structure, can then correspond to the highest level that represents a structure consisting of two or more related subdialogues, that we call **dialogue**. Thus, a dialogue model is constructed incrementally by relating the tutorial task-oriented levels(intentional structure) to corresponding dialogue levels(structure).

We can also consider a top level structure in our hierarchical model of dialogue that stands above all the dialogues that constitute the dialogue level and which we call **conversation**. Finally, we complete our conversation model with two more lower level structures, one called **intervention**(or **move**) and is part of the exchange, and the other called **speech act** and is part of the move.

5.2 Use of Discourse(Rhetorical) Relations in Building a Coherent Conversation (Teaching Session)

Our conversation structure is constructed incrementally as we walk depth-first through a lesson plan and execute the steps(goals) that each level of the plan consists of. Our model of conversation is thus a hierarchical model that reflects the tutorial task-oriented structure(intentional structure) that a tutor builds. Then, goals at each level form a tree structure in which each node represents a different task or goal (which can be a pedagogical, teaching, or discourse goal) that the tutor is performing and wishes to achieve. The children of a node represent the subgoals pursued in order to perform(accomplish) the parent goal.

Therefore, at any level of the tutorial(intentional) structure, the children of a node are all part of an overall goal at the immediately higher level and they contribute together toward the accomplishment of the latter. Thus, the children of a node are **related** in some way and we should be able to identify that relation as well as the role that each of them plays with respect to the other and to the higher goal it contributes to. These relations can then be used to model, i.e., coherently relate, the parts of the conversation that these goals correspond to.

By being able to determine discourse relations between any parts in the conversation, the whole conversation can be modelled in a coherent manner and we will be able to identify at any moment what function each part(dialogue, subdialogue, exchange) has in relation to other parts at the same level. Thus, in addition to the two intentional relation *dominance* and *satisfaction-precedence* of [Grosz and Sidner 1986] that are used to identify

reationships between goals on adjacent levels and account for informative satisfaction , in case of dialogue generation we have also to account for the discourse relations that relate goals and identify their role in the dialogue while being at the same conversation level as well as provide suitable ways in modelling interruptions (see section 6).

As a consequence of building such a sophisticated representation of the conversation(dialogue) structure, we will be able to identify similar situations that may occur later in the dialogue (if we know that the same relations hold between similar goals as in previous parts of the dialogue) and thus reuse them, make comparisons, and address stylistic issues by making use of parallel constructions for coherence, avoiding repetitions, etc.

Furthermore, our conversation structure is a combination of different particular discourse structures provided by the discourse relations identified at each conversation level. Each of these structures serves to reflect the sources it results from and a partial ordering on discourse.

At this point, [Suthers 1991] argues that structural relations should be distinguished according to their intentional(informative), pedagogical, conceptual, and epistemological sources. Our strategy is also to factor the discourse structure relations of [Hovy and Meier 1991] network into classes of relations that provide a particular structure of the dialogue. This fact will also provide a different dimension to the current network of rhetorical relations at least as concerns the tutorial-type dialogues we are dealing with.

Thus, at the pedagogical and teaching levels of our dialogue model, the relationships recognized among the intentions(pedagogical/teaching goals) form a set of **pedagogical relations** and provide the **pedagogical structure** of the discourse according to a pedagogical theory of comprehension and learning.

Similarly, at the discourse level, the relationships among the intentions(discourse goals) form another set of **instructional relations** and provide the **instructional structure** according to a theory of what instructional steps are usually followed in interactive tutoring.

Apart from the general and typical pedagogical and instructional reasons, discourse relations at each level are determined by the nature of the domain(pedagogical/teaching) and discourse goals themselves. In particular, pedagogical goals arise according to a student model and the current situation(teaching session context) and describe the intentions and situation underlying a general teaching strategy chosen by the Tutoring Strategist of [Fernandez et al. 1988].

Teaching goals are (a) part of a higher pedagogical goal and (b) each of them plays a particular role toward the completion of the pedagogical goal. Thus, teaching goals are coherently related and the role that one plays with respect to the other can also determine

what relation holds between them. The same reasoning (a) and (b) applies to discourse level as well, so that to capture and determine the relationships among the discourse goals at that level.

Therefore, goals have the capability to **highlight** certain relations. It is then from the moment these goals are posted by the tutor, when the whole generation process and conversation structure building begins.Thus, pedagogical relations are highlighted by pedagogical and teaching goals whereas instructional relations are highlighted by discourse goals.

Furthermore, our claim is that these relations will further influence: the choice of other rhetorical relations, the determination of certain communicative goals, the operations(content selection) on the knowledge base, as well as operations on other linguistic resources such as selection of certain cohesion relations and theme types(thematic progression).

For example, consider the lesson plan shown in figure 5.1. This lesson plan is one of possible plans that can be generated during a tutor-student interaction given a student model and the context of the teaching session. For further details of how lesson plans are generated, see [Fernandez et al. 1988].

By examing this lesson plan, we see that *background* is a rhetorical relation that holds between the intentions underlying the pedagogical plans(tasks) of "reviewing a known concept briefly" and "introducing new concepts" to a student who learned about the concept in a previous teaching session. The background relation is thus a specific relation that models(relates) the discourse intentions that refer to two pedagogical plans and thus will be part of the class of relations that we called *pedagogical relations*.

Note that a certain set of relations is not exclusive, that is, the same relation may belong to two different sets (e.g., it may be both pedagogical and instructional). However, their effect and influence on subsequent decisions are different since they belong to different levels of the dialogue model and their role and function are distinct at each of these levels. Thus, both relations, though having the same name, have to be considered at any decision point at the lower levels which they affect.

Pedagogical structure is the top and middle level structure of a tutorial conversation or lesson plan that clearly defines the main pedagogical and teaching tasks performed during a teaching session, gives global coherence of the conversation as a whole, and annotates a partial ordering on discourse (by applying ordering relations between intentions at the pedagogical and teaching level). Further ordering is provided by the instructional relations that operate on the discourse level, the lower level of the lesson plan, and build the instructional structure of the dialogue by annotating the order according to which

instructional steps have to be carried out so that the teaching tasks be reasonably and successfully completed (e.g., first you describe a concept through an explanation text and then student's support on text's comprehension is sought, by means of a verification question).

As a consequence, ordering relations facilitate the comprehension, reconstruction, acceptance, and retention of the material included. Furthermore, on other levels, that belong to a text planning facility, further ordering is provided by other relations derived from other sources, such as conceptual and epistemological. Ordering is also derived from focus of attention.

Figure 5.1 presents a lesson plan annotated with the pedagogical and instructional relations, identified for this plan, at both pedagogical-teaching and discourse levels respectively.

So far, we have presented a synoptic representation of the dialogue model, in particular that of the exchange model, although we have built it so that choices(move options) open to the participants at each decision point have been consistently shown to depend on what has already been generated. This has been shown, in section 4, by means of the special-type relations *proposition-delay*, *proposition-completion*, and *proposition-support* and the recognition of exchange states(context) such as *completed* and *supported* and allowable move options(places in structure) such as *initiate(elicl/inform)* and *respond(inform/support)*.

