Dialogue structure models: an engineering approach to machine analysis and generation of dialogue

Cerian E. Jones

Dialogue Structure Models: An Engineering Approach To Machine Analysis And Generation Of Dialogue

Cerian E. Jones

Laboratory For Natural Language Engineering
Department Of Computer Science
University of Durham

Ph.D. Thesis, 1994

Abstract

The problem area addressed within this research is the processing and understanding of natural English dialogue by computer. The presented work fundamentally constitutes the development of a theory for modelling natural dialogue, and the practical implementation of that theory.

There are two major aspects of consideration within the theory: the modelling of dialogue structure is balanced against certain individual factors. The structure of the dialogue is modelled via the mechanism of the Dialogue Structure Models and its constituent parts, accounting for situational context, participant motivation, participant role(s) and other contributory factors. The individual factors, on the other hand, are peculiar to the dialogue currently in progress, and cannot be pre-determined in the way that structure can. These factors include:

1. The personal characteristics of the participants (their personalities, backgrounds, interests and belief systems);
2. The overall mood of the participant (how (s)he is feeling today; the emotional state of the participant);
3. Instantiation factors relating to the events and circumstances of the particular dialogue in progress. (For example, how a participant reacts intellectually or emotionally to what the other person has just said).

A description of the implementation of this theory is presented, followed by a discussion of the testing techniques used to ensure that the original criteria for success have been met.

To my father, John Raymond Jones,

And the memory of *his* father,

Michael Ellis Jones

Acknowledgments

The research described in this thesis is funded by a Science and Engineering Research Council studentship and a CASE sponsorship by Headway Systems Ltd., of Mountjoy Research Centre, Durham.

I wish to thank my supervisor Dr. R. Garigliano to whom I am indebted for his invaluable help and encouragement throughout this research. I also thank my industrial supervisor, Mr. Ted Fuller, and all others at Headway Systems Ltd. who have offered advice and guidance. I very am grateful to Ms. Olive Hopker, Dean of the Faculty of Computing at Swansea Institute of Higher Education, whose realistic and supportive approach permitted me to complete this thesis.

On a personal level, I wish to thank Nathan and Thelma Abram, for keeping me well fed during the final stages of writing up, and Mark Bowman and Debbie Myers for letting me camp at their flat. Thanks also to everyone in the LNLE at Durham University who made my life and work in Durham so enjoyable, and especially Mark Smith, who saved my sanity on more than one occasion. Thanks to John, Dave, Deborah and Russell for proof reading. My sincere gratitude is expressed to Paul Abram, for sticking by me throughout the arduous months of writing up, as I suspect it was often as painful a process for him as for me, and for listening to all my winges. Finally, my very big thanks go to my parents, especially my father, who sweated every step of the way with me, and whose encouragement, support and interest proved invaluable on numerous occasions.

All work presented in this thesis is the author's, and the views expressed do not necessarily reflect those of the University of Durham or S.E.R.C..

Contents

Extended Abstract

The aim of this extended abstract is to provide the reader with an overview of the work which is described within the main body of this thesis. It is intended to equip the reader with enough knowledge about the research as to understand it and its goals on a surface level, without an excess of detail, and before reading the document as a whole. Obviously, a full understanding of the work can only be acquired after having read the complete document.

The problem area of the research is that of dialogue analysis, and is concerned with the understanding and generation of natural dialogue by computer. The type of dialogue dealt with is restricted to being "standard" dialogue. "Standard dialogue" can be described loosely as dialogue which is stereotypical or "the norm" in a given situation.

The major driving force of the work is that dialogue structure is considered in relation to the individual factors which control and influence dialogue. While it has long been recognised that dialogue structure can be analysed, the concern with structure has all too often over-ridden the consideration of those features which are unique and vital to the specific dialogue in progress. This work addresses this balance between structure and instantiation. For example, a common structure may be found to fit all dialogues occurring within the scenario of a job interview, but no two job interviews are exactly identical. This is due to the individual factors such as the participants' personalities and backgrounds, their beliefs and interests, the mood they are currently in, and instantiation factors such as the way they feel about the dialogue they are presently involved in. Whereas dialogue structure can be pre-determined, these other peculiarities cannot. These two threads of dialogue are of equal importance within the work. The structure of the dialogue is modelled by the DSM mechanism, which is then

influenced by mechanisms for modelling the individual factors and the instantiation factors. Before considering the individual and inastantiation factors, let us consider the way in which we model dialogue structure.

Modelling Dialogue Structure

The framework for modeling dialogue structure is arrived at by using a hierarchical theory. At the top level, a number of template-like schemas have been constructed for a variety of dialogue situations (e.g. informal chat, job interview, lecture), and these are called Dialogue Structure Models. Each Dialogue Structure Model (DSM) contains all the basic information necessary for describing the dialogue structure expected to be found within that particular situation.

Within each DSM, dialogue structure is broken down into a set of fundamental elements, each of which affects the structure of the dialogue in some way by its presence. These "fundamental elements" are called Dialogue Elements (DEs). These can be viewed as the parameters which define a dialogue, each producing a unique behavioural result within the dialogue. The full list of Dialogue Elements is given below, along with a brief description of the influence they bring to bear upon the dialogues in which they are found:

- **Goal**. Indicates whether a dialogue is conducted in order to achieve a process or a task, and what that goal is.
- **Persuasive**. Indicates whether one or more of the participants are attempting to make the other(s) believe something.
- **Information Seeking.** Indicates whether the participant is attempting to acquire information about a topic.

- **Emotional Exchange.** Indicates whether one or more of the participants are attempting to alter the emotional state of the other(s).
- **Number.** Indicates how many participants are involved in the dialogue.
- **Time Limit.** Indicates the time by which the dialogue must terminate or the length of duration (hours/mins).
- **Temporal Progression.** Concerns the manner in which the dialogue unfolds in the "real world" and its constituent sections. (E.g. introduction, main body, conclusion).
- **Register.** Indicates the appropriate vocabulary/vocabularies permitted within the dialogue.
- **Rhythm.** Indicates whether the length of sentence is Short, Medium or Long.
- **Distribution Of Time.** Indicates the length of "speaking time" a particular participant has in relation to the other participant(s) (who can talk the most, who must keep relatively silent etc.).
- **Colour.** Indicates the variety of linguistic phenomena used within the dialogue (e.g. rich use of adjectives)
- **Fixed Topic.** Indicates whether or not the topic of the dialogue must adhere to specified subject(s) only, or may wander freely.

These DEs may vary in strength, which also affects the dialogue structure. For example, *Persuasiveness* is a Dialogue Element present in a variety of dialogue situations, but plays a more prominent role in some dialogues than others; A political debate, for example, is likely to have a higher *persuasive* element than an English lesson.

Each Dialogue Element has a set of constraints attached to it which combine to make up a dialogue action. The constraints can be seen as sets of instructions concerning high-level actions without actually stating precisely *how* to perform the task (e.g. ask a question, pay a complement etc.). The DEs and their attached constraints provide a means for informing the system about the set of possible actions it may take due to the fundamental features of the dialogue, but the selection of appropriate constraints depends upon the DSM in "play". For each DSM, the system is told which constraints it may select from the complete set attached to each DE, thus ensuring "expected" behaviour. For example, if a participant has *Dominance* within a dialogue (s)he may have the power to terminate the dialogue before its natural conclusion. However, this would be highly inappropriate behaviour to take within a number of other situations. Furthermore, a particular constraint may not be appropriate at a specific point in the dialogue. Consider the example of a job interviewer. Although (s)he has greater dominance than the interviewee, it would be highly inappropriate for him/her to terminate the dialogue and walk out before it had reached a natural conclusion. This would only occur in an unexpected instance. Therefore, the constraints are ranked according to their "appropriateness", with the least acceptable/expected behaviour scoring the lowest grade. The constraint with the highest grade is the one which is selected as an action. Therefore, the selected constraints inform the system as to which actions are appropriate/inappropriate and are prescriptions of what to do at each point in the dialogue.

Once the appropriate action has been selected, however, the problem of actually performing that action must be overcome. This is achieved by the existence of plan boxes and analysis boxes. Plan boxes can be viewed as large clusters of detailed instructions regarding the performance of a particular action and are, in effect, the essential core of any implementation of the theory. Thus, the constraints inform as to *what* is to be said, and the plan/analysis boxes inform as to *how* this is to be performed. The analysis of, and reaction to, input is performed by "analysis boxes" which are similar to plan boxes in concept and construction, but contain instructions for understanding the input instead of constructing the output. The number of possible plan/analysis boxes is potentially infinite, and the more that are added, the more sophisticated and wide-ranging the resulting dialogue actions capable of being performed become. Two very large and important plan boxes and one analysis box, are those pertaining to the exchange of emotions. These plan boxes have been carefully considered and constructed due to the complexity and significance of the task of modelling emotional interaction. The exchange of emotions within a dialogue dictates the mood of the participants. Thus, the exchange of emotions needs to be addressed in detail because it directly relates to the instantiation aspects of dialogue discussed earlier. The analysis of emotional exchange is dealt with by the following analysis/plan boxes:

1. **Reaction to emotion.** This is an analysis box, as it concerns the interpretation of input and provides information on how to react and behave if the input carries a certain emotional nature. For example, on the receipt of input, the **Emotional Reaction** analysis box will be triggered and rules retrieved concerning the emotional reaction to different types of input. If, for example, the input is derogatory about oneself, then the result may be to feel angry or maybe surprised. Certain parameters govern the reaction, e.g. if the derogatory input has been received from a friend, the expected response may be shock. If, on the other hand, the other participant is an enemy, then an appropriate response would be to feel angry. Thus, the current "emotional state" of the participant is altered according to the input received.

2. **Expression of emotion.** The **Emotional Expression** plan box instructs as to the expression of certain emotions, which may or may not be the emotion which one is

currently feeling. For example, if someone is happy then this may be expressed in their dialogue. However, there are many occasions when it is improper to show one's true feelings, or to do so may conflict with one's underlying motivations (a car salesman does not express to the customer his joy at having sold him/her a bad car).

3. **Causing emotional reactions**. This plan box contains instructions concerning the invocation of a particular emotion within the other participant(s). For example, a charity organiser making an appeal for funds may try to make the audience feel pity or sympathy for victims. The **Cause Emotion** plan box contains rules on how to go about achieving such an emotional effect.

Note that the plan boxes form only a reactive planning mechanism, and are not long-term planning devices. They only formulate a plan for how to react to what has just been said, rather than formulating strategies which cover longer segments of dialogue, or dialogues in their entirety.

Modelling Individual Factors

As well as emotional interaction, other factors such as personal character and background also influence dialogue structure, and must be considered when attempting to model external behaviour, as is our aim. Rules and instructions are, therefore, constructed within plan boxes concerning the effect these factors produce upon the dialogue structure. For example, if a particular participant has had a close relative killed by terrorist activities, the subject of terrorism will naturally produce a more emotive reaction from that participant. Thus, personal history may affect our dialogue. Consider furthermore, the example of a particular individual who has an exceptionally

shy personality. The dialogue produced by this participant will vary greatly from that produced by an extrovert. This illustrates the bearing which character and other such factors can play upon dialogue structure.

As many pieces of research in the Artificial Intelligence (AI) field are criticised for providing a lack of "hard evidence" to substantiate claims and theories, a practical implementation of this theory has been constructed and is presented within the thesis. The system has been written using a functional programming language, and takes advantage of the feature of lazy evaluation. The main components of the dialogue system are discussed, and the dialogue system is put into context within the LOLITA Natural Language Processing system of which it is a module. The implementation is also discussed in the light of testing techniques used to evaluate its performance in terms of the criteria for success laid out in the methodological discussion.

Chapter One: Methodological Introduction

1.1 Introduction And Chapter Overview

The work described within this thesis presents a theory for modelling human-machine communication at dialogue level. This area of study is known as Dialogue Analysis and Generation, and is a sub-branch of Artificial Intelligence (AI). Human beings frequently communicate with each other using natural language at dialogue level (as opposed to single-sentence interactions or one-participant discourses), and the purpose of the theory presented herein is to model such communication behaviour, thus enabling the development of a computer system capable of conducting human-machine interaction.

The theory and its implementation are dealt with in greater detail in the following chapters, but before the work may be presented it is essential to clarify some methodological issues which bear a relevance to this thesis and its contents. This opening chapter serves, therefore, to present the reader with a clear view of the methodology chosen, and to also outline the boundaries of the work described herein. Once our methodological position has been presented, there follows a discussion relating to the criteria used for evaluating the work. Without these criteria it would be impossible to judge the results of the research, as they constitute a yard stick with which to measure and compare the outcomes of the project. It shall be shown that our criteria have been selected after due consideration and according to methodologically sound principles.

This chapter also includes an explanation of terminology used within the thesis which may be deemed controversial, or from which confusion may arise. Explanations are given for terms which are used in the thesis according to our own definitions.

Firstly, in order that we may place the work within perspective in its field, let us consider the field of Artificial Intelligence, and the aims of this branch of study.

1.2 The AI Goal

Dialogue Analysis is the youngest branch of Natural Language Processing (NLP), itself a sub-branch of Artificial Intelligence. The aim of AI (and, therefore, of its sub-field NLP) is to simulate intelligent human behaviour. Beardon defines AI as:

> *"...the field of research concerned with making machines perform tasks which are generally thought of as requiring human intelligence."*
>
> [p. 12, BEA89]

Under this definition, the focus of AI is not the way in which humans physically process natural language in the brain, or create, understand and perceive natural language in the mind. It is primarily the external behaviour that is to be mimicked within AI work, and as a result the mechanics of human cognition are of small relevance to the research presented in this thesis. Although it is always interesting to read psychology and cognitive science works regarding human processing of natural language, and it may even be encouraging if their beliefs coincidentally provide evidence that our AI models reflect the way in which human beings physically process language, these concerns are not the primary goal of the AI scientist. The ultimate aim of the NLP work is, therefore, purely to produce a system capable of processing natural language in a way which produces tangible, observable behaviour comparable to that expected from a human being within the same situation.

Due to the infancy of the field, no established methodology has yet been laid down as standard for this line of research. Furthermore, AI has traditionally attracted researchers from a variety of backgrounds, and so many confusions can arise as a result of the methodological expectations and preferences an individual reader may bring to the thesis. If the reader has a linguistics background, for instance, his/her preferred methodology may be different from that of someone with a history in software engineering or cognitive science. As NLP receives input from a wide variety of disciplines, there is clearly a need to state one's chosen methodology.

1.3 Approaches To AI

It has long been recognised that natural language is an extremely complex and sophisticated communication tool. Linguists, amongst others, have studied language for many years now, and so it comes as no surprise to the NLP researcher striving to model on a computer the way in which we use natural language, that the task is extremely difficult and large. Due to the enormity and complexity of the NLP problem (discussed in more depth in the following chapter), researchers have endeavoured to introduce scientific approaches to their work which have resulted in either restricting the domain area, or simplifying the task, or both. As a consequence of these measures, the field of NLP has progressed and developed by generally following two different research methods:

- research via a formal approach
- research via a heuristic approach

These two approaches are discussed below.

1.4 NLP Via A Formal Approach

Adopting a formal approach is by no means a new concept within scientific research, and is a widely accepted method of working within a variety of disciplines. The advantages of adopting a formal approach include the provision of a suitable framework with which to compare parallel theories and work, and also the facilitation of testing and proving results. A formal approach, however, is not always the ideal path to choose for two main reasons:

1. Formalisation can often reduce the work domain to a scale which is over-simplified, thus producing a restriction in wider applications of the work. Furthermore, as the formal approach focuses on the general case as opposed to the specific, the parameters of the research are adjusted, or hypotheses underlying the work are over-simplified in order to accommodate the methodological approach. Bearing in mind that NLP is so complex and difficult a task, and that natural language is in itself hugely varied and intricate, it may be argued that the purely formal approach is unsuitable for adoption within this field, at least at this stage of its development.
2. Doubts about the usefulness of the formal approach within practical implementations cannot be dismissed without thought. It is unlikely that a formal model could be built for "real-world" situations, but even if this was possible the

resulting model would probably require a large amount of time and space, and the behaviour of the model would include a greater level of detail than necessary.