Selection of a particular move option was then possible, depending on the current state(context) of the exchange. However, move option selection was restricted to those moves mentioned above since they were the only ones that contributed toward the completion of the exchange(discourse goal) as required by the synoptic model.

There are several other move options that can be available to the participants at each decision point and which do not complete or support the discoursal predictions set up by the preceding move. In this case then, when a move , i.e., the current state of the exchange, is *not completed* or *not supported* by a following move, we say that an **interruption** occurred. This is a dynamic(opportunistic) phenomenon which describes a not desired, though quite legitimate, situation and which our dialogue model should handle properly and be capable of relating any subsequent interrupting utterances to the evolving model in a coherent manner. It is to this dynamic perspective of language that the next section now turns.

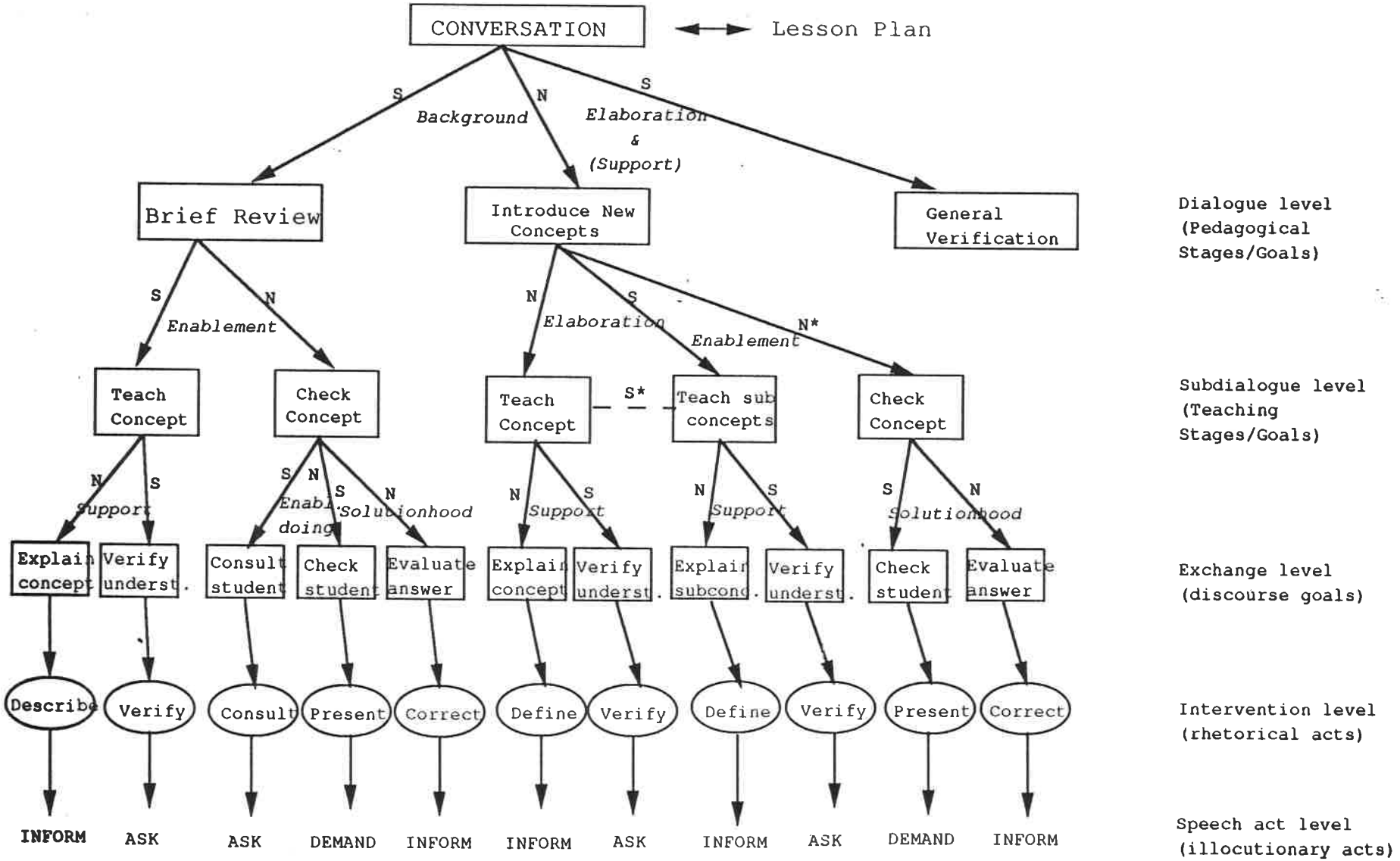

Figure 5.1: Lesson plan generated for a teaching session

6. HANDLING DYNAMIC PHENOMENA BY MEANS OF DISCOURSE(RHETORICAL) RELATIONS

Exchanges can be normally led to completion when both participants contribute to the satisfaction of the presuppositions set up by each participant's move. Discourses, however, typically include many points at which a potential contributor can fail, quite legitimately, to fulfil the predictions set up by the earlier contribution of the other interactant. It is exactly these points that constitute the various types of interruptions that may occur in a dialogue.

As a consequence, it is the types of interruptions and the failed predictions(presuppositions) of the earlier contribution(move) that we should precisely determine in order to account for the determination of the most appropriate relation that will coherently link the interrupting move and the answer to it to the evolving model of dialogue as well as for the generation of the most suitable answer to the interruption.

First, we discuss what accounts for coherence in a discourse by looking at what is allowed to occur if the discourse is to be coherent. Then, in order to determine what type of interruption has occurred and which type of relation appropriately relates it to the evolving model of dialogue, we classify interruptions into two broad classes which correspond to an existing classification of discourse(rhetorical) relations. Let's now turn to the first point.

6.1 Coherence in Discourse

[Tsui 1991] has suggested a sequencing rule that governs what *is allowed to* occur if the discourse is to be coherent. This rule governing coherent sequences is referred as the **Coherence Rule** and states that: *in order for an utterance to form a coherent sequence with the preceding utterance, it must either fulfil the illocutionary intention of the latter, or address its pragmatic presuppositions; if neither, it will fail to form a coherent sequence.* When the rule is violated, the resulting incoherence will be noticed and attended to by participants. Moreover, the violation, in most cases, can be accounted for.

Here, however, we are not dealing with the violation of the rule. Rather, we are concerned with defining and representing what *is allowed to* occur following the occurrence of an utterance, if the discourse is to be coherent.

To that end, we have to look for the candidates for a coherent next utterance. In other words, to account for the coherence, we need to look at the illocutionary intent and the pragmatic presuppositions of the previous utterance. (The term *pragmatic presuppositions* is used loosely to refer to the background beliefs or assumptions of the speaker. All *presuppositions* under discussion here are pragmatic presuppositions).

The type of illocutionary intentions and pragmatic presuppositions depend on the kind of move we are considering. Let's take, for example, an *elicit* move that is used to initiate a *solicit-information* exchange. For this type of exchange, after an elicit move is produced, the state of the exchange is *uncompleted*, i.e., an *inform* move, which completes the proposition uttered, is expected in order to change the state to *completed*. For example, for the solicit-information exchange below, let's consider the candidates for a coherent next utterance:

Tutor: Do you have any questions about concept C?
Student: (a) No, Iunderstand it.
(b) Go ahead.
(c) I don't know what to say.
(d) Well, I am not interested in it.
(e) You perfectly know whether I should have a question or not!
(g) Why do you ask me that?
(h) What do you mean by that?

Student's utterances are all coherent with tutor's question. The illocutionary intent, or purpose, of tutor's question is to get the student to provide a piece of information. It presupposes that:

(i) The tutor has reason to believe that the student has the information.
(ii) The tutor has reason to believe that the student is willing to supply the information.
(iii) The tutor does not nave the information and want it sincerely.
(iv) The tutor has the need and the right to ask for the information.
(v) The student can understand the meaning conveyed.

As [Tsui 1991] also notes, the above list of possible ways of not fulfiling the presuppositions of a preceding utterance is by no means exhaustive. However, it is sufficient for the time being. Most importantly, we explicitly represent both illocutionary intentions and presuppositions in our content-planning operators(discourse plans) that plan and generate text(utterances) for each move in the exchange.

By having all the presuppositions represented explicitly, we can account for the one(s) that failed to *complete* or *support* the current exchange state(move), and that will allow the planner to plan and generate the most appropriate answer back. This fact also accounts for the *extensibility* of our generation facility that is supposed to be able to handle many different types of questions. In other words, if some questions have not been forseen at

design time, it will be possible to easily extend the generation facility to respond to additional question types by adding and explicitly representing the new presupposition types in the plan operators.

Presupposition representation is done using the logic, proposed by [Cohen and Levesque 1985, 1990], for describing and reasoning about these mental states in a world in which agents have to interact with others.

Let's now turn to our second point: identyfying types of interruptions that may occur in a dialogue as well as types of discourse relations that can account for the coherent link of the interruptions to the evolving model of dialogue.

6.2 Types of Interruptions and Discourse Relations

Two main subclasses of interruptions will be distinguished: First, interruptions that focus on the *interpersonal content* of a preceding utterance and attack its validity, called *challenging interruptions*; and second, interruptions that focus on the *experiential content* of a preceding utterance and are usually generated for clarification purposes, called *tracking interruptions*;

We have initially based this distinction on Martin's work on dynamic exchanges [Martin 1991]. However, we extend this work further by (1) defining precisely the notions of *experiential* and *interpersonal* for the type of tutorial dialogues we are dealing with; (2) identifying the most appropriate set of discourse(rhetorical) relations that can be mainly used to model the two classes of interruptions, based on our intuition of what *experiential and interpersonal content* means; and (3) providing a computational method, by means of discourse plans, to handle dynamic exchanges. Next, we look at both these subclasses of interruptions in some more detail.

Challenging Interruptions:

We consider the example of the *solicit-information* exchange presented in the previous section. We can see that student's utterances (a) and(b) fulfil the illocutionary intention of tutor's question. So, they modify exchange's state from *uncomplete* to *complete* and since the type of the exchange is *solicit-information*, the exchange is considered successfully completed(terminated). The difference between answers (a) and (b) is that in (a) the completion is provided explicitly(directly) while in (b) in an implicit(indirect) way.

Utterances (c) through (h) , however, do not provide the information that the tutor seeks. Rather, they fail to fulfil the presuppostions, (i) through (v), of tutor's question

respectively. As a consequence, the current state of the exchange doesn't change. It remains *uncomplete* and an interruption is introduced.

As we can see, from the kind of presuppositions we have presented, these interruptions involve the participant's knowledge, beliefs, desires, or attitudes toward the propositional content of the previous move and attack its validity by expressing disinclination or disability to participate (toward the completion of the current exchange state). We will categorize these types of interruptions into a specific class that we call **challenging interruptions** and we will refer to them by this name thereafter. (Note, that the term *challenge* is used in oposition to *completing* or *supporting* the assumption or pragmatic presuppositions of the preceding utterance. It does not have the connotation of aggressiveness).

Thus, challenging interruptions focus on the **interpersonal meaning** of the preceding move and not only have the potential to suspend(interrupt) but also to abort an exchange and even the current dialogue(lesson plan) itself.

In addition to the pragmatic considerations identified above, specific expressions and gramatical forms also convey *interpersonal* characteristics(meaning) and help in determining whether an interruption is a challenging one.

In particular, interrupting moves that involve interlocutor's knowledge or beliefs are considered as challenging moves when: (a) the interlocutor avoids grading modality(probability or usuality) completely by claiming ignorance. Expessions usually used to describe this situation are: "I don't know", "I don't remember", I don't understand". (b) the interlocutor probes modality by asking, for example after an information-giving exchange: "Is it true?", "Are you sure?".

On the other hand, interrupting moves that involve interlocutor's desires or attitudes are considered as challenging moves when: (a) the interlocutor expresses disinclination or disability towards an action by using expressions such as: "I will not do it", "I can't do it". (b) the interlocutor arbitrates inclination after a offering or solicit action (e.g., T: "Shall I give you an exercise?" -- S: "*Could you/Would you/Must you?*"; or, S: "Will you get an exercise for me?" -- T: "*May I/Will I/Must I?*").

By considering the network of rhetorical relations proposed by [Maier and Hovy 1990], we can see that the set of **interpesonal relations** in the network necessarily involve in their definition the hearer's knowledge, beliefs, desires, or attitudes toward the propositional content of the text. Therefore, it will be these relations that can be used to coherently link a challenging interruption to the interrupted dialogue.

Then, discourse plans used to generate an answer for an interrupting move, involve specifying *revision, support(motivation, evidence, or justification)* of the interrupted

utterance, or *enablement* of an action described in the interrupting move in order to ensure(contribute to) its success (performance).

Grading with respect to attitude, modality(probability, usuality) and modulation(inclination, obligation) can be used to set parameters with respect to which a discourse plan can be mounted to deal with the interruption as well as which rhetorical relation can be selected to coherently relate the interruption with the interrupted exchange.

Setting of parameters will involve: identification of the presupposition challenged(failed); selection of an appropriate local teaching strategy by the tutor which will cause a goal to be posted for the discourse planner; highlighting of the appropriate relation(s) by the discourse goal; selection of the most suitable discourse plan(plan operator) from the set of possible(alternative) plans associated with the failed presupposition(i.e., from the set of multiple plans that can be used to accomplish the goal at hand); knowledge retrieval by means of the plan operator; theme selection; and so on.

For instance, considering again the example presented in section 6.1 above, presupposition (c) is described by the discourse goal "enable the student to provide the information". This goal will then highlight the *enablement* relation, and then an appropriate plan operator will be selected to retrieve content sufficient to achieve the current goal.

Presupposition (d) can be described by the goal "persuade the student so that he/she accepts the previous proposition". This goal may highlight a set of alternative relations such as, *motivation* or *concession*, from which one will be chosen depending on criteria such as the importance of the content of the preceding utterance, i.e., if the tutor considers that the content presented before is very important for the student, *motivation* should be chosen, and so on.