Although useful for abstract applications, a formal approach, therefore, carries certain shortcomings within a practical implementation. These are usually overcome by adopting a more heuristic approach to the work.

1.5 NLP Via A Heuristic Approach

The second approach to the NLP problem has been to tackle the task using a much more pragmatic methodology. Due to the complexity and size of the NLP task as mentioned above, this approach has lead researchers to examine and model more realistic situations but on a smaller, more restricted scale. It is then hoped that the scale can be enlarged, and principles derived from the specific example applied to the more general. This approach, however, is not without its risks or undesirable side-effects:

1. Adoption of the heuristic approach runs the risk of producing ad hoc rules which lack structure and sound basis. The models or theories produced by working on a small example may not then be portable to the more realistic larger scenario, resulting in the construction of "toy" systems. Ad hoc heuristics should, therefore, be avoided, although this is not always possible with this approach.
2. Another drawback of this more pragmatic approach is that there may be a tendency to select the example cases to fit the model rather than vice versa. The criteria for success are rendered invalid in such an event, because they are not

objectively produced. If this occurs, then the results would be of little use in other applications or wider domains.

3. Thirdly, it is not an infrequent criticism that work conducted according to a heuristic methodology lacks sufficient evidence of a structured, concrete design. In cases where more general claims are made, the designs of solutions referred to are usually omitted. It becomes extremely difficult, therefore, to reconstruct a clear set of requirements and design by which the results can be repeated. There is a danger that work which cannot be seen, or tested or run in a tangible manner will be dismissed as an exaggeration of the truth. This is a factor which has lead to many pieces of AI work being dismissed as unsubstantiated claims.

As shown above, both the formal and heuristic approach to NLP research have their shortcomings. This has caused some attempts at working according to a methodology placed somewhere between the two, endeavouring to utilise the advantages of both approaches and avoid the drawbacks. This third approach involves the development of heuristics which are more general, and which are arrived at by the use of a structured, principled approach.

1.6 The Adopted Approach

The problems connected with the two approaches outlined above have serious consequences when conducting research within the field of NLP. Due to concern within the LNLE at the University of Durham [1] to avoid these problems, the third

[1] Laboratory for Natural Language Engineering

approach has been adopted of using a methodology which lies between the two extremes. This work has, therefore, been conducted according to formalised principles, with a view to developing soundly-based heuristics which can be used in the wider domain as well as the restricted.

This is partly achieved by the careful selection of example cases. For the reasons stated in section 1.5, it is essential to select the class of example cases according to independent criteria and not simply because they complement the model. The class of examples selected within this work is chosen from the class of dialogues conducted in natural English between human participants, but restricted to those dialogues which may be generally considered as "standard" or "the norm" within the given situation. The set is further restricted to exclude certain features of natural language which prove extremely difficult and complex to model by machine. A more detailed discussion of the example cases is given in section 1.7 below, along with a methodological justification for their selection.

The research method adopted within this project, and the group at Durham, has its roots in Natural Language Engineering (NLE) rather than the more traditional field of computational linguistics (CL). Aspects of NLE include:

- SCALE: the size of the system must be sufficient for realistic large-scale applications. Properties such as the vocabulary size and the grammar coverage (e.g. measured by the perplexity, the number of rules, or the types of text which can be manipulated) are critical.
- INTEGRATION: there are two aspects of integration. Firstly, the system components should not make unreasonable assumptions about other parts. This frequently does not happen when specific NLP problems are tackled in isolation.

Secondly, a system component (such as the semantic representation) should be built so that it can be used to assist other parts of the system (e.g. the generator).

- FEASIBILITY: For example, hardware requirements must not be too great and the execution speed must be acceptable. This process incorporates making the system and its components efficient.
- MAINTAINABILITY: how useful the system is over a long period of time. Maintenance of a large system has been proved to be an important aspect of the software life-cycle
- FLEXIBILITY: ability to modify the system for different tasks in different domains. This can be considered as adaptive maintenance but is separated to emphasise the difference between major adaptation to accommodate large changes in functionality (flexibility) and more subtle adaptation due to, for example, changes in the environment (adaptive maintenance).
- USABILITY: The system must support the applications end users want and be user-friendly.
- ROBUSTNESS: This is a critical aspect of large-scale systems. To quote Galliers and Sparck Jones [GAL93] "while it (robustness) may not be a serious problem for any individual application, it has to be faced up to in general". Robustness concerns not only the linguistic scope of the system but how it deals with input which falls outside of this scope.
- USE OF A WIDE RANGE OF TECHNIQUES: NL engineered systems must use a wide range of AI techniques. Where they are available it is of course advantageous to use long-standing well-worked and general theories from computational linguistics and logic (for example, set-based semantics or multi-sorted object oriented logic). A key aspect of NLE is that, when these theories are not available, alternative methods are employed. These methods range from more localised theories (which despite being unable to cover global possibilities are

sufficient to handle what is required), knowledge based approaches, individual heuristics and adaptive or evolutionary techniques. Incorporating such a range of methods means that the development of the system does not get stalled due to the difficulty in following a particular logical or linguistic theory while the benefits of such well established theories can still be enjoyed.

1.7 The Example Cases

In accordance with the adopted methodological approach, the following example cases were used within the work.

1.7.1 "Standard" Dialogue

"Standard dialogue" is dialogue which could be seen as stereotypical or "the norm" in a given situation. For example, if we are listening to a conversation taking place in a job interview, it would be expected that the interviewer will choose which topics to discuss, will ask more questions than the interviewee, and will have a greater control over the dialogue. If the interviewee suddenly decided that (s)he did not want to discuss a certain topic, but would rather talk about another topic of his/her own choice, then this would be unusual, and not expected. What may be classed as "standard" dialogue, therefore, depends upon the context in which the dialogue is taking place. In the case of an AI system playing the role of a dialogue participant, the computer must react to input and create output in a manner which would be expected of an "intelligent" human with respect to the context of the interaction.

Due to the infinitely large number of example cases of dialogue which one may produce, some degree of restriction is necessary . However, by restricting the class of examples to those dialogues which are "standard", as opposed to freak or inappropriate within their situational context, we exclude only those examples which are the exceptions rather than the rule. The task of modelling the majority class of "standard" dialogues certainly cannot be dismissed as trivial, and so to exclude anomalous example cases can hardly be viewed as restricting the example class to a toy scale.

1.7.2 Other Restrictions

Due to the complexities of natural dialogue, it was found necessary to restrict the variety of natural language to a certain degree by not considering language containing metaphors or humour. These aspects of natural language pose extremely difficult phenomena to model within NLP, and to include them would take the work far beyond the scope of the project. However, we feel that the selected class of examples is still sufficiently large and contains an adequately rich variety of features to avoid claims of having selected the examples to fit the theory. The development of a theory and its implementation for modelling dialogue void of these features would in itself be an innovation in the field at this present time. It would be a remarkable achievement at this present time to build a dialogue system capable of handling "straight-forward" dialogue, and intricate complexities such as metaphor or humour can come later. A sound base is needed before we can add optional features to it and make it more varied or sophisticated in its behaviour and the type of dialogue it can handle. It is argued, therefore, that the restrictions we have imposed upon our example cases have been necessary, and yet avoid the possibility of reducing the work to a trivial scale.

1.8 Analysing Dialogue At The General Level

The challenges of dialogue analysis have been tackled from a variety of angles, as is illustrated by Chapter Three in which we present some of the pieces of work already conducted within the field. Some have adopted the approach of analysing the linguistic phenomena of dialogue in order to model it on the computer. Such analyses are really operating at the discourse level according to the definitions given later in this chapter, and do not tackle the specific features of rich interaction present within dialogue as we define it.

Others, however, have recognised the need to develop a more generalised approach to dialogue processing, (although it shall be seen in Chapter Three that such theories are far less common than the linguistic approach, and evidence of implementations are even less frequent than the theories). The problem is not tackling one aspect of the task alone, per-se: it is how to integrate this with the rest. Without a general theory of dialogue, it would seem futile to attempt to construct a dialogue system. A structured framework must be set out to deal with dialogue on a general, abstract level as well as examining such finer details as the linguistic, grammatical construction of the language. By accounting for a wide range of motivational aspects of dialogue, it is possible to conduct a dialogue, albeit on a very basic level, which does not include such features as anaphora, ellipsis etc. These "niceties" can come later. It is a well known principle of scientific methodology that it is often better to tackle the general problem rather than a particular one. This is so because at the general level the unimportant but eye-catching details tend to disappear, and patterns emerge which could not be seen at the local level. Thus, by working on dialogues in general, the aim is not only to encompass situations not yet analysed in a unique framework, but also to produce better explanations for the classic examples in the literature.

1.9 Criteria For Success

In order that the work can be evaluated and judged, a set of criteria are needed against which the outcomes and results can be measured. These criteria are set out here, and a discussion of how the project results compare with them is presented in Chapter Six, which concerns the testing and evaluation of the research. The criteria which we use to judge the work are:

- the definitions of AI given in section 1.2 apply to the work;
- that the work is on a "real world" and not "toy" scale;
- that the behaviour produced by the system passes certain tests, ensuring that it can be classed as "natural";
- that the programming and coding of the system passes tests which ensure that system has been constructed in a sound manner, both from a coding and a logical point of view.

The type of evaluation tests performed within the project conform to Galliers and Sparck Jones' definition of investigations. They define two types of evaluation test [GAL93]:

- **Investigations** concerning the operational systems/setups of a work to establish performance characteristics;
- **Experiments** concerning attempts to answer the questions "what would happen if so and so?" as opposed to "what is happening?"

The reason why our tests are largely of the investigation type is that to conduct experiments within the system would necessitate a greater amount of time than was available under the constraints of the project. The type of tasks involved in dialogue analysis are not trivial tasks, and the sort of questions which one would like to ask would demand a large amount of time and effort to be devoted in order to arrive at some answers. For example, we might like to ask "what would happen if we built a module to handle the long-term planning issues involved in dialogue?". However, the development of tools, theories and models necessary for constructing such a system module would constitute a research project in its own right. Therefore, due to the enormity and difficulty of the dialogue analysis problem, our tests largely fall into the category of investigations.

The criteria for success are discussed in greater detail below.

1.9.1 The AI Goal

As the work presented in this thesis is a piece of AI research, the ultimate aim of the work must, therefore, be to produce a system capable of mimicking intelligent human behaviour. The general aim of the work is to produce a theory for modelling natural English dialogue, the implementation of which would allow human-computer interaction in English at dialogue level, and in a reasonably "natural" manner. (The term "natural" is defined in section 1.9.2 below). This can be checked by performing the behavioural tests discussed in the next section.

1.9.2 Behavioural Tests

As stated, the aim of our work is to imitate human behaviour as closely or "realistically" as possible. In other words, we strive to produce behaviour which is as "natural" as possible. By "natural", it is meant in a manner which could be considered representative of the dialogue expected between two human participants in the same situation. The intention is to facilitate human-machine interaction which could easily be believed to be human-human interaction. Stilted language, inappropriate registers or grammatical constructions within the given context, and unexpected responses to input are all to be avoided. In order to evaluate the behaviour of the model, a number of tests may be performed:

1. The output may be measured against "real" data. This means that the theory developed within this work could be used to model a dialogue which has occurred naturally in the physical world, and the results examined to see if a correct analysis of the dialogue is given. If not, then the project has not been successful. If, on the other hand, the theory can be used to correctly analyse a naturally occurring dialogue, then this can be reasonably deemed successful. Of course, we recognise that the precise definition of a "correct analysis" is a somewhat grey area. The understanding of dialogue is a subjective matter, and thus it is difficult to give a definitive description of what constitutes a successful analysis. However, we take the view that if the analysis of the dialogue can be judged as an interpretation acceptable to the majority of people, then this will be classed as a correct analysis.

2. The implemented system can be made to actually perform a task in practice. This could possibly involve the system playing a role within a human-machine dialogue.

In order to consider the work to be successful, a human-machine dialogue should be possible with the implemented system.

3. A test which is becoming more and more popular for use in the development of dialogue systems is the so-called "Wizard of Oz" test [AND93, DYB93]. In this situation, human beings interact with a machine without knowing whether or not the actions of the machine are simulated by a so-called wizard. The human "wizard" is concealed from the other human participant. This method of testing allows us to judge how a human-human dialogue participant would behave in a given situation, thus providing us with information about the behaviour we are attempting to emulate.

4. Another method of assessing the behaviour of the model is to measure its performance against certain benchmarks. For example, dialogues which have been used by other people, and are possibly deemed as classical test data in the area, can be used as input for the undertaken work. If the input is not analysed correctly, or reacted to in an inappropriate fashion, then the criteria for success has not been met and the work must be amended. However, it must be stressed that in the relatively new, recently emerged field of dialogue analysis, few such "benchmarks" have been firmly established, and those which are well-known are often such small fragments of language as to be of little use for us. Despite this, however, we choose a dialogue which is well-known in the literature and test the model using it.

1.9.3 Work On A "Real World" Scale

As was claimed earlier, methodological shortcomings in existing dialogue analysis work can result in the domain of the research being constricted to a size which leaves the end result of little practical application. As has been stated, the example set for this work has been left sufficiently large as to avoid criticisms of over-simplification or restriction. The resulting implementation will, therefore, be less application-specific than other systems constructed from models using miniaturised domains. No restrictions of application or end use are placed on the work, and the resulting system should be capable of handling a wide variety of phenomena (excluding those which were named earlier as outside our chosen set of example dialogues). A toy-scale system would only allow a very limited interaction with a user, restricted possibly in terms of domain, or richness of interaction, or both. The work undertaken for this project aims to avoid all of these restrictions.

Also, a real world dialogue analysis system cannot practically live in isolation of other NLP work. Dialogue analysis work assumes the existence and support of other NLP systems such as a grammar, parser, semantic analyser etc. The results of dialogue analysis also has effects upon other work such as dialogue generation. In order to build systems on a real world scale, these features should be linked together to form a whole functioning NLP system.

1.9.4 Programming Tests

Another level at which the success of the work may be evaluated is found at the programming level. As was declared earlier, this work has been produced with a strong commitment to implementation. As the thesis progresses, the reader will be presented with a description of both the theory and the practical implementation of that theory which have been developed for this project. Any computer system needs to be tested for syntax errors which would prevent successful running of the system, and also for logical errors, which could result in the system behaving in a manner other than the intended manner. The software written for this implementation should be no exception. If the system crashes due to coding errors, then this would be unacceptable. A certain degree of robustness in programming terms is necessary.

In addition to the above programming issues, dialogue analysis systems must also address the issue of forming an integral module within an overall NLP system. They must exist in relation to other NLP modules, such as semantic analysers, syntactic parsers, grammars etc., even if these modules are only simulated and not actually implemented. Input from such modules has to be received and understood so that the dialogue analysis module can make use of the information the rest of the system produces. Similarly, output may need to be passed on to a generation module or back to some other module to provide feedback. In order to build real world systems, dialogue analysis work must combine with work in other areas of NLP, and this should be reflected in the design and implementation of our practical systems. The implementation should, therefore, be capable of receiving input and sending output to and from other parts of the overall NLP machine.

1.10 Terminology Issues

Throughout the thesis, technical terms which are non-controversial and have standard or established definitions within the NLP field are briefly defined in the glossary. Technical terms which may not be considered standard within the NLP community, or are used in a special manner by a particular author are defined both in the glossary and also when first encountered in the thesis. Terms which may be deemed controversial because they are either undefined or used ambiguously within the field are defined below in order to avoid confusion or mis-interpretation of the intended meaning:

1.10.1 Controversial Terms

Below follows an explanation of controversial terms and how they shall be used throughout the thesis.

1.10.1 (a) Discourse

The term "discourse" will be understood in this thesis to mean a sequence of sentences which are linked by context and linguistic features, and is one level "higher" than the sentential level. Multi-participant interaction is not a pre-requisite of discourse under this definition. Thus, the words on this page, a newspaper article or monologue may be considered as discourse. There is no rich interaction present in discourse, which is the subject of our next definition.