In general, for the solicit-information exchange, the presuppositions identified above, i.e., their associated goals, highlight the following discourse relations:

presupposition (i) highlights the *enablement* or *support* relation.

presupposition (ii) highlights the *motivation* or *concession* relation.

presupposition (iii) highlights the *justification* relation.

presupposition (iv) highlights the *evidence* or *motivation* relation.

presupposition (v) highlights the *evidence* or *background* relation.

Thus, we see how challenging interruptions can be linked, by means of interpersonal relations, to the evolving model of dialogue in appropriate and coherent ways.

Tracking Interruptions:

On the other hand, if the current state of an exchange does not change to *completed* or *supported* by the participant's next utterance and this fact has not been caused by failure of any presupposition identified for the preceding(interrupted) move, then we consider that the participant's interruption is **experientially**, rather than interpersonally oriented.

In such a case, a participant's interrupting move explores the experiential content coded in a previous turn(the interrupted turn) in the sense that it has the function of clarifying the experiential meaning of what has been proposed and elaborating it in specific ways. In other words, its purpose is to obtain some further(may be more specific) information, from the system's knowledge base, that is missing in the preceding(interrupted) utterance so that the participant be able to finally complete or support the initiated exchange(discourse goal).

We classify these types of interruptions into a class that we call **tracking interruptions** and they are normally considered as requests for *clarification*.

Given what we mean with the notion of *experiential*, we note that tracking(or clarification) interruptions involve both the *semantics* of the previous utterance and the participant's *knowledge*. One could say then that these interruptions fail to satisfy a type of presupposition similar to (i) or (v) above.

The difference between tracking and challenging interruptions then, as regards presuppositions (i) or (v), is that, by challenging interruptions, interlocutors avoid grading modality completely by claiming ignorance(e.g., I don't know). However, by tracking interruptions, interlocutors refer in part to their knowledge but their purpose is to extend it wirth some more specific one, by means of *elaboration, circumstance,* or *comparative* methods.

That leads the planner to track the content of its previous (interrupted) utterance and then explore the possibilities of its knowledge base to find the appropriate information that will extend or elaborate the previous content with new one.

For example, the tutor has just introduced two concepts which are subclasses of a previous generic concept. The student may then wish to know any possible differences between the two concepts. This fact is interpreted as an attempt of the student to amplify his/her existing knowledge as regards the two concepts, thus making the tutor choose some *comparative* method to elaborate on the experiential meaning conveyed in a previous turn about the two concepts.

This observation leads us to consider the set of **ideational relations** in the network of [Maier and Hovy 1990] as the best way to model tracking interruptions since it provides discourse relations such as *elaboration, circumstance, comparative*, etc. that describe

specific ways that new information can be appropriately related to the knowledge the interlocutor already has. For example, in the above example, student's tracking interruption posts the goal "explain differences between the two concepts" which highlights the *comparative* relation in the network which will further highlight more specific relations like *similarity* and *comparison* (by walking through the network) depending on the domain knowledge and student model at hand.

To conclude, we have classified interruptions into two main classes: challenging and tracking, depending on whether they are considered as interpersonally or experientially oriented. A clear definition, of what the notions of*interpersonal* and *experiential* mean, was given. Furthermore, we identified two sets of relations(interpersonal and ideational) that can be respectively used to model the above types of interruptions in a coherent manner. What remains to be done now is how to apply the RST theory and its notion of nuclearity in order to appropriately model and represent these relations as well as relate the interruptions, through these relations, to the evolving model of dialogue built so far. This is the point that we turn now.

6.3 Extending RST Further to Model Interruptions

Since interruptions is a dynamic(opportunistic) phenomenon where several options are open to the participants at each decision point, an initially generated text, by a move in an exchange, can be further extended and elaborated in various ways because of subsequent moves(interruptions) of a participant before the exchange completes. Thus, a particular text(move) can be referred to (clarified or challenged) multiple times until the move is finally *completed* or *supported.*

Therefore, although in our synoptic model of exchange, a move may be considered as a satellite, when this move gets interrupted(i.e., not completed or not supported), its role and function in the exchange changes.

Our idea is then to consider the *interrupted move* as a new *nucleus* which is to be completed or supported by subsequent utterances (if we wish our exchange to come to a successful completion). Moreover, two more things have to be modelled: the interlocutor's *interrupting move* and the *answer* of the other participant to the interrupting move.

At this point, we note that the *interrupting move* does not offer any real new information to the nucleus so that, according to the RST theory, to consider it as something different, such as a satellite to the nucleus. Rather, it simply refers to the nucleus by "prompting" it to some further expansion. Then, it should be really considered as *part of the nucleus*, as a criterion and motivation for the nucleus further growth which has not been predicted by the

other participant before. The interrupted move(nucleus) is still the main and important thing, that's why it is considered as the nucleus, while the interrupting move has as function to *provide a reason for extending the nucleus*, that's why it is made part and is attached to the nucleus.

Therefore, it is reasonable to consider that the *answer* to the interrupting move plays the role of the *satellite* whose function is to support the nucleus(i.e., both participants' interrupted and interrupting move) by extending it in specific and appropriate ways, just as requested by the interrupting move. The nucleus must be finally *completed* or *supported* just as a normal move in the synoptic exchange model.

However, many interrupting moves may occur in sequence that refer to the same interrupted move(nucleus). That is, many satellites may be needed to bring nucleus into the state of *completion* or *support.*

Furthermore, a satellite which played the role of an *answer* to a preceding interrupted move(nucleus) may itself become the topic of conversation for a while, be interrupted itself by subsequent utterances, and thus temporarily be viewed (together with its interrupting move) as nucleus as before, and thus needed to be *completed* or *supported* by an appropriate answer(satellite), and so on to an arbitrary depth of nesting.

As a consequence of all the above, a further extension of the RST theory is proposed in order to deal with and model interruptions:

First, we extend RST from relating adjacent pairs of text units to *relating a text unit to a multiple sequence of other text units.*

Second, we introduce the idea of *multiple two-party nuclearity,* in the sense that a nucleus consists of two or more individual units of texts corresponding to two participants rather than to one. A satellite can also be considered as composite in the sense that it can include moves by both participants. As a consequence of the above two extensions, a multiple two-party nucleus is related to multiple, but independent, satellites.

And, third, the satellite role that a text unit had adopted at a certain time can be, quite legitimately, *switch to nucleus role* if that text unit becomes the topic (i.e., the most important part) of conversation for a while. This phenomenon may be repeated as many times as necessary, thus allowing our model to handle and model interruptions to an arbitrary level of nesting.

The latter is further supported by the fact that at any moment in the dialogue, we precisely know at which point of the exchange we are, since we know which is the current state of the exchange, what is expected to modify this state to its next one, and when and how the exchange is finally successfully completed. Thus, we never loose control of the dialogue, though many continuous interruptions may occur. This allows us to smoothly

resume the interrupted move, by usually generating a metacomment that connects the interlocutor to what was going on before, thus maintaining global coherence in the dialogue.

Our dialogue model and theory are further illustrated through selected examples from tutorial dialogues. Two main examples are presented. One that describes a synoptic model and another that discusses the dynamic perspective of discourse. That's the subject of the next section.

7. ILLUSTRATING OUR DYNAMIC DIALOGUE MODEL AND THEORY IN TERMS OF RST

In this section,we present two examples taken from real tutorial interactions ([Fernandez et al. 1988, Diaz et al. 1990]) which further demonstrate the functionality of our new dialogue model and approach. The illustration is done in two steps, following the way our new model has been developed. First, an example with no interruptions is shown describing our initial synoptic model. Then a different example that involves various types of interruptions is presented which clearly illustrate and discuss our new approach to model interruptions coherently by relating them to the evolving model of dialogue in appropriate ways. Finally, for both examples, an analysis is shown in terms of our extension of RST.

7.1 Illustrating our Synoptic Model of Dialogue

Our first example is presented in figure 7.1 and is related to the lesson plan presented in figure 5.1. Since the purpose of this section is illustrative, we present the analysis and discuss the modelling of part of the conversation that could generated according to the lesson plan. In other words, a complete dialogue("Brief-Review") and the first two-thirds of the second dialogue("Introduce-new-concepts"), are presented and analyzed in terms of our extension of RST. The RST analysis of the this example is presented in figure 7.2.

(Brief Review)

(Teach concept)

T1: Last day you learned about logic expressions.....
T2: Do you have any questions about it?
S1: No it's OK.

.....

(Check concept's understanding through an exercise)

T3: Do you want to solve an exercise now?
S2: Yes, please.
T3': OK, here you go.
T4: Specify the following problem: Express in polar representation...
S3: Here is my solution to the problem: Input: ..... Output: ...
T5: OK. Your answer is right.

Figure 7.1: Example of a dialogue with no interruptions (cont'd on next page)

(Introduce New Concepts)

(Teach concept)

T6: Now, we are going to study how to specify functions appearing in the output. Since there are two kinds of output functions (function declaration and function call), we first start with the explanation of function declarat.
T7: Any questions before proceeding?
S4: No, we can go on.
T8: A function declaration is.... and consists of a function head and a function body.
T9: Is it clear?
S5: Yes, I understand it.

.....

(Teach subconcepts)

T10: Now, we are going to descibe the parts of the function head which are the name, domain, range and definition of the function.
T11: Let's first start with defining the name of the function and then we continue with the rest of the parts.
T12: The name of the function.....
T13: The domain.....
T14: The range.....
T15: The definition.....
T16: Do you have any questions?
S6: It's OK. I understand how a function is defined.
T17: OK.

.....

(Check concept's understanding through a problem)

T18: Let's now practice with an exercise.
T19: Given ...
S7: { buggy answer }
T20: You have some little problems.
T21: Remember that the type of arguments has to be indicated.
T22: Try it again.
S8: { buggy answer }
T23: You still have a problem.
T24: You need to use a conditional for the definition.
T25: Try it once more.
S9: { right answer }
T26: Ok.
T27: Now it's right.
T28: Let's now do another problem.
T29: Specify ...
S10: { right answer }
T30: Very good.
T31: Ok, I think we're done with this concept.
T32: So, let's continue with the next one now.

Figure 7.1: Example of a dialogue with no interruptions

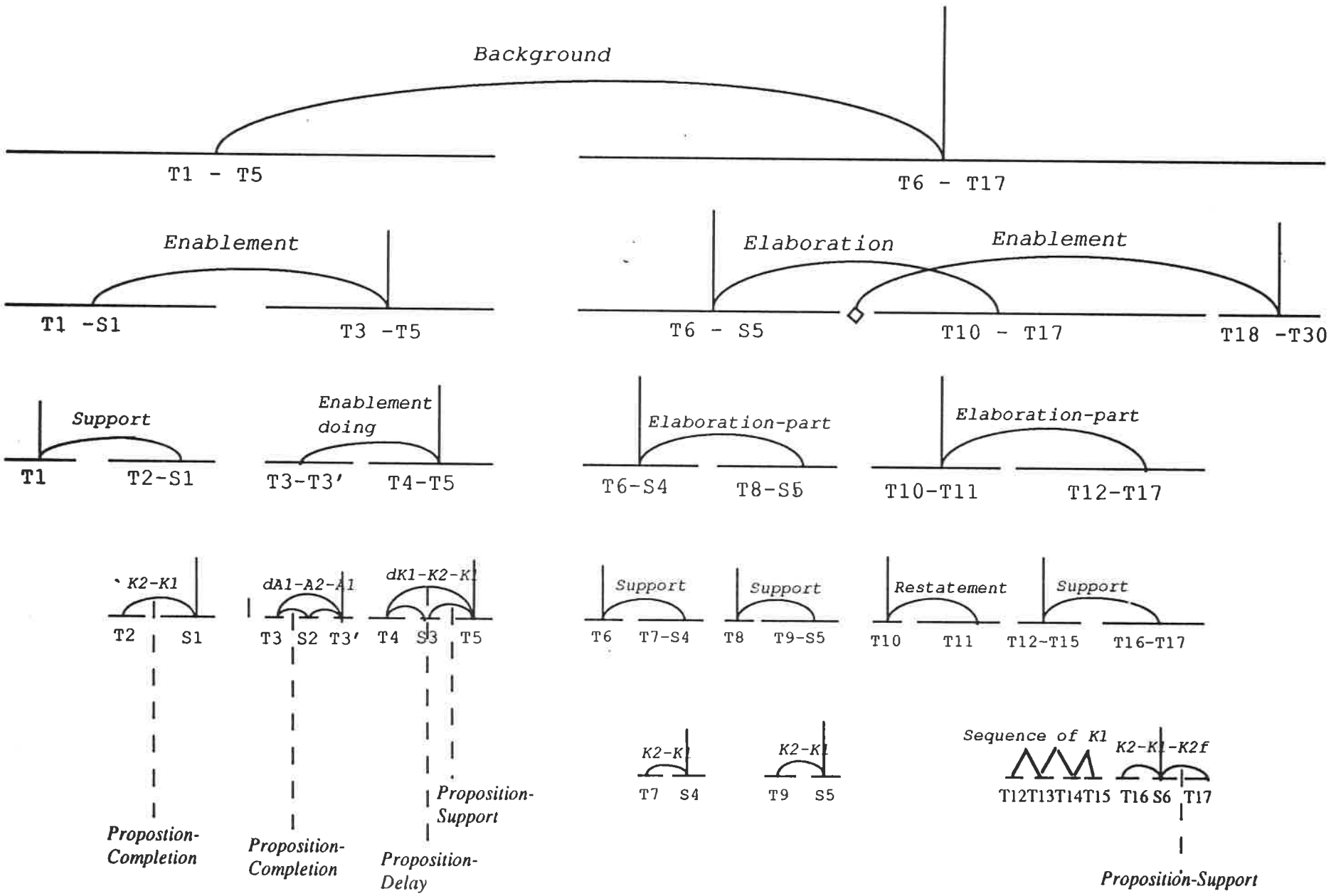

Figure 7.2: RST analysis of example with NO interruptions

Considering figure 7.2, our analysis begins from the upper levels of the dialogue model. In other words, the discourse relations that model dialogues, subdialogues, and exchanges at each level (also annotated in the lesson plan in figure 5.1) are presented here.