1.10.1 (b) Rich Interaction

A rich interaction (as opposed to a straight forward interaction) is an exchange between two or more parties, and may involve such features as sub-dialogues, interruptions, complex shifts in focus and other linguistic phenomena which occur when two or more participants are involved in communication. For example, a rich dialogue would not simply involve a user asking a question and the system replying, but would possibly involve the system asking a question of its own, refusing to answer a user's question, changing the topic, going off on a tangent, the involvement of emotional exchanges etc.. It is the presence of this complexity and wider variety of features/phenomena that defines a rich interaction.

1.10.1 (c) Dialogue.

Dialogue is defined herein as a *rich interaction* between two or more participants (either human or machine) which consists of a sequence of sentences which are connected both linguistically and contextually. Monologues and straight-forward prose do not, therefore, class as dialogue under this definition. Dialogue can be seen, therefore, as discourse with the added presence of rich interaction.

1.10.1 (d) Standard Dialogue

Standard dialogue, in this thesis, means dialogue which is typical or *"the norm"* within a given situation (as opposed to *unexpected* or *unlikely* dialogue). Of course, it is difficult to say exactly what is "expected" dialogue, or "the norm", as this is a somewhat subjective matter. The difficulty is added to by the absence of any statistical data to say what "the norm" is. We cannot turn to any data in order to look up what is the norm for a given situation. This term is, therefore, used from an intuitive point of view, in that standard dialogue is defined here as dialogue which occurs in a situation which the system user could expect to find his/herself and would have certain expectations about the dialogue likely to be encountered in that situation. If, for example, we attend a job interview, we expect a certain type of dialogue to occur. Note that the content is not specifically expected, but the style and structure of the dialogue encountered, and the roles and motivations of the participants are.

1.10.1 (e) General Purpose Base System

According to Galliers and Sparck Jones [GAL93], a *system* is defined as an entire software and hardware entity. An NLP system performs a task, and any system used to carry out a task within a specific domain is classed as an application. Under Galliers and Sparck Jones' definitions, a *generic system* has the goal of performing a certain task within different domains, although this can then be tailored by the addition of domain specific resources to a variety of applications.

A *general purpose* system, on the other hand, is designed to be directly usable for any application, without the need for prior tailoring. They point out that "*general purpose systems do not exist even for any one NLP task, let alone a range of tasks.*" In addition, they remark that "*within certain limits, or on certain assumptions about the scope of language processing, generic NLP systems are essentially general purpose. i.e. they will serve language-processing needs within any task system*".

However, Smith et al [SMI94] have extended the terminology to include the term *general purpose base*, which they use to describe a system which is more than a generic system because it is not restricted to a single task type, but does not, as yet, fulfill the criteria for a general purpose machine as it is unable to be used for any task in any domain. Throughout this thesis, we shall use the definitions set by Galliers and Sparck Jones, and the definition for a general purpose base devised by Smith et al.

1.10.1 (f) Understanding

The term "understanding" can cause confusion, as the way in which a computer "understands" dialogue is not the way in which humans understand. Obviously, a computer does not have the cognitive powers or mechanisms possessed by a living human individual. Throughout the thesis, the term "understanding" shall be taken to mean those operations which enable a computer to abstract a representation from, reply to and translate a piece of natural language input.

1.11 The Logical Progression Of The Thesis

This section has been included in order give the reader an overview of the thesis structure, and to inform him/her as to where different information may be found within the document. The thesis is organised according to a plan, in which every chapter performs a specific role and function.

The extended abstract is included before, and separate from, the main body of the thesis. Its purpose is to provide the reader with a surface knowledge of the work without entering into too much detail, and before reading the thesis in its entirety. It is recommended, however, that a full and complete understanding of the work can only be gained from reading the thesis as a whole.

Chapter One deals with the methodological issues which bear relevance to this piece of work. Unless these issues are clearly expounded, and the position of the work in relation to the rest of the field made apparent, confusion may arise. In this chapter, we state what we intend to do, how we intend to do it, and how we can finally assess the success of the work.

Chapter Two introduces the reader to the problem area of natural language processing, and in particular to the area of dialogue analysis. Here we will look at what is difficult about the problem of modelling dialogue on a computer, and what makes the problem interesting and complex.

Chapter Three follows on to present ways in which other people have tackled the problem area described in Chapter Two. This chapter discusses work which is of relevance to the presented research, and is intended to provide the reader with an up to date view of the state of the art. Significant pieces of work which have influenced the field are discussed, along with other work which bears relation to the Dialogue Structure Models theory. The structural approach to dialogue analysis is discussed, followed by the semantic and finally intention based approach.

Chapter Four presents the actual theory developed during the project, and is, therefore, a crucial chapter in the thesis. The theory is described in full, and it is shown how natural English dialogue may be modeled using the hierarchical structure of Dialogue Structure Models (DSMs), Dialogue Elements (DEs), constraints, plan boxes and analysis boxes. A discussion of how the individual factors of dialogue are dealt with is included, as this is an integral part of the work.

Chapter Five describes the implementation which was constructed using the DSM theory. The design and construction of the implemented system is given at algorithm level, (although examples of the actual code are included in the appendices, not in the main body of the thesis).

Chapter Six address the evaluation issues of the work. In order to evaluate the design and implementation stage of the work, and in order for the work to be assessed in its entirety, a testing chapter has been included in the thesis. This serves to provide the reader with an indication of how the system's performance was tested and the results which were obtained. This also enables us to evaluate the work in relation to the criteria for success which were laid down in Chapter One.

Chapter Seven, the final chapter of the thesis contains the conclusions to the work, in which the contents of the documents are drawn together and summarised. The results and findings of the research are presented, with an indication of possible future work which may serve to extend the project presented for the Ph.D.. In this concluding chapter we recap on what we set out to achieve, and discuss what actually *was* achieved.

A table of references, glossary and bibliography are inserted in the thesis after the Conclusions.

Chapter Two: The Problem Area

2.1 Introduction

Having declared the methodological position of the work, the problem area in which the research is embedded may now be discussed. The aim of this chapter is to inform the reader regarding the general problems addressed by the work, and discuss the issues and difficulties involved in such a project, and why it is difficult to model natural language dialogue by computer.

2.2 The Need For Established Definitions

The two fields of discourse analysis and dialogue analysis are often merged within the literature, and all too frequently the terms dialogue/discourse are used interchangeably. However, we argue that they are two identifiably separate behavioural acts and under our definition of AI, this is an important distinction. The literature rarely makes a clear distinction between these two behaviours. This may be due to the fact that researchers within the field usually come from a wide range of academic backgrounds, and so use terminology according to their own definitions and approaches. However, as far as the field of dialogue analysis is concerned, this leaves us with a need to understand an individual's terminology before being able to appreciate what that person is trying to do or say. Unless we are completely clear and sure of what it is we are trying to model, the field can only be left in a state of confusion.

2.3 The Ad Hoc State of Dialogue Analysis

The field of dialogue analysis as it stands today is somewhat dis-jointed and fragmented. Different people have produced valuable work on different topics, but these theories appear to give accounts only of specific phenomena in selected situations. There is a distinct lack of integration between the work carried out by different people, with little or nothing to unify the various pieces of work. This has resulted in a collection of approaches to a variety of problems, leaving the field in an ad hoc state of development. Systems or theories have been developed to tackle only the special interest of the authors, and not to address the general problems of the field of dialogue analysis as a whole. For example, work has been produced on handling dialogue failures, but it is necessary to have a dialogue system working efficiently before one begins to tackle the problems of communication failure. The ad hoc state of dialogue analysis must be overcome if overall progress is to be made and the ultimate goals of the field achieved. The field requires some type of cohesion which would pull all of these ideas together into some kind of unity. Mykowiecka phrases it as follows:

> *"...Many systems are intended to lead a dialogue with their users... Until now no universal theory of dialogue has been defined, consequently we cannot describe all the possible situations and requirements of the dialogue partners. Because of this, many attempts dealing with a particular application have been undertaken."*
>
> [MYKO '91]

Obviously, work on linguistic features such as focus, anaphora and so on are extremely important, but there must be more to a dialogue system than simply these components which tackle dialogue at a local level. The obvious solution to the problem of ad-hoc working would be to construct a system which would provide a more general answer to the task of dialogue processing, but which could also incorporate work developed on the linguistic features of dialogue. This is vital in order to eliminate the "ad hoc" nature of the field, integrating work already conducted in the area, so that progress can be achieved.

2.4 The Problem Of "Standard" Dialogue

Dialogue analysis is a difficult task partly because it is so difficult to even define exactly what it is we are trying to model. The need to deal with standard dialogue arises from the complexity of the dialogue analysis problem. As no existing dialogue systems can as yet handle "standard" dialogue in a sophisticated manner, the complexities involved in understanding non-standard dialogues and what makes them the exception rather than the rule are beyond the capabilities of the field in its current state.

Furthermore, the problem of actually defining what is "standard" as opposed to "non-standard" adds to the complexity of the problem. Actually providing an example of "normal" dialogue poses some difficulty. Whether or not a dialogue is normal can depend upon the given context and a number of environmental influences. Factors such as culture will influence one's perception of what is the norm. Also, the dialogue which one individual would consider "standard" within a given situation may dramatically differ from that of another individual. The "standard" dialogue is, therefore, a subjective notion and what one person claims is "standard" may not be what another person would argue. This will always pose a difficulty within the field, because people will always differ from each other in culture, personality and background. This leads us on to another problem area within dialogue analysis.

2.5 The Conflict Between Instantiation And Dialogue Structure

The number of influences affecting an individual dialogue are enormous. The participant's personality (shy, aggressive, confident, curious etc.) will affect the way in which that person interacts with others using dialogue. Likewise, personal history and background may influence the dialogue they produce and the way in which they react to or interpret dialogue. Add to this the factor of mood (if you are having a "bad day", the dialogues you have with other people may vary drastically from the sort you have when you are feeling good) and the problem of instantiation becomes increasingly complex. However, as our criteria for success stipulate that a dialogue system should behave in a "natural" manner, these factors cannot be overlooked.

On the other hand, we cannot attempt to model dialogue without assuming that there is some sort of structure behind the language, as if dialogues depend solely upon instantiation factors and no others, then we would need a modelling theory for each and every dialogue which is ever created, and the task would be too complex to even consider. As Chapter Three shows, researchers have tried a number of approaches at formalising dialogue, and while we recognise that a framework needs to be developed in which to model dialogue structure, we also realise that in order to meet our criteria for success, instantiation influences must also be accounted for. It is only in the combination of structure and individuality that we will achieve the level of flexibility and naturalness called for in the previous chapter.

2.6 The Problem Of Describing Dialogue Structure

The task of defining that underlying structure of dialogue is not simple. Discovering, defining and examining the sub-units of dialogue and the way in which they interact with each other is a problem which has been the attention of academic study for many years, and no definitive last word has yet been declared on the subject. When humans use dialogue to interact with other people in our daily lives, a large proportion of the rules they are using are hidden at the sub-conscious level. For the NLP researcher to endeavour to extract the structure behind that external language and the way in which smaller units build into larger ones, and why, is an enormously difficult task. This remains one of the largest dilemmas involved in tackling the dialogue analysis problem.

2.7 Conclusions

The field of dialogue analysis is not without its difficulties and problems, many of which stem from the infancy of the field, others from the inferently difficult nature of the subject we are trying to model, namely natural dialogue. These problems are not trivial, and cannot be solved overnight. However, by identifying and addressing these problems within research projects, it is hoped that progress will be made within the field as a whole. We now move on to consider the ways in which other people have tackled the task of dialogue analysis.

Chapter Three: Related Work

3.1 Introduction

The problem of modeling dialogue has lead to a variety of approaches being adopted by different researchers, [see the following for overviews: CAR90, MCK92a, MCK92b, LAM93] but generally it is accepted that work falls into three major categories:

1. studying the *structure* of dialogue/discourse and how constituent parts connect to build up into the whole.
2. studying the *meaning* of the utterances within the dialogue/discourse (i.e. a semantic approach).
3. studying the *intentions* of the people involved in the dialogue/discourse and their goals, (much of the work in this area concerns plan recognition and construction).

The three approaches named above are discussed in the following sections of this chapter, but before moving on it should be noted that in this chapter we refer to "dialogue" or "discourse" according to the term used by the particular author under discussion. However, it is important to bear in mind that this will not necessarily be in accordance with the definitions of dialogue/discourse given in Chapter One.

3.2 Structure-Based Approaches To Modeling Dialogue

Much of the work on dialogue analysis has been produced from the approach of using theories to analyse and model dialogue structure. Central to this approach is the assumption that such structures do indeed underlie the dialogue. The structural approach portends that in order to model dialogue/discourse, structural units within the language can and must be identified, and the way in which component segments inter-relate is of crucial importance when extracting the meaning of, or generating, natural language. The markers which indicate structural segments are often not explicit, and a great deal of the work produced using this approach concentrates upon schemes for recognising such segments and then representing them in some way.

3.2.1 Discourse Representation Theory

One of the first attempts at formalising a structure-based approach to the problem is found in Kamp's work [KAM81]. Kamp claims that formal semantics of natural language have been dominated for a long time by two conceptions of meaning: the first conception is of seeing meaning as "that which determines conditions of truth.'' (page 277,[KAM81]). This is the view of meaning mostly adopted by philosophers and logicians; the second conception, according to Kamp, is "that which a language user grasps when he understands the words he hears or reads." (page 277,[KAM81]). Kamp argues that because these two views of meaning had been kept separate for so long, they had become an obstacle to the development of semantic theory. He regards his work on Discourse Representation Theory (DRT) as an attempt to overcome this obstacle.

Central to DRT is the notion of Discourse Representations (DR), which are described by Kamp as the mental representations which a person forms as a response to the input he/she receives. The DRs are formed in response to the discourses they represent, and their formation is governed by certain rules. According to Kamp, these rules operate on the syntactic structures of the discourse sentences, and it is by means of these rules that syntactic form determines the nature of the DRs that are produced. The notion of DRs is illustrated by the sentences: "Pedro owns Chiquita. He beats her." (page 284,[KAM81]). Symbols are assigned to the two elements contained in the first sentence, hence Pedro and Chiquita become 'u' and 'v' respectively. A DR is then produced as follows:

1. Pedro owns Chiquita.

u = Pedro, v = Chiquita

u owns v

2. He beats her.

u beats her

u beats v

page 284[KAM81]

Kamp recognises that the natural interpretation for 'he' as referring to Pedro is only natural because one knows that the name Pedro usually refers to male individuals. He also adds that often humans cannot do any more than to guess whether a referent is male or female, but some guesses are more "educated" than others, although no advice is offered on how to incorporate such knowledge.

Some sentences give rise to Discourse Representation Structure, called DRS for short. Sometimes a representation will involve structured families of DRs, (often representations of conditional sentences), and then this structured grouping will be called a DRS. Kamp claims that each sentence or discourse induces the construction of such a DRS, and only if the sentence or discourse is fairly simple will the DRS comprise of only one DR. He adds that among the DRs within a DRS there is always one which represents the discourse as a whole.

Even though this work is significant because it was one of the first attempts to take account of discourse structure, the theory is extremely limited. For example, the theory takes no account of motivation, intention or goals. Kamp's theory breaks the discourse sentences down into a series of relations which attempts to resolve such linguistic phenomena as anaphora, but this does not reflect the rich complexities of discourse or dialogue which are now generally accepted to be of importance if one intends to construct an efficient natural language system. These complexities include goals and intentions. Kamp's "theory" is really only a "notation".

Kamp's hierarchical approach to analysing and ordering the structure of dialogue is present in other pieces of work with the same aim. One example is that of the work of Polanyi and Scha, which we now move on to consider.