The lower level of each dialogue describes the synoptic representation of our exchange model. The moves of each exchange are represented by means of both their labels and the special-type relations that we defined.

7.2 Illustrating our Dynamic Model of Dialogue

Our second example, presented in figure 7.3, is also related to the lesson plan in figure 5.1. This example presents just one dialogue, the "brief-review" dialogue, that involves various types of interruptions. Its RST analysis is presented in figure 7.4.

Let's concentrate in the second subdialogue "check concept's understanding through a problem" after utterance T5 is generated.

T5 is an *elicit* move initiating a *demand-information* exchange. So, initially the current exchange state is *uncompleted*, and the move plays the role of satellite in the exchange. An *inform* move, which *completes* the proposition presented in T5, is expected. However, student's both utterances S4 and S5 cause an interruption that indicates that lack of knowledge unables the student to provide the answer to the problem presented (i.e., to *complete* the T5 move).

Thus, the role of move T5 is immediately changed, from satellite role that the move was going to play in the synoptic representation of the current exchange to the nucleus role that the move is now giong to play until its state changes to *completed*. Moreover, student's interrupting move S4-S5 becomes part of the nucleus, its attachment to it is denoted by a "diamond" ◊.

Then, by examining the presuppositions underlying T5 move, we see that presupposition (i) - "the tutor has reason to believe that the student has the information" - fails to be accomplished. This results in classifying the interruption as a *challenging* one (since it challenges a presupposition of the preceding move) and causes the goal "explain concept *polar representation*" to be posted to the planner by the tutor, as well as the set of *interpersonal* relations to be highlighted in the network. The posted goal then highlights the rhetorical relation *enablement* which is used to relate the subsequent utterances T6-T7 to the composite nucleus represented by utterances T5 and S4-S5.

In the same way, the challenging move S6 and tracking move S7, and their answers T8 and T9-T10 are treated and modelled by the discourse relations *support(solutionhood)* and *circumstance* respectively.

In particular, S6 challenges presupposition (i) again, but this time the *support(solutionhood)* relation is highlighted, since the posted goal was just to provide a supporting(positive) answer to the student.

BRIEF REVIEW dialogue

(Teach Concept)

T1: Last time you learnt about the concept "logic expressions" which has been defined as.....
T2: Do you have any questions about it?
S1: No, you can go on,

.....

(Check Concept's Understanding through a Problem)

S2: but, could you give me first an exercise?
T3: Yes, sure. I was actually going to do so.
T4: Specify the following problem:
S3: -----
T5: Express in polar representation a complex number which is given in a form of an ordered pair.
S4: I don't remember what is the polar representation.
S5: Could you tell me?
T6: Of course.
T7: Given a complex number X coordinate ---> real part, Y coordinate ---> imaginary part, in order to express it in a (module, argument) form, you must ....
S6: Must I also use the form of a pair?
T8: Yes, you must.
S7: Could I use the concatenation instruction CONS?
T9: You are not programming.
T10: The form to specify the par you want, is by saying that ...
S8: I don't understand the form to express it.
S9: Could you please show me an example?
T11: OK.
T12: (the tutor presents the example).
S10: Al right.
S11: Here is my solution to your problem:
S12: INPUT: ..... OUTPUT: .....
T13: OK, the solution is correct.

Figure 7.3: Example of a dialogue with interruptions

On the other hand, S7 is considered as a *tracking interruption* (thus, highlighting the set of *ideational relations*) since it doesn't directly challenge any of the presuppositions underlying T5. The student believes that he/she can provide the answer to the exercise but

he/she is not certain about the *means* that he/she could use to do it. Since the student's means to solve the exercise are wrong, the tutor corrects and extends student's knowledge by first stating what is the real *circumstance* of the situation - utterance T9 - and as a *consequence* of that, the tutor provides the right *means* to the student - utterance T10 (note the cue word "by").

Finally, tutor's explanation in T9-T10 was not clear to the student, since his/her utterances S8-S9 present student's inability to understand the meaning conveyed by T10 (presupposition (v)). In fact, in S9, the student him/herself proposes an alternative way - explanation through example - to the tutor to make student understand the previous content in T10.

Thus, the satellite role that T9-T10 played in the previous interaction, now changes to nucleus role that T9-T10 is going to play in relation to subsequent utterances. This change, from satellite to nucleus, also marks a shift in the focus(topic) of conversation. This means that the nucleus always contains the current focus of conversation; in linguistic terms, in cases of interruptions, nucleus contains the *theme* which corresponds to the point of departure of conversation, whereas the satellite contains the *rheme* - the point of arrival of conversation.

Therefore, by challenging proposition (v), S8-S9 is considered as a challenging interruption, highlighting the set of interpersonal relations and making the tutor post the goal "explain concept through example" to the planner. This goal then highlights the *evidence* relation which is used to relate the subsequent utterances of the composite satellite T11, T12 and S10 (all together) to the composite nucleus represented by utterances T9-T10 and S8-S9.

Utterance S10 is part of the satellite of the *evidence* relation since it *supports* T12 and thus S10 together with T11 and T12 are used to indicate evidence(*completion*) to the nucleus(T10), as well as indirectly to the previous nucleus (T5). Thus, S10 serves to terminate both the nested and the initial interruption, providing expectations for the completion of the dK1 move. In fact, student's next utterances S11-S12 modify the current exchange state to *completed*, and finally tutor's utterance T13, changes it once more to *supported* which results in the successful completion of the current exchange.