3.2.2 The Work Of Polanyi and Scha

Polanyi and Scha [SCH84, SCH88a, POL86, POL88], for example, also put forward a theory (the Linguistic Discourse Model) which builds discourse out of smaller building blocks. Polanyi and Schas' work uses a formal grammar in conjunction with "world" knowledge (semantics and pragmatics). Within this theory, combinations of clauses form "Discourse Constituent Units" (DCUs), which in turn combine and are known as "Discourse Units" (DUs), which again combine to form events within the discourse. Discourse is constructed bottom-up, building from the smallest unit to the largest, with each sentence being analysed individually before being related to the overall discourse. However, Polanyi and Schas' work does not lend itself directly to an implementation, although it does deal with issues such as topic selection and internal context.

Polanyi and Scha's work is important in that it treats discourse as a structured phenomena rather than some random entity. They use this structure to illustrate how certain topics can be more likely to be brought into the discourse than others. In addition, the notion that dialogue understanding necessitates more considerations than simply surface factors, and requires, for instance, contextual, social and extra-linguistic analysis, is an important contention.

However, Polanyi and Scha do concede [SCH88a] that their model only adopts a bottom-up approach, lacking consideration of speaker's intentions despite the fact that the model relates to goal-oriented discourse units.

3.2.3 The Work Of Grosz And Sidner

Another theory of discourse structure is presented by Grosz and Sidner [GRO86], whose model of discourse consists of three major strands:

1. Linguistic structure. Relates to the way in which strings of utterances form larger discourse segments, which then in turn build up into the total discourse.

2. Intentional structure. Relating to the purposes behind the segments and how these purposes are connected.

3. Attentional state. Relates to the focus of attention at a given point in the dialogue; a stack of discourse segments is derived from the linguistic structure, in which every item in the stack is a list of objects in focus within that particular discourse segment. It forms a representation of the purposes, propositions and intentions which relate to the current point in the dialogue. (Note that this relates to the focal point of the discourse, and not necessarily of the participant(s)).

Certain cue phrases are indications of segment boundaries (e.g. "On the other hand", "Moving swiftly on" etc.) and also convey information about intentional structure. When a segment boundary is crossed and a speaker moves from one segment to another, the cue phrases can be argued as indicating a switch in speaker *attention* or speaker *intention.* Along with cue phrases, another indication of segmentary boundaries is the Discourse Segment Purpose (DSP), one of which is associated with each segment. The DSP of a segment dictates its role and importance within the overall discourse, and concerns the intention of that segment. The discourse itself has a purpose (Discourse Purpose or "DP"), and recognition of the DSPs or DP is crucial to the fulfillment of the intention. Furthermore, the *Dominance Hierarchy* imposes a level of order upon the DSPs, which exists when an action satisfies one intention (e.g. DSP1 is intended to partially satisfy DSP2). A second level of organisation is known as the *satisfaction-precedence,* which dictates when a particular DSP must be satisfied before another.

Although the theory is supposedly three-pronged, the major driving forces of the theory come from the notions of linguistic structure. Lambert [LAM93] complains that while the concepts of DSP relations may provide a framework for representing discourse structure, they do not enable the understanding of that discourse, because the information derived from the theory is insufficient. Objections may be raised to the principle of representing discourse as a stack (which suggests that segments and topics are "pushed" onto the stack and "popped" off again after discussion), as it would intuitively appear that dialogues are not always structured in this fashion. There is often the criticism that the work disregards meaning and that interest in the structural and linguistic features of the dialogue overshadow concern with intention. The concern with topic and focus are central to the work by Grosz and Sidner, and while this important and influential piece of work can be seen as a milestone in discourse/dialogue analysis research, it is argued that the main thrust of their ideas on focus and topic rely too heavily upon syntactic features, and that the work is thus too restricted to deal with dialogue/discourse which is not as strictly structured as their theories need them to be. The work by the following author, Rachael Reichman [REI78, REI85, REI89a, REI89b], builds upon the notion of focus, but concentrates more heavily upon the relationship between the segments of focus rather than the recognition of focus boundaries and shifts.

3.2.4 The Work Of Reichman

Reichman's work argues that dialogue participants maintain and update a model of the dialogue as they are interacting. In updating the model, participants use a set of implicit rules which enable them to understand what is being said within the dialogue, and Reichman claims that any machine capable of engaging in dialogue communication with a human agent must be able to use the same rules. Reichman's theory of Context Space approaches this problem by decomposing dialogue into segments known as *context spaces*, and then examining the way in which utterances relate to one another, how one particular utterance may relate to its predecessor, and also the types of utterance which may legitimately follow it. So, for Reichman, discourse structure can be modelled by the recognition of context spaces and how they relate one to another:

> *"Discourse participants "package" pieces of discourse into these separate units and selectively bring these units in and out of the foreground of attention...the structure of a discourse can be specified by the identification of its context spaces and the relations between them."*
>
> page 24[REI85]

In processing context spaces, a number of cues are used, and discourse is decomposed into a hierarchical framework. The ten types of context space are categorised into four main groups:

1. **Issue spaces:** these play an important role within the work, and act as topic setters. These are the independent constituents which influence the generation and interpretation of all subordinate context spaces.
2. **Non-issue spaces:** This category of context spaces encompasses the following context space types:
    - **Comment and Narrative Context spaces:** A comment context space is a "temporary subconstituent of another kind of space that is currently active",

      page 58 [REI85]. Narrative spaces concern story telling.
    - **Support Context Spaces:** These are further sub-divided into two major types of support context spaces: **narrative-support** and **nonnarrative-support** spaces.

The narrative-support spaces concerns the telling of a story in which one event, built up to by preceding told events, is an instance of the claim which is being supported. The nonnarrative-support spaces relate to the description of a particular state of affairs which, using some generic principle of support along with some rule(s) of inference, leads to the belief of the claim being supported.

Inside each context space, Reichman defines a fixed number of slots which carry different types of information. All slots must be filled, with possible values for the slot being dependent upon the type of the space in question. The slots are as follows:

- **TYPE**: specifies the category name of the context space.
- **DERIVATION**: specifies whether the substantive claims within the context space are explicitly stated by the speaker or inferred by the grammar.
- **GOAL**: specifies the function served (the conversational move performed) by the context space
- **CONTEXTUAL-FUNCTION**: this slot is structured into two sub-components:
    1. **METHOD**: specifies the particular method used to achieve the goal (e.g. analogy, Modus-Ponens, flat denial etc.)
    2. **CORRELATOR**: this slot contains specific reference to the context space in relation to which the conversational move is being made.
- **SPEAKERS**: consists of a list of persons who have made utterances within the context space.
- **STATUS**: Specifies whether the context space is in the foreground/background at the current point in the discourse.
- **FOCUS**: specifies the influential status of individual elements within the context space. (Further sub-divided into categories of High, Medium, Low and Zero).

The context space grammar used by Reichman has been criticised for claiming that context spaces contain goals, just as Grosz and Sidner claim that focusing structures contain purposes. (Mc Kevitt et al argue that intentions are not substructures, but are, indeed, at least equal to topic in importance, with discourse structure aiding the revelation of this in some way or another [MCK92]). Reichman's earlier work ([REI78, REI84]), on the other hand, has been criticised for providing no mechanism for the recognition of a speaker's goal and intentions [LAM93].

Further criticisms have been levelled at Reichman's work for not fully detailing how certain tasks may be achieved. For example, Reichman says:

> *"In addition to the context space types we have so far described, other discourse phenomena - such as descriptions, small talk, question-answer pairs - may ultimately be incorporated into context space structures for the purposes of discourse analysis"*
>
> page 62, [REI85]

Unfortunately, this is not elaborated upon, but dropped into the work as an after-thought with no accompanying explanation.

Also, Lambert [LAM93] complains that Reichman's work "fails to provide a computational model that can be implemented and tested as a working system". This lack of "hard evidence" is, as we claimed in Chapter Two, a major plague of the discourse/dialogue analysis field.

Other more recent work on dialogue structure can be found. For example, Yamashita and Mizoguichi [YAM93] present a method of predicting a user's next utterance within a spoken dialogue based upon their theory of two types of dialogue model. The first dialogue model, the SR-plan, is described as a low-level model which captures the basic structure of the dialogue, and works along the principle that in a goal-oriented dialogue a stimuli is followed by a response. The second dialogue model, the TPN (Topic Packet Network), is introduced with the intention of improving the level of utterance prediction produced by the SR-plans. This TPN model uses the assumptions that the relationship between a topic and its descendant topics, (descendants are simply more detailed elaborations of the general topic), is often independent of the actual dialogue itself. The authors examine this relationship, representing dialogue is a hierarchical fashion comprising of an initial "rough" topic, which becomes focused into more detailed ones, and can return to previous topics in either an explicit or implicit fashion. Thus, the work draws together two types of dialogue structure models, involving topic and utterance pairs. While some degree of success is achieved by this approach, the authors state that it is impossible to identify a topic without the aid of "bottom-up" information which is derived from the word lattice produced by their speech recognition system. Keywords are associated to each topic, and if a keyword is found to match one found in the lattice, then the "score" of that topic rises, meaning that the topic is a more likely candidate for selection in the next utterance. Thus, Yamashita and Mizoguichi use the imposition of dialogue structure to aid in the prediction of utterances in spoken dialogues.

The purely structural approach is now largely considered as only one aspect to consider in the understanding of natural dialogue/discourse. While an important factor, the strictly structural analysis in isolation is not wide-reaching enough to encompass the rich complexities of naturally occurring dialogues/discourse. For this reason, other ways of modelling dialogues have been considered. We now move on to discuss some of the work conducted from a largely semantic point of view.

3.3 Semantic-Based Approaches To Modeling Dialogue

The semantic-based approach to modelling dialogue concentrates on the meaning of what is said as opposed to how it is said or why. The representation of semantics has been attempted using propositional logic, semantic networks and other formal approaches such as Conceptual Dependency and Preference Semantics. In this section, we will look at the two main works which have influenced work in this area to the greatest extent. Firstly, let us consider the work of Yorick Wilks.

3.3.1 Wilks' Preference Semantics

The Preference Semantics approach developed by Wilks [WIL73,WIL75a, WIL75b, WIL75c] was originally used for machine translation. A bottom-up approach is adopted, as Wilks constructs a semantic analysis of the overall discourse from the semantics of individual sentences within it. The discourse is represented in terms of semantic items, (Wilks uses paragraph sized chunks of text), and is known as a Semantic Block. This is made up of Templates (relating to clauses or easily recogniseable surface items). Templates are related to one another via "paraplates" and "common sense inferences" made up of formulas (which express word sense, one formula corresponding to one word sense). Formulas are constructed out of "elements", which form the building blocks of the discourse. These elements represent basic semantic concepts such as entities (e.g. MAN), actions (e.g. FORCE) and cases (e.g. TO).

In later work, [WIL75c], the concept of Pseudo-Texts is introduced, which can be seen as similar to Schank and Abelsons' frames [SCH77]. A psuedo-Text is, in essence, a semantic block, and is constructed from a linear sequence of templates. These templates are connected to each other by means of case ties. Consider the Pseudo-Text for "car":

[(MAN)	inject	liquid	]	
[MAN)	(USE)	#tube	]	insertion of fuel
[ICengine	(USE)	#liquid	]	petrol moving engine
[*	(MOVE)	[]	]	moves the car

Within this Pseudo-Text, a NLP system can attempt to find the template in it which matches most closely the template of the source. Thus, an understanding can be derived for a variety of natural language expressions, including those with such rich meanings as "My car drinks gasoline". In this case, Wilks' use of *projection* replaces "use" with "drink" within the formula for "drink".

3.3.2 Schank's Conceptual Dependency

The use of basic semantic units for representing and constructing the semantics of discourse is an approach also adopted by Schank within Conceptual Dependency [SCH72, SCH77]. The main axioms of the theory are as follows:

1. Where two sentences have identical meanings, (irrespective of language), there should only be one representation.
2. Any implicit information contained in a sentence must be made explicit in the representation of the meaning of that sentence.

Schank claims that these two axioms have necessitated the use of an economical form for representing meaning. In doing this, he claims, he has produced the following framework:

1. The meaning propositions which underlie language are called conceptualizations. These conceptualisations may be active or stative.
2. An active conceptualisation has the form:
   Actor Action Object Direction (Instrument).
3. A Stative conceptualisation has the form:
   Object (is in) State (with Value).

This form has lead to the principle of primitive actions. Because an object is defined as an actor doing something to an object in a direction, Schank argues it is necessary to determine exactly what it is an actor can do. A primitive act is language independent, and provides the basis for representing the meaning of actions. Because sentences with identical meanings should be given the same representations, Schank says that it is necessary to examine the actions which seem similar to see if it is possible to extract the essence of their similarity.

Schank also claims that the second axiom makes it necessary to make explicit whatever differences there may be between two actions and to express them accordingly. He uses the example of the two verbs 'give' and 'take' which both share a similar primitive element, (transfer of possession), but they are also different. The best representation for a verb, he argues, is the primitive element it has in common with other verbs, plus the explicitly stated concepts which make it unique. He states that:

> *...often these explicitly stated concepts turn out to share similar elements with other verbs, which means that a verb is represented as a particular combination of primitive actions and states, none of which are unique to that verb, but whose combination is entirely unique.*
>
> page 12 [SCH77]

In a nutshell, Schank's work calls for the translation of different natural language sentences into an internal representation, which is known as the Conceptual dependency. This is achieved by the use of the set of primitives, which Schank argues can be used to represent all natural language actions. Schank also accounts for causality and the influence one sentence has upon the subsequent ones by means of rules for determining the causal relationship between events. From observations of the physical world, Schank arrives at two causal syntax rules:

1) Actions can result in state changes.
2) States can enable actions.

Associated with every primitive action within the CD theory is a set of states which it may affect, along with a set of states which are necessary in order to effect it. Thus, Schank claims that by using the causal rules, it is possible to build up a representation for a complete multi-sentence piece of natural language.

Criticisms of the Conceptual Dependency theory have mostly centred upon Schank's concept of the world as being condensible to a finite, relatively small set of primitive actions. It can be argued that the world, and indeed natural language, is far more complex than that and defies being explained in the simple terms that Schank argues for. While this is probably true, the dialogue analysis field owes a great deal to the

Conceptual Dependency theory because of its (then) innovative approach to representing natural language. We now continue to look at another piece of Schank's work.

3.4 Intention-Based Approaches To Modeling Dialogue

Another school of NLP researchers approach the task of modelling dialogue using participant motivation or *intention* as a centrally cohesive factor. This line of thought recognises that whenever a dialogue participant makes an utterance, they do so with some purpose in mind, and subscribers to this approach view intention as the main factor which under-pins the coherence of a dialogue. That is to say that the individual segments of a dialogue are created and inter-related due to the participant's intention. Work of this nature frequently refers to the goals, plans or beliefs of the speaker.

3.4.1 The Work of Austin

One of the first and most significant pieces of work to be included in this category is that of Austin, [AUS62], who argued that not all utterances can be explained in terms of "true" or "false". He calls some utterances "performatives" because they change the state of the world when created under certain circumstances (e.g. "I guarantee that this will work"), and actually *perform* some action. He then goes on to argue that in making utterances, dialogue participants perform three types of act:

- **locutionary act:** actually making the utterance

- **illocutionary act:** an act which is performed via the process of making the locutionary act e.g. informing someone of something. (These are also referred to as "speech acts" within the literature).
- **perlocutionary act:** an act which is the *result* of making that utterance (the effect it produces upon the other participant(s)). E.g. making a threat might scare the other dialogue participant, whether fear was actually intended to be an effect or not.

It is the illocutionary act (or "speech act") which is more frequently the subject of interest within the area of dialogue analysis. A fundamental aspect of the work concerning speech acts is the focus on the functionality of natural language, and the contrast between the illocutionary force of an utterance and its literal content.

3.4.2 The Work Of Grice

Another significant advocate for the approach of using intention to model dialogue is Grice, [GRI57, GRI69, GRI75]. Grice [GRI57] follows the opinion that dialogue is, above all, goal-oriented, and that the true understanding of a dialogue can only be achieved by the correct recognition and interpretation of these participant goals. Grice defines meaning in terms of intention, arguing that when an utterance is made, the speaker intends that utterance to cause a certain effect upon the listener, and that this is achieved because the listener recognises the original intention. Later, [GRI75], Grice presents his notion of the "Cooperative Principle", which comprises of four sub-sections which he calls "maxims". The Cooperative Principle relies on the assumptions that dialogue participants are collaborating within the interaction, and that the maxims are not being violated. The maxims are of **Quantity** (concerning the degree of information conveyed by a participant), **Quality** (concerning the moral truth of what is said), **Relation** (concerning the relevance of what a speaker says to the dialogue as a whole), and Manner (concerning the way in which the speaker contributes to the dialogue, declaring the need to avoid obscurity and ambiguity).

3.4.3 The Work Of Searle

A third significant contribution to the intention-based approach is that of Searle [SEA69] who produced work combining and expanding on features of the work carried out by Grice and Austin. He uses Grice's ideas about participant intentions and these are attributed to psychological motivations, whereas Austin's notions of intentions have their roots in social conventions and organisations. Searle's work extends the ideas initiated by Austin, further detailing the illocutionary act into the five sub-branches of Assertives, Directives, Commissives, Expressives and Declarations. These are an attempt to expand upon the types of utterances capable of being explained by Austin's speech acts, which he sees as too restricted and not comprehensive enough. As a result of these expansions, Searle takes account of a wider range of utterances than simply those containing straightforward performatives. But neither is Searle completely satisfied with Grice's notions about meaning, which he views as misplaced and restrictive. Objections to Grice's definitions of meaning lead to Searle's combination of Austin's and Grice's work, using the conventionality produced by Austin's speech acts with Grice's notions of intention. In order for an utterance to be classed as having meaning according to Grice, the hearer must interpret the speaker's utterance exactly as it was intended. However, this is not always the case, and utterances can be interpreted or misunderstood in a variety of ways. For Searle, an utterance is classed as having meaning as long as the hearer indicates to the speaker that the utterance has been received. Whether the original intention is the meaning derived from the utterance by the hearer is irrelevant to Searle's pre-requisite for utterance meaning.

3.4.4 The Work Of Cohen, Perrault And Allen

The problem of ambiguity in interpreting speaker intentions mentioned in the previous section is the subject of much work. Cohen, Perrault and Allen [COH82] deal with this issue, developing ideas about inferences and expectations relating to intention interpretation.

By conducting experiments on users interacting with a natural language system called PLANES, Cohen, Perrault and Allen were able to draw a number of conclusions about the way in which dialogue participants engage in a question-answer interaction, what they expect from it and the inferences they draw. This work (and likewise many other projects conducted within the field) presents an overlap in the intention-based approach and the plan-based approach to modeling dialogue, because both intentions and goals are considered as inter-related. Cohen et al argue that speakers plan certain effects upon the hearer(s) to be the result of an utterance, and that a hearer achieves these effects by reasoning. Any good natural language dialogue system should, according to Cohen et al, be capable of reasoning about the user's intentions, and use plans to fulfill and cooperate with these intentions. The use of plans falls into two stages:

1. Plan construction. This involves predicting the outcome to events without having to actually perform them. In this way, a participant can attempt to foresee the effects of making an utterance before he/she makes it.
2. Plan recognition. This concerns the prediction of events resulting from actions executed by another participant (i.e. making inferences concerning plans being followed by the other participant). This process involves making assumptions or judgments about the participant's beliefs, the behaviour which is likely or expected from that participant, and certainties which occur as a result of a given action.

Work presented by Cohen, Perrault and Allen, therefore, discusses three basic aspects of the plan-based approach to modeling dialogue: reasoning about the actions and utterances of the other participants, planning one's own utterances in order to achieve goals inferred about the other participant, and also to recognise individual utterances as forming part of the other participant's overall plans. For Cohen, Perrault and Allen, intentions and the plans inferred from those intentions are the important aspects of dialogue coherence.

3.4.5 The Work Of Allen

James Allen [ALL79, ALL80] based his work on that of Austin and Searle, treating language as a sequence of actions. Combined with this, Allen's work is also influenced by Grice and Searles' goal-oriented approach to discourse processing, applying a planning formalism. Allen develops schemas for dialogues originating within the domain of boarding and meeting trains, and also for speech acts such as "request" and "inform". Allen uses a set of plan inference rules to infer goals from participants' speech acts. These rules are set out below:

1. Precondition-Action Rule: If an agent's goal forms a precondition of a certain action, then it may be the case that the agent desires that action.
2. Body-Action Rule: If an agent wants the body of some action, then it may be the case that the agent wants that action.
3. Action-Effect Rule: If an agent wants an action, then it may be the case that the agent desires the effect of that action.
4. Want-action Rule: If an agent wants another agent to want some action, then it may be the case that the agent wants the other person to perform that action.

5. Know-Pos-neg Rule: If an agent wants to know if P, then it may be the case that the agent wants P, or perhaps the agents wants $\neg$P.
6. Know-Value Rule: If an agent wants to know if P(a) for some predicate P and some value, a, of a parameter, then maybe the agent wants to know the value of that parameter of P.
7. Know-term Rule. If an agent wants to know the referent of x in D(x), then perhaps that agent wants P(the x:D(x)) for some predicate P.

Three main assumptions underlie the work that Allen has developed. Firstly, there is an assumption that human beings act as rational dialogue agents, capable of forming and also executing plans which serve to fulfill their goals. Secondly, that human dialogue agents often infer the plans and goals of other dialogue agents by observing the actions of that other agent. Thirdly and lastly, there is the assumption that human agents are able to detect obstacles which stand in the way of other agents' goals becoming achieved. Allen uses plan inferencing not only to detect these obstacles, but to also provide responses which will help to remove those obstacles. This aspect of "helpful" dialogue is central to Allen's work, which offers a mechanism for the generation of responses which convey more information than was explicitly asked for; consider the example:

(1) **a** patron: when does the Montreal train leave?
b clerk: 3.25 at gate 7.

Although much of Allen's work deals only with single utterance interactions as opposed to multi-sentence formations, his work is of significance because of the influential impact it has had upon subsequent thought and work. The ability to handle both direct and indirect speech acts and the ability to allow of an excess of information to be imparted in addition to that requested, makes Allen's work an important milestone in the history of dialogue analysis.

3.4.6 The Work Of McKevitt, Partridge and Wilks

The work of McKevitt, Partridge and Wilks [MCK92a, MCK92b, MCK92c] centralises around the development of the OSCON system (Operating System CONsultant) which acts as a question-answer system which human participants may query with regard to computer operating systems. OSCON's aim is to understand the questions put to it by the human participant, and respond using natural language utterances. Note that the natural language used is English, and the authors take the term "utterance" to refer to any unit of language, such as a word, phrase, sentence or exclamation.

McKevitt et al are very much concerned with the *coherence* of discourse, which they claim is determined by the coherence of participants' intentions. The analysis of coherent discourse for McKevitt, Partridge and Wilks involves examining the way in which individual utterances combine and inter-relate in order to build up an overall dialogue. The central thrust of the work is the claim that coherent dialogue can be modelled as a result of analysing sequences of intention. As a result, there follow two main tenets:

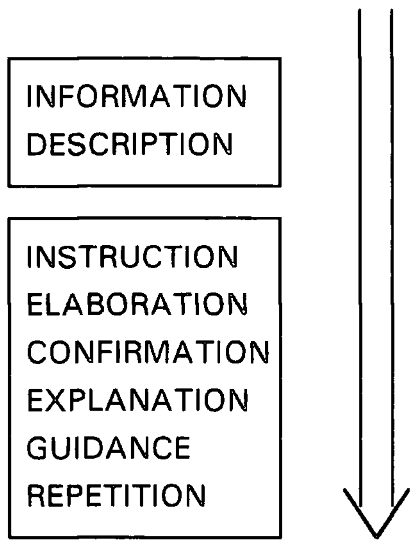
1. That intention is, in fact, recogniseable
2. That the representation of intention is a feasible objective.

On the analysis of intentions, (intentions may also be known as "speech acts" or "communicative acts"), it is argued that natural language dialogues normally consist of utterances "in a specific temporal order" [MCK92b] with each utterance reflecting some intention. So, if dialogues consist of sequences of utterances, those dialogues may be represented as sequences of intentions. It is noted that there may be a "many-many" relationship between utterances and intentions, where a given utterance may be the product of more than one intention, and similarly, many utterances may be the manifestation of a single intention. McKevitt et al argue that in order to recognise the intention behind natural language utterances, syntactic, semantic and pragmatic knowledge must play an influential role in that analysis.

In order to make the task of intention recognition easier, the authors categorise intention into types which could be considered expected within a given domain, in which the user is trying to perform a task. These types are defined as relational to the user's plans and goals, a goal being " some state of affairs which a user wants to achieve", and a plan being "a set of actions which need to be executed in order to obtain some goal". The domain which the authors are most interested in is that of consultancy, as the computational model they develop in order to test the theory is consists of the computer system acting as a consultant expert which a user may interrogate regarding operating systems. While McKevitt et al point out that it may be impossible to identify all the intention types within a given domain, they do offer us some intention types which may facilitate in the intention analysis task.

Intention	Definition
information	An intention requesting a PLAN to achieve a specific goal, where the goal is described.
description	An intention requesting a description of a concept or object
instruction	An intention acting as an instruction to achieve a goal, rather than how to achieve the plan to achieve that goal
elaboration	An intention requesting more information on a plan or goal
confirmation	An intention requesting confirmation of a belief, or some plan believed to execute some goal.
explanation	An intention requesting explanation or clarification of an item which occurred during the execution of a plan for a goal.
guidance	An intention requesting a plan for a goal where there is no explicit goal expressed.
repetition	An intention which is a repeated request.
nointention	An intention which is not immediately relevant to the domain, or not understood by the hearer as being relevant to the domain

Having grouped intentions into types within a domain, the authors then move on to consider the ordering of intentions, stating that as their interest lies in the dialogue interaction between computer and human, where the computer is the agent supplying information to the human, the satisfaction of the user may be a meaningful ordering of intentions. Thus, the human user may indicate through intentions the level of ease with which (s)he is understanding what the computer is saying. Influencing factors such as the interest which the human user has in the dialogue will affect the ordering of intentions. The ordering of intentions is calculated upon the intuitive basis of the type of requests human participants make when understanding or not understanding what is being said. The authors quote the example of asking straightforward information requests: a human participant will do this if (s)he is experiencing no difficulty understanding the discussion topic. However, if more difficulty is experienced, elaborative, explanatory or confirmative requests for information may be given.



Satisfaction ordering of intention [MCK92b]

Having thus recognised and ordered intention, the representation of intentions then comes under consideration. Here, the authors are interested in how the intentions combine to construct representations of intention sequences within a dialogue. At the "smallest" end of the scale, an utterance consisting of a single intention can be useful in the analysis of user intention, but more information is gleaned, it is argued, by examining intention pairs (e.g. an *information* intention type followed by an *explanation* type). The representation of intention pairs is achieved by the use of intention graphs, which represent the links between pairs of intentions. The intention graphs allow a variety of information to be inferred about intention:

1. The complete set of intention types modelled within a dialogue
2. The local temporal relations of intention pairs
3. The frequencies with which each of the intention pairs in 2. above occur.

The OSCON system draws upon the intention analysis work described above, and the tasks of intention recognition and intention representation are handled by different components within the system. As well as enabling a question-answer type dialogue to be conducted between the machine and the human user, the computational model is also used as a means of user modelling i.e. the modelling of a user's level of expertise within some domain. Experiments and their results are given relating to the testing of what the authors call the Intention-Computer hypothesis, which states that by analysing intention within natural language dialogues we facilitate natural language dialogue communication between a variety of types of people and the computer. The Intention-Person hypothesis, on the other hand, states that the analysis of intention found in natural language informs us about the level of expertise held by different persons.

The work produced by McKevitt et al is refreshing in that it includes a computational model which is used as a means of testing the theoretical aspects of their work, and that their findings and methodology are fully presented. They relate their work, to some extent, to the work of Schank and Wilks which uses the concept of decomposing dialogues into primitive units. The primitive units dealt with here is that of intention primitives, although the authors make no claim as to the completeness of the primitive set, suggesting that the set of intentions they define in their work may only form a minor fragment of the potential set of intentions occurring within the consultancy domain.

The work can also be related to structure-based work. The authors suggest that there may be some relation between intention analysis and discourse segmentation, in which intention analysis could possibly aid in the identification of discourse segment boundaries. Furthermore, the work of Reichman discussed earlier may be related to the intention analysis work by viewing Reichman's notion of conversational moves in terms of intention types.

3.4.7 Scripts, Plans And Goals

Schank and Abelson [SCH77] have produced work useful in story understanding. The notion of "scripts" was created for dealing with a very stereotyped sequence of events and is described by the authors as:

> *...a structure that describes appropriate sequences of events in a particular context...a predetermined, stereotyped sequence of actions that defines a well-known situation.*
>
> [page 41][SCH77]

The most well-known, "classic" situation which Schank and Abelson use as an example is that of going into a restaurant to eat, arguing that when we do so we know that there will be a waiter/ess, a menu, a chef, food and so on. We also know that a certain sequence of events is likely to take place, such as sitting at a table, being shown a menu, ordering food, eating, paying the bill etc. Sometimes there may be situations which are not so typical, but then there will be sub-sequences of events/actions which are standard. For example, if a robbery occurred at the restaurant, this would be highly unexpected, but there would still be a general plan to the robbery, such as:

- The robbers enter the restaurant

- They point guns
- They demand money
- They threaten to injure/kill if they don't receive money
- They ensure a safe getaway from the restaurant

The first part of a script declares who the actors in the actions are, and also which objects may participate. In the restaurant script, the actors are the customer, waiter/ess, the chef etc. The objects taking part are the dining room, the tables, the chairs, the menu, the cutlery and so on. In a text, the script's actors and objects take the definite article even when mentioned for the first time. The authors argue that as soon as someone specifies which situation they are entering, these objects and actors already exist in our minds. For example, after being told that someone went to a restaurant yesterday, we have no difficulty understanding what is meant when the notions of 'the menu' or 'the waitress' etc. are introduced.

The second part of a script is the *releaser*, which defines the necessary criteria to trigger a particular script into motion. These are rules for activating scripts, and described as "dependent on certain key concepts or conceptualisations when found in certain contexts" ([SCH77], p. 48). In other words, the restaurant context should not be initialised simply because a sentence refers to a restaurant.

A different script is triggered according to the role of the actor. (It is worth noting that the same point is reflected in the theory presented in this document, as Dialogue Structure Models vary according to the role or "direction" of the dialogue participant. This concept is dealt with in Chapter Four, when the theory is discussed in detail).

The rules which govern the decision to invoke a script or not are categorised into four classes called *headers*. These are grouped according to how strongly they predict that the associated context will be instantiated.

1. "Precondition Header" (PH) makes the weakest predictions and is so called because the script is activated when a main script precondition is mentioned in the text. For example, if the text says that an actor is hungry, this would be a PH for the restaurant script, as it is normally assumed to be true within the script that people who go into restaurants as clients are hungry.

2. "Instrumental Header" (IH) makes stronger predictions about the associated context than the PH. An IH normally comes into play when the input refers to two or more contexts, of which at least one may be seen as an 'instrumental' for the others. Schank and Abelson use the sentence, "John took the subway to the restaurant" as an example, because following input concerning either the subway or the restaurant would make sense. It is more likely that the subsequent input will be concerning the restaurant, and the subway would be viewed as an instrumental means of reaching the location in which more important script goals will be expected.

3. "Locale Headers" (LH) produce stronger predictions still, and arise from the notion of a time-place locale for situations. Many situations are known to take place within a certain location, which is defined as the 'residence'. When someone is said to be near, or more importantly *inside*, such a residence it may be fairly accurate to predict that a certain script will be activated.

4. "Internal Conceptualization Header" (ICH) makes the strongest prediction of all. The authors claim that any conceptualisation or role from a script may occur in a

script, and that sometimes it will activate that script, but at other times it will not. The most obvious cases of these alternatives occur when a role name, (e.g. waitress.), is used in the locale of the role or away from the role, (e.g. "My cousin is a waitress").