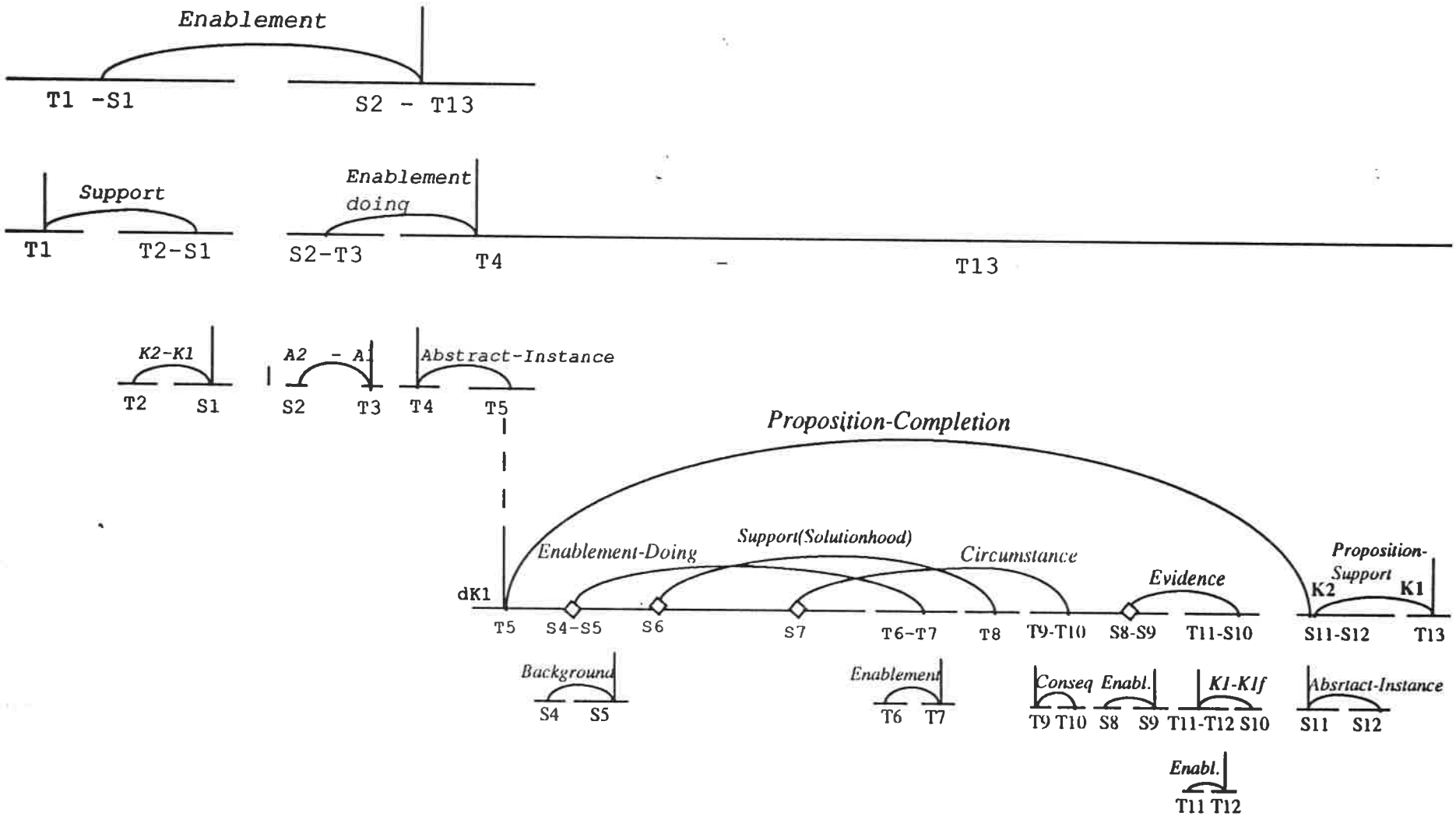

Figure 7.4: RST analysis of example WITH interruptions

8. CURRENT AND FUTURE WORK

Our current and future work involves research in two main directions that both contribute to the building of a new two-dimension planner.

The first is concerned with planning the content of the dialogue by distinguishing between pedagogical, teaching and discourse plan(goals) and modeling each of these three kind of goals(actions) as separate tree structures with links between the actions on adjacent levels. Discourse plans are used to plan and generate text given a discourse goal.

The second is concerned with organizing the conversation with the student and applying control on the presentation and generation of text planned and produced by the content planning operators(rules) above. This is the level we have been mostly concerned here.

In particular, this organization closely follows the tutorial task structure by providing a clear correspondence between the levels of the tutorial task structure(*intentional structure*) and the levels of the dialogue structure. Thus, a hierachical model of dialogue structure is built by relating: pedagogical level with the notion of *dialogue*, teaching level with *subdialogue* and discourse level with *exchange*. Two more lower levels have been identified: the *intervention(move)* and *speech act level.*

Current work is in progress, beyond this preliminary idea, to build a full basic dialogue planning mechanism that will consist of a set of dialogue planning rules at any conversation level. These rules will be then used to structure dialogues and subdialogues on a topic, exchanges with the student and interventions(moves) within that exchange. These rules will actually function as *metaplans* and will be used to control the content planning rules(domain and discourse plans) by calling them each time content is to be further planned.

The principal idea, which differs this model from any other models so far, is to make use of discourse(rhetorical) relations, identified at each level of content planning, in order to model the whole conversation. Discourse relations will be then the dialogue planning rules which will be applied to structure the entire conversation from the dialogue level down to speech-act level, as well as to model any type of interruptions that may occur within an exchange.

For example, considering figure 5.1 again, the *background* relation will be a top-level dialogue planning rule that will be used to relate the dialogue(generated by the pedagogical plan/goal "Brief-Review") with the next dialogue(generated by the pedagogical plan/goal "Introduce-New-Concepts"), by thus building the high level of the conversation structure. The *enablement* relation at the next level(subdialogue level) is another middle-level

dialogue planning rule that is used to model the dialogue "Brief-Review", by coherently relating its two subparts(subdialogues), and so on.

At the exchange(discourse) level, we still need to precisely define and represent every type of exchange that we have identified by determining all the possible choice points open to participants within an exchange so that to be able to cover and model a sufficient number of interruptions. Parallely, for an efficient handling of interruptions, discourse plans should be defined so that all possible presuppostions and effects underlying the plans are explicitly represented.

Furthermore, moves consisting of more than one illocutionary act or moves that contribute to the same discourse goal can be treated in two different ways, depending on whether information is presented to the interlocutors as negotiable or non-negotiable. Then, in the first case, the moves will be treated as separate exchanges, whereas in the latter case will be treated as one exchange, i.e., as comprising a *move complex* according to [Ventola 1987].

Moreover, in the latter case, special type of rhetorical relations, called *logico-semantic relations*[Ventola 1988], will be used to link the several acts of a move or the individual moves together in a move complex. Then, one of these acts(or moves) will be treated as the *head act or move*; it will constitute the *nucleus* of the move complex and will carry the dominant pragmatic value of the whole move. The other acts will be *subordinate acts or moves*; they will act as *sattelites*, and their presence may be necessary to support and contribute to the nucleus, and will in generally have a variable function (such as, argumentative, reformulative, informative, etc). This function will actually determine the rhetorical relation between the nucleus and the satellite moves) of the move complex.

Thus, interventions(moves) can have their own structure as well. An example of a move complex is shown by the utterances S4 and S5 in figure 7.3, where S4 provides the background data enabling the information request (head move) to be presented by the student and be accepted by the tutor. So, futher work is on progress in formalizing and representing the two lower levels of our dialogue model.

Finally, future work involves identifying the ways these high-level relations, together with the domain and discourse(communicative) goals can influence future planning decisions, including highliting(selecting) other relations(that will coherently relate text parts), knowledge selection, thematic progression, lexical cohesion, and text plan growth.

9. CONCLUSIONS

We have presented a dynamic model of conversation based on extensions of linguistic theories on exchange structure and dynamics as well as on an extension of the Rhetorical Structure Theory(RST) so that to account for global coherence of the dialogue especially in cases when interruptions break the normal course of a tutor's curriculum planning.

In building a dynamic model of exchange in terms of an extension of the RST theory, we initially considered some of the advantages of a synoptic representation of the exchange structure to determine special types of relations(*propostion-delay, proposition-completion,* and *proposition-support*) that precisely and neatly define *the exchange context*(the state of the exchange at any point) and the *places in the structure* of the exchange(the available move options open to participant at each decision point).