"Interferences" and "Distractions" occur when a script has been called into play and then a sentence is encountered which does not have any relation to that script. The only way to tell if a new topic, and therefore script, is being introduced is to wait and see what comes next. Schank and Abelson claim that it is unusual for people to intentionally mislead others, and that if something "non-standard" occurs, then it will probably be mentioned explicitly. Interferences are defined as "states or actions which prevent the normal continuation of a script" ([SCH77], page 52), and are broken down into two smaller sub-divisions: obstacles, (where some enabling condition for an impeding action is missing), and errors, (where an action is completed with a result which is unexpected and inappropriate). If an actor encounters an obstacle, he/she may react by taking corrective action in an attempt to produce the missing enabling condition. These corrective measures are called "prescriptions". On the other hand, the actor may simply exit from the scene, either immediately after the obstacle, or after one or more prescriptions have failed.

Schank states that every act within any script is susceptible to obstacles and errors, and each will suggest its own appropriate prescriptions or loops. If this happens often enough, then someone who frequently experiences this script will learn these prescriptions or loops as he/she learns the script.

The second type of events which may cause detours or abrupt script endings, 'Distractions', are defined as unexpected states or actions which cause the actor to reevaluate his/her goals, causing him/her to leave the script permanently or temporarily. Distractions can be caused by many possible sources, are not tied to any particular script, and it is possible for an event to present both an interference and a distraction. If the distracting events run their course within the location of the script, then Schank and Abelson argue that it is likely that control will return to the script at the place where the interruption occurred.

Scripts are categorised into different types. The easiest script to deal with is the situational script, the type described so far, in which the situation is specified, the actors have connecting roles to play, and they all share an understanding of what is expected. However, if one party wishes to channel the events along a different path, a new situation arises. This actor may have a Personal Script which he/she is following, and this type of script does not behave in the same stylised manner as the situational scripts. It may be that not all of the participants in a Personal Script are aware of their involvement. For example, someone may not know if he/she is being seduced or swindled.

Another type of script is the Instrumental Script, which is similar to the Situational Script in structure, because they work with a predefined series of events. The difference is in the types of actions involved, the variability of the ordering and the use of the script in understanding. Examples provided include starting a car, frying an egg and working a keypunch. It is obvious from these examples that there is little variability within the order of events, and each step must be taken in its correct place within the order. These Instrumental Scripts are numerous and normally involve only one person. The three types of script may occur together in juxtaposition, and all depend upon our shared experiences and expectations of certain situations.

Suitable uses of scripts include the completion of a system's semantic representation and also for programs which "summarise'' texts. Each action in a script can be marked to indicate whether it is essential, very important, secondary, optional, trivial etc. De Jong, [DEJ77], has attempted to develop a system which can summarise reports of accidents or disasters, based on the observation that when a car accident takes place, the condition of the victim(s) is of vital importance, whereas the colour of the car is a minor detail of little significance.

Scripts are also useful for displaying pragmatic knowledge in a way which is clear and precise. This is then easily passed on to other people who do not share the same experiences, such as people from other cultures, or children. In an unknown situation, it is useful to have knowledge about the exact sequence of events/actions which are needed to achieve the desired goal.

However, many events cannot be predicted, and so Schank and Abelson offer *plans* as a way of explaining non-stereotypical events. Like a script, a plan is a means of understanding the relationships between the various actions which serve to achieve a goal. Plans are seen as a means of connecting items of information which can then be used to explain unexpected actions. The definition of a plan is given as "(being) made up of general information about how actors achieve goals. A plan explains how a given state or event was prerequisite for, or derivative from, another state or event" ([SCH77], p. 70).

Schank and Abelson state that plans describe the set of choices which a person has when he/she sets out to achieve a goal, and that when humans listen to discourse, they use plans to make sense of seemingly disconnected utterances.

A plan is made up of goals, which in turn may call other plans into motion and each individual goal is linked to a number of plans. Schank and Abelson describe two types of goals: the first type is always achieved by the same methods, regardless of the context of the text and the other type of goal is achieved by methods which are dependent on the object with which the goal is concerned. For example, a goal which states: "prepare an object for use", could involve a variety of actions in order for the goal to be achieved.

3.4.8 Other Significant Work

This work on scripts, plans and goals led to the construction of a plan-recognition system called PAM [WIL81]. Wilensky attempts to explain the goals which a particular character may have, in the hope that this will explain the behaviour. Wilensky's plans include information relating to domain goals and meta-goals, (e.g. combining plans to achieve particular goals). As a result of this, Wilensky was able to make observations and conclusions concerning a participant's goals, using this information to understand the behaviour of that participant.

Sidner [SID81, SID83, SID85] takes the view that not only should a speaker's goals be recognised, but that in addition the hearer should recognise the intended response which the speaker desires. This is called the Intended Meaning. Information from a range of sources is required by the hearer if the intended meaning is to be received, including domain knowledge, shared beliefs about the hearer's abilities and also knowledge concerning the speaker's plan. If the speaker's plan is recognised, then all subsequent utterances are interpreted in relation to this plan. Although Sidner suggests that more than one individual utterance may be needed in order to evaluate an intention for a discourse, no clear framework is provided on how to achieve this end. Strengths of the work include, however, the consideration of contextual knowledge and participants' levels of expertise.

Other significant and/or relevant pieces of work relating to the plan-based approach include work by Lambert [LAM93] on the development of a tripartite model of dialogue which adopts a plan-based approach, which distinguishes between domain, problem-solving and discourse actions. An algorithm is presented which claims to combine world, linguistic and contextual knowledge thus enabling the recognition of complex discourse acts such as the expression of doubt. The dialogues used by Lambert are negotiation dialogues in which the participants are working cooperatively.

Carberry ([CAR87], [CAR89], [CAR90]) also presents a plan-based approach to her work, which is significant because of its ability to recognise plans in an incremental manner, and keep track of the focus of attention within those plans. A system is presented by Carberry which constructs a representation of the discourse called a context model, which represents the system's beliefs about the plans of the user. The context model presents the user's goals in a tree formation. Each goal which has been considered or explored by the user is shown as a node on the tree, with the "root" goal being the overall discourse goal, and one specific goal being nominated as the current goal of the user at the present time. The system is used for recognising pragmatically ill-formed utterances and elliptical fragments, but the algorithm cannot recognise discourse goals (i.e. what the speaker is attempting to do by saying something) for larger sections other than elliptical fragments.

Young [YOU93] presents an approach which seeks to bring together the structural and plan-based approaches. Working with Spoken Language Systems (SLS) in mind, she argues that spontaneous spoken dialogues have a "predictable structure that is defined by the properties of discourse plans and their interaction with the domain plans for an application" [p. 1169, YOU93]. It is argued that these structural properties dictate "what can occur when". Young presents a dialogue system which relies upon a knowledge base which is domain specific, and which contains information regarding the plans which can be executed in the application domain. Thus, a new knowledge base is required for each potential application. In a nutshell, the underlying notions of the system is that by keeping a track of all information communicated in the dialogue up until the current point, it is possible to infer speaker goals and plans. Furthermore, this information should also enable certain predictions to be made regarding the next utterance in terms of the type of discourse action it will be. Young provides evidence of her processing methods, giving an example of a clarification subdialog in the domain of ordering a pizza, and represents a growing number of workers who now realise that evidence and substantiation of claims are increasingly important.

3.5 Summary & Conclusions

Approaches to the analysis and modeling of dialogue have traditionally fallen into the three categories discussed above; structural, semantic and intentional. However, some recent works by such people as Lambert and Carberry have recognised that more than one aspect of interpretation must be considered. To model only the structural features of dialogue produces a very restricted interpretation, as does a purely semantic or plan-based approach. It is the author's belief that progress will be made in the dialogue analysis field only if the three "traditional" approaches are augmented and work moves in the direction of combining the strengths of each individual approach into one. Only by investigating all three aspects of natural language dialogue, (structural, semantic and intentional) will NLP systems be capable of understanding the full and complex meanings conveyed by those dialogues.

Much of the work conducted to date can only handle single utterances (e.g. [ALL79], [ALL80]). Others can only handle discourse, not dialogue, as defined in Chapter One. Another limitation of existing work is that there is often a lack of implementation presented. Even worse, there is often an absence of even an algorithm or design, or the work may be presented at such an abstract, theoretical level that it must be considered computationally infeasible. While theories have their uses and place, working implementations are essential if the field is to progress, and hard evidence proves that claims are not simply exaggerations.

A further limitation of existing work is that while systems and theories have considered the goals of one dialogue participant, none have yet encompassed the goals of *all* the participants. This is recognised by Lambert [LAM93] as "crucial" for systems which handle negotiation dialogues, as participants often discuss the acceptability of what has previously been said. We would argue further, that it is not only systems for modeling negotiation dialogues which should account for the goals of all participants, but that the roles being played by the different agents should always be recognised, as they always influence any dialogue.

A major limitation of existing dialogue systems and theories is that they are designed specifically using only one type of dialogue (e.g. negotiation dialogues, question and answer dialogues etc.). Furthermore, the domain of the work may be restricted to a point which renders any ensuing implementation useless in the real world. So many of the dialogue systems in existence today are severely limited in their over-all abilities, even though they may handle one specific aspect of dialogue quite well. The ideal system, however, would be capable of handling a wide and rich variety of dialogues, with unrestricted domains and shifting participant intentions. Lambert claims that:

> *"understanding ... dialogues requires recognising and modeling initial and changing beliefs, acceptance and non-acceptance of claims, uncertain belief in addition to belief and disbelief, and complex communicative actions such as expressing doubt. No previous model has all of these capabilities"*
>
> [LAM93, p. 22].

We have seen, therefore, that while progress has been made within the field of dialogue analysis, there exist some severe limitations to the work conducted so far. The theory presented in the next chapter, along with the computational model of that theory presented in Chapter Five, is offered as a move towards overcoming these limitations. Chapter Four also contains a discussion of the project in relation to the work discussed in this chapter.

Chapter Four: The Theory

4.1 Introduction

We now move on to the presentation of the research itself. This chapter deals with the theoretical side of the work, namely the Dialogue Structure Models theory. This work is intended as a solution to the problems discussed in Chapter Two, and aims to fulfil the criteria for success laid out in Chapter One.

Firstly, an explanation is presented of the hypotheses which underpin the theory, and the assumptions which were made when developing the work. Section 4.4 presents the theory itself, describing the constituent concepts and definitions in more depth than was presented in the Extended Abstract. The chapter is then "rounded off" and drawn to a close by the Conclusion.

4.2 Assumptions Behind The Theory

In attempting to build such a general theory of dialogue, a central hypothesis has to be made. As the work focuses largely upon dialogue structure, this assumption is that such a structure to dialogue exists, and that we, as human beings, use it every day, albeit subconsciously. Also, by applying a DSM approach to dialogue analysis, the theory assumes that people in the same situations use the same dialogue structure. These assumptions form the central hypothesis of the theory presented here.

Another assumption relevant to the theory concerns the existence and presence of supporting NLP modules which enable a dialogue system to be constructed. As dialogue analysis is operating at the top level of the NLP system, it is assumed that the "lower" building blocks are in place, e.g. that there exists a semantic analysis module, pragmatic module, a grammar, a parser and so on, which enables human-machine interaction at sentential level. The dialogue analysis module developed within this research project could not, within the time constraints, have produced a whole NLP system from scratch. It is assumed that the natural language input has been parsed by the base NLP system, and the results of its analysis passed on to the dialogue system.

Before moving on to a discussion of the theory itself, it is important to look at how we approach the problem of standard dialogue discussed in Chapter Two, section 2.4. As was pointed out, trying to define what is a "standard dialogue" is an incredibly difficult, if not impossible, task. Therefore, we do not claim that we can produce an example of a "standard" dialogue. Instead, we claim that we take dialogues which statistical evidence would suggest are expected by most people within a certain situation. Within each dialogue situation, certain social conventions and constraints are experienced by the majority of people. (We know what is acceptable in a job interview as opposed to a party, for example). Most people are aware of these constraints, and so we share certain expectations of what is acceptable or unacceptable in certain situations and dialogues. So while we do not say that we can provide a definitive example of a standard dialogue, we have based this work upon dialogues which would fall into that category.

There now follows an in-depth discussion of the theoretical aspect of the work.

4.3 The Dialogue Structure Model Theory

The remainder of this chapter deals specifically with the concept of Dialogue Structure Models, and the presentation of the theoretical side of the project. Definitions and concepts are discussed in greater detail than was given in the extended abstract. This forms the crux of the thesis. The principles set out in this chapter are then used in the design of the implementation discussed in the next chapter. These principles are presented in a hierarchical manner, with the more general level aspects of the work discussed initially, moving on to the more "low level" aspects of the theory, which relate more directly to implementation concerns.

4.3.1 Dialogue Situations

The starting point of the Dialogue Structure Model theory is the dialogue situation. Dialogue situations exist in the physical, real world and are the tangible embodiments of dialogues. For example, a lecture is a physical situation not a dialogue in itself. The language produced is the dialogue, the "lecture" (someone standing in room talking to an audience) is the dialogue situation. Thus, the dialogue situation is made up of a number of factors, only one of which is the dialogue being conducted. The physical environment of a dialogue can influence that interaction (for example, a fire alarm rings and the dialogue ends abruptly), and so it is a vital, although not sole, consideration within the dialogue analysis task.

Dialogue situations cannot be overlooked or dismissed as human beings exist and communicate in the "real" world. Schank's theory of scripts [SCH77] has shortcomings because it deals *only* with the physical world, but neither should the world around us be completely overlooked by a dialogue system as dialogues do not occur in a vacuum. Dialogues always occur for a reason and in some external context or another. Even if the situation involves simply passing the time of day, the dialogue still has a context and a motivation which affects the dialogue structure. The influence of the physical dialogue situation is reflected by the presence of the "External Elements" presented in section 4.4.3, which we shall discuss later.

4.3.2 Dialogue Structure Models

A Dialogue Structure Model (DSM) contains standard information about a stereotypical dialogue situation. However, note that within this work the stereotypical subject is the dialogue structure rather than physical, temporal events. Each DSM is a template-like schema containing basic information about the dialogue structure which may be expected to be encountered within the given dialogue situation. A DSM can be built for any dialogue situation specified, but the construction of each is identical in design with different "slots" filled in or quantified in a different manner. Basic information held within each DSM includes information about the goal of the dialogue. The details are filled into this schema by means of a unique combination of Dialogue Elements.

The DSM represents, therefore, the highest level of the theory, into which all the other constituent parts slot. The next entity within the DSM hierarchy is that of the Dialogue Element, which we shall now move on to consider.

4.3.3 Dialogue Elements

Dialogue Elements (DEs) are the factors which control and influence the dialogue structure, putting the "meat" onto the skeleton DSM. They combine in unique sets to fill in the details to the basic DSM, and the presence of each individual DE produces an effect upon the way the dialogue is structured. The DEs can be viewed as corresponding to Schank's notion of primitives [SCH72] and the primitives used by Wilks [WIL75a]. However, instead of physical primitives these DEs are the underlying features present in the dialogue structure rather than direct actions.

Without the existence of the DEs, the DSM is nothing; the DSM is, in essence, the set of unique DEs which come together to form the DSM, in the same way as a cake is nothing until all the individual ingredients are brought together in their correct quantities. By using the concept of DEs, dialogue can be stripped down to its basic units providing a compositional approach to the analysis of dialogue. By reducing dialogue structure to a set of basic components, it is possible to investigate the way in which they reassemble to form the whole. This is akin to the structural approach described in Chapter Three and developed by such researchers as Grosz and Sidner, Polanyi and Scha, Kamp and Reichman. The theory for modeling dialogue presented within this thesis is concerned ultimately with deep dialogue structure as opposed to what is actually said or the specific words involved. Whereas Grosz' dialogue structure is comprised mostly of linguistic segments, our building blocks relate to the motivational, linguistic and external aspects of dialogue.

Rather than behavioural factors, these DEs are motivational factors (e.g. to obtain information), external factors (e.g. time limit), or verbal factors (e.g. rhythm) which all act as constraints upon the dialogue. In the same way as we expect certain actions and events to take place in well known, stereotypical situations, we also expect certain dialogue structures to occur in standard dialogue situations. For example, the informal dialogue between friends in a restaurant is likely to be casual, sharing the same goals and equally distributed between the participants in terms of the time each spends talking. It would be highly unlikely that two friends in a restaurant would speak to one another with a great deal of formality. This is modeled by the combination of DEs present within the "informal chat" DSM.

Some Dialogue Elements will be present in many dialogue situations, but in varying degrees. For example, both lectures and charity appeals may attempt to convince a participant that a particular belief is true, but the charity appeal is likely to have a higher *persuasive* motivation than a lecture. At the end of the day, the lecturer is not primarily concerned whether his/her students *believe* what is being said, as long as they are aware of the relevant opinions and can pass the final exam. The charity appeal, however, *must* persuade the audience that the beliefs being presented are true, or else they receive no donations. Therefore, it becomes apparent that in a given dialogue situation, the corresponding DSM will comprise of some elements which absolutely must be included, some which are optional, and others which are more than simply absent, but are actually forbidden. Furthermore, the importance or "strength" of a DE can vary depending upon the DSM in play. Continuing the cake analogy used earlier, if one ingredient is doubled in quantity, then the flavour of the final product is altered. Dramatically change one or more quantity, and you may even end up with a totally different, but recognisable, cake altogether. The same is true of the mixing of DEs and the way in which their presence and strength affect the overall dialogue.

At this point, it is necessary to examine the DEs within the theory, and the roles they play within dialogue. The set of DEs is listed below, categorised into the three groups we mentioned earlier, which are motivational, external and verbal DEs.

4.3.3 (a) Motivational Elements

In each DSM, at least one DE belonging to the set of Motivational Elements (ME) *must* be present. These are DEs which are connected to a purpose; they embed the fact that the dialogues take place in order to achieve some goal. The set of MEs is as follows:

- **Goal**. Indicates whether a dialogue is conducted in order to achieve a process or a task, and what that goal is. The nature of that goal may dictate that other DEs are used to achieve the ultimate goal of the participant, (e.g. may use Emotional Exchange DE in order to make some one feel scared and therefore disclose information during an interrogation).
- **Persuasive**. Indicates whether one or more of the participants are attempting to make the other(s) believe that X is true (e.g. political debates). This concerns the exchange of beliefs (as opposed to plain facts or emotions).
- **Information Seeking.** Indicates whether the participant is attempting to acquire information about a topic (e.g. asking for directions to the nearest bank) or is attempting to impart information to the other participant(s). This DE concerns the passing or receiving of factual information.
- **Emotional Exchange.** Indicates whether one or more of the participants are attempting to alter the emotional state of the other(s), (e.g. jokes, comedies, tragedies).

It is clear from the above list that a dialogue can have more than one ME present, (a Shakespearean comedy, for example, intends not only to make the audience laugh but to raise certain moral issues and be persuasive).

4.3.3 (b) External Elements

- **Number.** Indicates the number of participants involved in the dialogue. Under the definition of dialogue given in Chapter One, there must be at least two participants involved. Input received must be acknowledged to the particular participant from whom it came, and similarly it must be made clear to which participant(s) the output is being addressed.
- **Time Limit.** Indicates whether time constraints are to be considered within the dialogue, either because it must end at a specified time, or because it must only last for a given length of time. This affects the dialogue structure in a variety of ways: if, for example, the time limit is reached and one participant has much greater dominance than the other(s), (s)he may terminate the dialogue; alternatively, if the time limit is approaching, the participant(s) may begin to draw the dialogue to a close (e.g. moving from the main body of a lecture to the conclusion).

4.3.3 (c) Verbal Elements

- **Distribution Of Time.** This relates to the sentence length of each participant in relation to the others' sentence length. In crude terms it can be seen as the length of "speaking time" a participant is allowed compared with that of the other participant(s). If, for example, the current DSM is the *Job Interview*, then the interviewee knows that (s)he is expected to talk for a greater length of time than the person(s) conducting the interview.

- **Rhythm.** This DE relates to the length (or "rhythm") of the actual sentences given as output by a participant. Rhythm is measured in terms of Short, Medium or Long. Note that whereas *Distribution Of Time* relates to overall speaking times at dialogue level, this DE is operating at *sentence* level. Furthermore, phrase length does not dictate sentence length, (a phrase may be lengthy but the individual sentences short, thus producing a staccato effect).
- **Colour.** This is measured in terms of None, Low, Medium, High or Total. The "higher" the rating, the more "flowery" the language in terms of a frequent use of adjectives, sub-clauses, metaphor etc.
- **Register.** This relates to the vocabulary used within the dialogue. These include such common registers as "formal", "informal", "slang", "jargon", "swearing" and "polite", or may even include less frequently encountered registers as "poetic" and "archaic" (e.g. old English). The DSM may impose a specific register upon a participant's dialogue or it may be unspecified, left to the participant to select according to other parameters such as current emotional mood, personality or relationship with the other(s). In addition, the register may fluctuate within a dialogue as a result of different occurrences (e.g. although the appropriate register for a job applicant within the Job Interview DSM is the "formal" register, the register may change to "swearing" if the interviewer asks a taboo question such concerning religious beliefs).

- **Temporal Progression.** This conveys information regarding the manner in which the dialogue structure unfolds in the "real world", and the stages through which the dialogue may pass or must pass. The complete set of stages is: Greetings, Introductions, Main Body, Conclusions, Closing Phrases and Farewells.
- **Fixed Topic.** This relates to the subject(s) which may be discussed within the dialogue, and whether the topic is fixed or wandering. For example, in a lecture the topic is fixed and it would be unusual for a lecturer to suddenly discuss

unrelated subjects. The dialogue held between friends having a casual chat, on the other hand, may cover a large number of different topics.

As explained above, the individual DEs may be divided into a selection of sub-options, allowing a great deal of flexibility within the theory. DSMs may, therefore, be made up not only of a combination of DEs, but also of a variety of 'slots' within each DE. The DEs are descriptive, because they describe what a dialogue is made of, and the slots are prescriptive, because they indicate what should be done in certain circumstances. A description does not inform as to the actions to take in a particular situation, but simply offers an understanding of that situation. A theory should be both descriptive and prescriptive, as then it provides both an understanding of the situation and a course of actions to be taken in response to the situation. One without the other is unsatisfactory.

4.3.4 Direction

It cannot go unnoticed that information or messages of some type are being transferred within all of the DSMs. The whole purpose of language is to communicate (whether that be facts, beliefs or emotions). However, an NLP system needs to know when it is meant to be receiving and when it is required to give out information. Information exchanged by the dialogue can be seen in terms of input or output. Obviously, whether a certain dialogue produces input or output depends upon which role you are playing within a particular DSM. For example, in an *Information Seeking* dialogue, one participant will ask for information and the other participant(s) will reply, thus producing output. The inquirer receives that information as input and the cycle is iterated building into a dialogue. This is not an innovative view of dialogue, as many others have constructed systems using this notion, especially for the development of natural language database front ends, which treat questions as database queries and statements as database updates. However, dialogue is obviously much more complex than that, as a participant may be receiving and giving information simultaneously. Overall, however, the motivation will be one or the other.

Direction can be seen intuitively as the "role" which is being played. A different DSM may exist for a number of participants within the same dialogue situation, as the combination of DEs present may vary for each participant.

4.3.5 Constraints

Once the system has a set of DE's which correspond to the Dialogue Structure Model for the particular situation, it is necessary to move from the DEs to a set of constraints upon possible actions. These were arrived at intuitively. Each DE carries a set of attached constraints which begin to close down on the set of actions permissible and "the norm" within the particular situation. The constraints can be seen as sets of instructions concerning high-level actions without actually stating *how* to perform the task (e.g. ask a question, pay a complement etc.). If a DE does not carry a constraint, then fundamentally it is useless as it cannot be used for an action. If this is the case, then the DE simply serves to describe the situation, not actually prescribe anything. As previously mentioned, description and prescription are both important, and both should be present. A sample of some constraints currently within the DSM theory, and the DEs to which they belong, are given below. For a fuller list of constraints, the reader is referred to appendix I:

Dominance

1. Initiate the dialogue, (if other person has not already done so)
2. Terminate the dialogue
3. Interrupt a sub dialogue
4. Choose the topic
5. Change the topic
6. Initiate a sub-dialogue
7. Ask a question
8. Refuse to answer a question

Time Limit

1. Determine whether there is a time limit
2. Determine absolute time limit, (i.e. precise time dialogue must end at)
3. Determine duration of dialogue, (i.e. how long the dialogue must last)
4. If time limit is reached, terminate dialogue
5. If no time limit is found keep going until:
    a) something external happens to stop you.
    b) fundamental goal fails, (may apologise or not)

Temporal Progression

1. Make greetings
2. Initiate introduction
3. Make small-talk
4. Move to main body of dialogue
5. Move to summary of main body
6. Terminate dialogue with closing phrase, (good-bye, thank you etc.)
7. Scale dialogue according to time limit

8. Exit the dialogue, (other person has terminated the dialogue but you still need to say good-bye etc.)

Persuasive: (from direction of persuader)

1. Use only information which is beneficial to goal
2. Use *Emotional Exchange* constraints
3. Determine whether the other(s) believe what is being said

(from direction of person being persuaded)

4. Determine whether other(s) is/are telling the truth
   a) Believe what is being said
   b) Reject what is being said

When a DSM is called into play, the set of corresponding DEs is automatically known and, therefore, the full set of constraints designated to it. However, not all of the constraints attached to each DE may be appropriate within that particular DSM or at a particular point in the dialogue. Furthermore, more than one constraint may be applicable at a given point in the dialogue, and a choice has to be made as to which is the most appropriate selection. Clearly, a mechanism is needed for ranking constraints and sorting them into order of priority. This mechanism is provided as part of the function of the plan boxes and analysis boxes. We now move on to discuss these components of the theory below.

4.3.6 Analysis Boxes

Up until now, the focus of the theory has been upon the modelling of dialogue structure, rather than the instantiation factors. This imbalance is now to be corrected. It is via the analysis and plan boxes that much of the instantiation factors are taken into consideration, largely by those relating to the exchange of emotions and to the personalities of the participants.

The analysis boxes, like the plan boxes we shall consider next, are made up of a number of detailed instructions which provide information about how to do something. However, while the plan boxes relate to the production of output, the analysis boxes relate to the understanding and interpretation of the input. Although a large amount of analysis is already assumed to have been performed upon the input by the existing "base" modules of the NLP system (e.g. grammatical, semantic and pragmatic knowledge), some information still needs to be gleaned at the dialogue level.

The analysis boxes deal with such issues as "Joe Bloggs has just said X. How do I know if that's an insult or not?". In this example, the analysis box would use the rules contained within it to analyse the input, by stipulating actions such as "If Joe has said something positive about me I should feel happy", or "If Joe has said something of a negative nature about myself or someone I care about, I should feel angry" and so on. If, for example, the rules determine that the input has been an insult, the plan box which concerns the reaction to insults is triggered. These plan boxes are discussed in the next section, and are constructed in an identical fashion to the analysis boxes, but whereas the former contains rules for understanding the nature of the input in the context of the dialogue so far, the latter contains rules for expressing and reacting to the effect of that input. We now move on to consider the notion of plan boxes.

4.3.7 Plan Boxes

When an input has been analysed, a response must be made and an action taken. Not only must the appropriate action be selected, but the task of actually performing that action must be achieved, or else the only powers the theory has are of a descriptive nature, and not prescriptive. This requirement is met by the notion of plan boxes, which can be viewed in simple terms as large groups of detailed rules and instructions relating to how to select the most appropriate constraint for the current point in the dialogue, and then how to execute that constraint.

Thus, the constraints inform as to what *may* be said or done at a given point, and the plan boxes inform as to *how* this is to be carried out and *when*. So, by way of an example, if the current DSM dictates that participant X has a much greater level of dominance in the dialogue than participant Y, then the constraints say that Participant X may terminate the dialogue if (s)he so wishes. However, this may not be an appropriate action to take at the current point in the interaction. Although a lecturer has greater dominance in a lecture than the students, it would be odd for the lecturer to walk out half way through the hour without any justification. Thus, the possible action *"terminate the dialogue"* would be marked as being a highly inappropriate, albeit possible, action to take at the current point in the dialogue. The way in which the plan boxes mark possible actions is by assigning a rank (an integer) to each of the possible actions which the system may perform according to the constraints. The higher the rank, the more "permissible" the action. Thus, the plan boxes are made up of rules relating to the appropriate/inappropriate points in a dialogue when a certain action may be taken.

The potential number of plan boxes is infinite, as the more plan boxes that are added, the more sophisticated and wide-ranging the dialogue actions capable of being performed by the system become. The plan boxes are independent of the DSM which is currently applicable, and each plan box may be triggered by any DSM. The way in which other factors affect the ranking of a potential action are discussed in subsequent sections.

It is important to point out that the plan boxes are not the be-all and end-all of planning. While they do provide a planning mechanism within the theory, the planning they perform is limited. In effect, the plan boxes provide standardised answers to situations which may arise at any point in the dialogue. Fundamentally, they serve to provide reactive planning, formulating plans "on the run", according to what has just taken place within the dialogue. The author is aware, however, that reactive planning is not totally sufficient within a dialogue analysis system, and that planning is much more complex an issue than the plan boxes alone reflect. Long-term planning which encompasses the dialogue as a whole is required if a completely "natural" dialogue is to be produced. This has not, as yet, been included in the work, although work is currently in progress to address this issue. The absence of a long-term planning mechanism is not the result of any particular cause, but merely because the issue was not covered in the project. Planning issues can easily form projects in their own rights. Furthermore, it is believed that a long-term planning mechanism is not needed in order for the work to meet its original aims, and that the work is still valid: the work models two aspects of dialogue (structure and instantiation), and provides the means of dialogue generation. While the inclusion of a long-term planning mechanism would enhance the work, the research as it stands is sufficient in its own right to meet the criteria for success.

Now we move on to discuss the most influential and important of the plan and analysis boxes, namely those which govern the way in which emotionally charged inputs are reacted to, and emotions are expressed and/or stimulated.

4.3.8 Plan And Analysis Boxes Relating To Emotions

The most important analysis and plan boxes are those which relate to the exchange and interpretation of emotionally charged language. Within the project, emotional behaviour is accounted for by the construction of two plan boxes and an analysis box to deal with the expression of emotions, the reaction to emotionally charged input, and the intention to cause an emotional reaction within another participant. A set of emotions were arrived at intuitively, such as happiness, excitement, self esteem, misery, pity and so on. The "emotional" plan boxes and analysis box are described below.

1. **Emotional Reaction.** This analysis box provides information on how to react and behave if the input has a certain emotional value, and how to recognise the emotional value attached to a piece of language. For example, if the input is classed as an insult of some kind, the current emotional value of the system will be set to "anger" or "shock". Depending upon the particular circumstances, the appropriate plan box concerning emotional expression will be triggered for the value "anger" or "shock" and the instructions found there will be followed. Certain parameters govern the reaction, e.g. if the insulting input has been received from a friend, the expected response will be shock. If, on the other hand, the other participant is an adversary, then an appropriate response would be to send an insult back as output. The current "emotional state" of the system/participant is also altered according to the input received. For example, the settings for an emotional

value (e.g. Anger, Happiness, Fear, Excitement etc.) can be set to None, Low, Medium, High or Total. If the current emotional state of the system is "Anger" (Medium) because the last input received was an insult, and the next input received is another insult, then the emotional value of the system remains "Anger", but will increase to "Anger" (High). These examples are the sort of rules held within the analysis box, which inform the system about the sort of emotional response required for different types of emotional input. (Note that information relating to the content of the input is received by the dialogue analysis module from elsewhere in the system). Thus, the analysis box examines each piece of input and determines the emotional charge of that utterance. This may mean switching the value of the emotional state (say, from happiness to anger), or altering the intensity of that emotional state (for example, from mildly annoyed to very angry).