Based on the notion of nuclearity of the RST theory, we considered the obligatory move(X1) of the exchange as the nucleus, while the optional moves as the satellites. Furthermore, *the nucleus of an exchange is allowed to be missing*. However, the completion of the exchange is manifest by the contextual information. That is, the nucleus can be fully recoverable. Moreover, omission of nucleus can also occur in a move structure where the move may constitute an *indirect speech act*. A similar point is made by [Fawcett et al. 1988] who state that the *head act* of a move may be omitted if it is reconstructable.

To handle interruptions, the general term *dependency relations* have been used so far to account for the relationship of the interrupting move to the interrupted one [Ventola 1987, Martin 1991]. However, no specific relations have ever been determined to account for this dependency.

By classifying the interruptions into two broad classes(*challenging* and *tracking*)[Martin 1991], and defining more clearly and explicitly the different meaning(*interpersonal* or *experiential*) these interruptions may have in discourse, we were able to relate them to two broad classes of discourse relations(*interpersonal* and *ideational*) respectively. Thus, every time an interruption occurs in an exchange, a discourse relation can be determined and used to connect it to the evolving model of dialogue in an appropriate and coherent way.

Moreover, an important contribution to interruption handling has been our further extension of RST that involved definition of new notions such as *relations holding between non-adjacents pairs* of text, *multiple two-party nucleus and satellites* and *satellite role changing to nucleus* according to focus(topic) change in conversation.

Thus, we can see how we can use the strengths and advantages of the RST theory to model lengthy conversations that include interruptions to an arbitrary depth of nesting in a coherent manner.

The most important point in our dialogue model and theory is that the notion of nuclearity is the same as in the initial conception of RST. Dialogue move components of the exchange possess an unequal status and an unequal function. Then, the component designating the obligatory part of the exchange is equivalent to the *nucleus* and carries (or indicates the completion of) the illocutionary intent(goals and effects) of the exchange. The other components simply raise the probability of success of the nucleus and are equivalent to the *satellites*.

To conclude, RST has provided us with the idea that allowed us to: build a rich(sophisticated) dialogue structure, in which we can always account for the role that each part plays in the dialogue; control when the completion of the exchange has been reached; facilitate natural passing to the next exchange after the current one is completed; allow smooth resumption of its interruption; handle interruptions to an arbitrary depth of nesting in a coherent manner; and most importantly account for global coherence and better representation of dialogue structure by identifying and providing a number of contributing structures: pedagogical, instructional, intentional, and rhetorical structure.

10. REFERENCES

[Berry 1981] Berry M. Towards layers of exchange structure for directive exchanges. *Network 2*, 23-32, 1981.

[Cawsey 1990] Cawsey A. Generating communicative discourse. In *Current Research in Natural Language Generation*. Dale R., Mellish C., and Zock M. (eds.), Academic Press, Boston, 1990 (75-102). Also presented at the *2nd European Workshop on Language Generation*, Edinburgh, 1989.

[Cohen and Levesque 1985] Cohen P., Levesque H. Speech acts and rationality. In *Proc. of the 23rd Annual Meeting of the Association for Computational Linguistics*, Chicago, IL, 1985.

[Cohen and Levesque 1990] Cohen P., Levesque H. Rational interaction as the basis for communication. In P. Cohen, J. Morgan, and M.Pollack (eds.), *Intention in Communication*, pages 221-255. MIT Press, Cambridge, Massachusetts, 1990

[Diaz et al., 1990] Diaz A., Fernandez I., Verdejo M. A natural language interface for intelligent tutoring systems. *Education and Application of Computer Technology*. M. De Blasi, E. Luque and E. Scerri (eds.); Fratelli Laterza Publisher-Bari, Italy, 1990.

[Fernandez et al., 1998] Fernandez I., Diaz A., Verdejo M. A cooperative architecture for tutoring tasks. In *Proc. of the 8th Int. Workshop on Expert Systems and their application*, Vol. 2, Avignon, 1988.

[Grosz and Sidner, 1986] Grosz B., Sidner C. Attention, intention, and the structure of discourse. *Computational Linguistics*, 12(3): 175-204, 1986.

[Fawcett et al. 1988] Fawcett R., van der Mije A., van Wissen C. Towards a systemic flowchart model for discourse structure. In R. Fawcett and D. Young (eds.) *New Developments in Systemic Linguistics Vol. 2: theory and application.* London:Pinter. 116-143, 1988.

[Hovy 1988] Hovy E. Planning coherent multisentential text. In *Proc. of the 26th Annual Meeting of the Association for Computational Linguistics*, pp. 163-169, Buffalo, 1988.

[Lambert and Carberry, 1991] Lambert L., Carberry S. A tripartite plan-based model of dialogue. In *Proc. of the 29th Annual Meeting of the Association for Computational Linguistics*, Berkeley, California, 1991.

[Litman and Allen, 1990] Litman D., Allen J. Discourse processing and commonsense plans. In P. Cohen, J. Morgan, and M.Pollack (eds.), *Intention in Communication*, pages 365-388. MIT Press, Cambridge, Massachusetts, 1990.

[Mann and Thompson, 1987] Mann W., Thompson S. Rhetorical Structure Theory: A theory of text organization. In Livia Polanyi, ed., *The Structure of Discourse*. Ablex Publishing Corporation, Norwood, NJ, 1987. (Also available as USC/Information Sciences Institute, TR: RS-87-190).

[Martin, 1991] Martin J. R. *English Text: System and Structure*. Benjamin Press. Amsterdam, 1991 (fortcoming).

[Maier and Hovy, 1990] Maier E., Hovy E. A metafunctionally motivated taxonomy for discourse structure relations. In *Proc. of the 3rd European Workshop on Language Generation*, Judenstein, Austria, March 1991 (38-45).

[Moore 1989] Moore J. *A Reactive Approach to Explanation in Expert and Advice-Giving Systems*. Ph.D. dissertation, University of California in Los Angeles, 1989.

[Moore and Paris, 1990] Moore J., Paris C. Planning text for advisory dialogues. In *Proc. of the 27th Annual Meeting of the Association for Computational Linguistics*, pp. 203-211, Vancouver, Canada, 1990.

[O'Donnell 1990] O'Donnell M. A dynamic model of exchange. In *WORD: Journal of the Int. Linguistic Association, Vol. 41, No 3*. Brend R., Costello J., Chang-Rodriguez E., Embleton S. (eds.), December 1990.

[**Suthers, 1991**] Suthers D. Task appropriate hybrid architectures for explanation. In *AAAI-91 Workshop on Comparative Analysis of Explanation Planning Architectures*, Anaheim, CA, July 14, 1991.

[**Tsui 1991**] Tsui A. Sequencing rules and coherence in discourse. In *Journal of Pragmatics 15*. North-Holland, 111-129, 1991.

[**Ventola, 1987**] Ventola E. *The Structure of Social Interaction: a systemic approach to the semiotics of service encounters*. London: Pinter, 1987.

[**Ventola, 1988**] Ventola E. The logical relations in exchanges. J.D. Benson and W.S. Greaves (eds.), *Systemic Functional Approaches to Discourse*. Norwood, NJ:Ablex. 51-72, 1988.

[**Winograd 1983**] Winograd T. *Language as a Cognitive Process, Vol. 1: Syntax*. Reading/MA: Addison-Wesley, 1983.