2. **Emotional Expression.** The **Emotional Expression** plan box instructs as to the expression of whichever emotion is the current emotional state. For example, if one is happy then the dialogue produced by that person will express that emotion. There is always a current emotional value, even if this is only set to "Neutral Emotion". The **Emotional Reaction** analysis box analyses the received input, decides which emotion is the appropriate response to take (if any change is dictated) and changes the current emotional value of the system to that emotion. This is then picked up by the **Emotional Expression** plan box, and the corresponding behavioural instructions picked up and sent to the generation module to be considered when compiling a response output.

3. **Cause Emotion Plan Box** This plan box contains instructions concerning how to cause an emotional reaction by another participant (in other words, how to make someone feel happy, sad, angry, excited and so on). This may be triggered by the

Emotional Expression plan box, the **Emotional Reaction** analysis box or by other, separate plan boxes. For example, if the input received has been from an opponent of some type and has been an insult, the **Emotional Reaction** analysis box may change the current emotional state to "Anger", and this in turn triggers the **Cause Emotion** plan box, and instructions picked up concerning how to express anger or make the other person feel miserable. Rules relating to the expression, reaction and causing of emotions have been arrived at via an intuitive approach.

4.3.9 Individual Factors

Consideration of the individual factors ensures that a complete understanding of the dialogue is achieved. In Chapter Two it was argued that in order to encapsulate the complete meaning and motivations of a dialogue both the structure and the individual aspects of the interaction should be handled. This interaction between the constant, underlying structure of dialogue and the peculiarities of the individual dialogue result in a more comprehensive use of dialogue. The individual factors discussed below, combined with the structural analysis provided by the DSMs, give us that desired flexibility. The individual factors are those which are unique to the dialogue but are pre-determined before the dialogue begins.

4.3.9 (a) Personality

Personality is sub-divided into the categories below:

- **Personal Beliefs And Interests:**

One way in which the dialogue may be affected is as a consequence of the participants' beliefs and interests. For example, if two strangers met at a party, and discovered a common interest in stamp collecting, the pursuing dialogue is likely to be more friendly and informal, as both participants relax and enjoy each other's company. This will affect the dialogue structure by altering the strength of the different DEs present in the DSM. Alternatively, if a person insults another's religious beliefs, then the subsequent dialogue may turn from a chat into an argument, a discussion or may even end abruptly if one party walks away. Note that in the latter case, the option chosen will depend upon other modifying factors such as personality, mood, and background (one may be less likely to storm away in anger if personal upbringing has taught politeness and good manners). The next sub-section looks at these factors in more depth.

- **Characteristics:** As just stated, another factor which may affect the specific instantiation of a DSM are the personal characteristics of the individual participants involved in the dialogue. For example, a highly confident person may be more assertive than a less self-assured person, and so the *Dominance* DE may increase in strength. This will naturally be reflected in the dialogue. Plan boxes enable factors to be taken into account such as the influence played by personal characteristics upon dialogue structure. (Chapter Five describes the plan boxes defining the "personality" of the system constructed during the implementation of this theory). In the latter example given above, a shy person may not feel confident enough to state that what the other person has said has insulted his/her own beliefs. Indeed, a shy person may not enter into much of a conversation at all, and the *Distribution of Time* DE may increase for the other participant because they have to talk more in order to sustain the dialogue. On the other hand, a highly aggressive individual may not only inform the other person that they are insulted, but may shout, be abusive and then terminate the dialogue. Thus it is apparent from observing human behaviour that personalities affect the dialogue structure.

- **Personal History:** The personal history which each and every human being has will affect the way in which they communicate via dialogue. We all have a history, and have all experienced life in different ways. Those experiences affect the way in which we think, feel and thus express ourselves. By way of an example, for most people the subject of terrorism is a theoretical or political topic. If involved in a discussion about terrorism, most people would talk objectively about the subject. If, however, your family have been killed by terrorists, your dialogue is likely to be far different from a political debate. The *Emotional Exchange* DE would probably increase, along with the *Persuasive* DE. Your emotional state is unlikely to be "neutral".

  Other historical aspects such as family upbringing will also influence our use of dialogue. For example, a person brought up in an environment which encourages freedom of speech and thought, will produce a different dialogue structure from that of an individual who has been taught to remain silent and never discuss personal opinions or emotions. Cultural differences relate to this, for example, British people are said to remain far more reticent and unemotional when speaking to people with different backgrounds.

Thus we can observe that the personal experiences, background, culture and social environment in which one has inevitably been in contact affects the dialogue structure which we produce.

4.3.9 (b) Emotional Mood

Human dialogue participants do not always begin an interaction with a neutral mood. They always come to a dialogue in one mood or another, which is rarely neutral or indifferent. Mood is determined within the theory as relating to the emotional state of the participant at the current point in the dialogue. The set of emotions was divided up into positive and negative emotions, and a value for Neutral. If the current emotional state of a participant equates to a positive emotion (such as excited) then the participant is classed as being in a good mood. However, if the current emotional state of that participant equates to a negative emotion, then their mood is classed as bad. Having a family quarrel before setting off for work may well affect one's mood (and, therefore, one's dialogues) until something occurs to change that emotional mood. As far as the work is concerned, an initial mood has to be set for each of the participants at the start of the dialogue. This is set manually, as it cannot be inferred prior to the dialogue having begun. The emotional state of a participant is limited within the work to only one emotion at a time, although it is recognised that human emotions are complex and it is possible for us to feel more than one, perhaps even contradictory, emotion at a time. Every input/output can be seen as having an emotional value, even if it is classed as "neutral". The emotional value of the participant is then picked up by the emotional analysis and plan boxes, and acted upon accordingly. The participants' mood may change as the dialogue progresses, which is dealt with in the next section.

4.3.10 Instantiation Factors

The instantiation factors are those relating to the dialogue at the current point. These are discussed below:

4.3.10 (a) Emotional Mood

The exchange of emotions affects the "mood" of the participants involved in the dialogue. For example, if a complement has been paid, the recipient's mood may well improve. Conversely, if an insult is thrown, the recipient of that insult may be put into a bad mood. Thus as the dialogue progresses, a participant's mood may alter, according to what has been said. The results of the emotional analysis/plan boxes discussed above is the alteration of the participant's emotional state according to the rules set out in the plan/analysis box.

4.3.10 (b) Interaction Between DEs

The DEs in an instantiated DSM will interact with each other, resulting in a modification of the set of constraints. Some DEs depend heavily upon each other, while others are more loosely interlinked. For example, if the *Time Limit* DE is present in a DSM, and the allocated time is running out, the *Dominance* of one or more participants may shift. This is illustrated by seminars/presentations, where the chairperson may step in to interrupt a speaker when time is short, whereas it would be inappropriate for him/her to do so at any other point in the dialogue. The constraints, therefore, will be modified according to the shift in strength of the *Dominance* DE, as control of the dialogue will pass from the speaker to the chairperson. The set of constraints for each participant will change from point to point in the dialogue as the DE's interact with each other. In general, if a constraint linked to a DE of the motivational type, (referred to as ME in section 4.3.3), is not satisfied for a length of time, this will become more and more important as the dialogue progresses.

Note that in the same way as DEs may influence each other and interact, so may the individual factors mentioned above. A person's character, beliefs, interests and history may all influence each other, affecting the strength of the various DEs in turn. For example, if a person is very confident in personality, has recently lost their job through redundancy and is told by another dialogue participant that all unemployed people are useless scroungers on the state, then it is likely that the reaction to that statement will be assertive and angry. On the other hand, if the person is less confident, and has no personal experience of being made unemployed, (s)he may choose to ignore the comment and change the subject, even if (s)he does not actually believe the statement is true. Thus individual factors influence each other, and as a result the DEs are altered to reflect these changes.

4.3.10 (c) DE Strength

The importance of a DE may vary from one DSM to the next, each having different strengths. The *Dominance* DE is present in both the interrogation and job interview DSMs, but is present in different levels. The interrogator has more dominance over the other participant(s) than the job interviewer does. The interrogator may verbally abuse the other participant, question him/her on any topic he/she chooses, may start and stop the dialogue abruptly when and as he/she decides and so on. The interviewer has a different set of rules which must be followed, as the level of *Dominance* in the DSM dictates. For example, it would be rude of the interviewer to suddenly interrupt the interviewee in mid-sentence, and certainly would not be expected to throw verbal abuse at him/her. This set of rules must be reflected by the set of constraints appropriate for each DSM, as the strength of the DE's will enforce a certain set of rules which regulate the appropriateness of the dialogue.

4.4 The Theory In Relation To Other Work

Having presented the theoretical aspect of the work, it is now possible to discuss the project in the light of the work discussed in Chapter Three.

In Chapter One we declared our definition of dialogue to be discourse plus the added feature of rich interaction. Much of the work conducted according to the structural approach investigates discourse. We would argue that work on handling linguistic features such as anaphora, ellipsis etc. is vital if one is to understand discourse, and one has to understand discourse if one wants to understand dialogue. However, it is dialogue analysis which is the goal of this work, not discourse analysis. Therefore, work on such structural views of discourse are not closely related to the presented work. It is the deep structure of dialogue which we are interested in, i.e. the motivational, verbal and external building blocks of the dialogue, and not the linguistic structure. We argue, therefore, that while structure is important, the work on structural analysis of dialogue discussed in Chapter Three looks at the language without taking its context into account.

With regards to the semantic analysis of dialogue, this is dealt with by the semantic analysis modules of the LOLITA system. We assume, therefore, that semantic understanding of the dialogue has already been achieved by the time the dialogue analysis module beigins its operations. As a result, the theory has not been developed with the intention of modelling semantic aspects of dialogue.

On the analysis of intentions and plans, we believe that this is a vital consideration but not the overriding factor. Social conventions and constraints must also be considered.

For example, it would not be expected for someone to go into a job interview and just keep stating "I want the job, I want the job". We argue, therefore, that it is only by modelling intention with other factors of dialogue that a complete understanding can be achieved.

4.5 Conclusions

As this approach provides the system with knowledge about the motivational, verbal and external features involved in dialogue, the theory is capable of modelling natural language dialogue and providing the means for generating responses appropriate to the dialogue situation. For example, if it is known that the current dialogue is an argument, (by being explicitly told the dialogue situation or by inferring it by matching the DEs which it recognises), then it is known which constraints and controlling factors go to make up that particular dialogue structure. Coupled with this ability to model dialogue structure is the consideration given to the individual factors discussed above, in particular the analysing and expression of emotions. Individual factors such as personality, mood, background etc. influence the dialogue and are therefore provided for in the theory. The combination of a constant dialogue structure, with a unique set of individual factors which are peculiar to the dialogue, provides the theory with flexibility.

The work combines a structural emphasis with the emphasis of intention and other factors. Others have advocated this multi-faceted approach to dialogue analysis [LAM93, PUS87], and it would appear that opinion is moving in the direction of believing that if robust dialogue systems are to be built which are capable of understanding the complete meanings behind natural language, then we need to consider a variety of dialogue features instead of only having one or two strings to our bow. The DSM theory is one offspring of this emerging body of opinion.

Chapter Five: Design And Implementation

5.1 Introduction

This chapter presents a discussion of the practical implementation of the Dialogue Structure Model theory described in Chapter Four. The implementation has been constructed using the functional programming language Haskell and has been designed to work in conjunction with a large NLP system developed at the University Of Durham, England. This system is called **LOLITA** (**L**arge-scale, **O**bject-based **L**anguage **I**nteractor, **T**ranslator and **A**nalyser). In this chapter we aim to present the reader with a description of how the theory we have already discussed has been built into a computer system. This system forms the physical embodiment of that theory. Examples of Haskell code have been given in places throughout the chapter where appropriate, but only where the code is easily understandable for the reader who is unfamiliar with either the language or implementation aspects of the LOLITA system. Lengthier code listings have been included in the appendices.

Before a discussion of the dialogue system can be appreciated, it is essential to at least have a general level understanding of the much larger LOLITA system, how it has been constructed and what it can currently do. Without some knowledge of LOLITA, it would be difficult to fully understand any system which has been built as an "add-on" module. For this reason, there now follows an overview of LOLITA.

5.2 The LOLITA System

The LOLITA system has been developed over a number of years at the University of Durham, principally by Garigliano and Morgan [GAR92, GAR93], within the LNLE. It is this LOLITA system which is used as a base for the dialogue analysis module to work upon. The LOLITA system has been developed in modular form, and currently runs on a 48Mb Sparc file server. The program is divided into approximately 150 modules, made up of around 35,000 lines of code, and contains more than 1500 English grammar rules

The LOLITA system has been developed as a general purpose base system, independent of any specific end application. As LOLITA is not restricted to a single task type, it is more than a generic system. In its current state, however, it does not meet the definition of a general purpose machine as it cannot, as yet, be used for any task in any domain. Although demonstrable prototypes have been constructed using LOLITA for a range of different tasks and domains (most notably contents scanning, machine translation and Chinese tutoring), no polished "final" application has yet been developed. This is due to the concentration so far of research resources upon the "base" of the system rather than on task-dependent system development. However, it is now felt that the base system is mature enough to allow more research activity to focus upon more restricted tasks and applications in the near future.

In order to understand and appreciate the LOLITA system for what it is, the following sub-sections will discuss the applications, main modules and other implementation aspects of the LOLITA system at a very general level, as this is all that is required in the context of this thesis. If, however, a more in-depth understanding of the system is desired, this can be gained from the literature referenced throughout this chapter or by contacting the authors.[1]

5.2.1 Applications Of LOLITA

As the LOLITA system has now reached a good degree of sophistication and maturity, specific tasks and domains are now being considered for application of the system. One such application is as a contents scanning tool, for which a prototype is already working, producing favourable results [GAR93]. A full semantic analysis of, for example, a 100 word paragraph from a newspaper article is achieved in a few seconds. Also currently under development is an intelligent tutoring system which uses the LOLITA system as a base and teaches Chinese to English students [WAN92, WAN94]. Other applications of the LOLITA system include a machine translation prototype (Italian to English), a query application for enabling a user to give information to and interrogate LOLITA using NL, an NL generator [SMI94, SMI94a], and a story application which allows the user to interactively build paragraph-sized pieces of text by passing a series of semantic network nodes, together with stylistic constraints, to the "realiser", (i.e. the final transformation stage of the generation process). For more detailed discussions of these, and other, aspects of LOLITA applications, we refer the reader to the literature cited. The dialogue analysis application of the LOLITA system is the work presented within this thesis.

[1]Contact Dr. R. Garigliano, LNLE, Dept. of Computer Science, University Of Durham, England.

5.2.2 The Programming Language

The functional programming language Haskell was selected as the construction language of LOLITA partly because of the lazy evaluation it features. Expressions are written as functions, and functions may call and be called by other functions. Lazy evaluation restricts to a minimum the amount of work needed to be performed by the system, as functions are only evaluated as and when they are needed. This is obviously an advantage when conducting searches on a large scale, as unnecessary work is ignored. Efficiency is thus kept at an optimum level and response time is kept at a minimum.

Another benefit of using Haskell is the high level of abstraction with which the system can be designed, enabling low-level functions to be "hidden away" from higher level functions. This produces programs which are easier to understand, as general level events and actions can be observed without any confusing details. For these reasons, this high level of abstraction also leads to programs which are simpler and easier to maintain. In NLP, this is an attractive feature, as the complexity of the problem leads to large systems being developed over long periods of time. If systems are to evolve and improve, maintenance and reusability must be a consideration at all times. This high level of abstraction also aids the task of building NLP systems in a modular form. As the need to deal with low-level code is removed, one can call other parts of the system into play without having to know the complete details of the mechanics. As most large NLP systems are built in modules, (e.g. parser, dialogue module, generation module, grammar, semantic network etc.), abstraction facilitates the way in which one module can interact with another in the system, by simply calling the top-level functions, thus linking the system sections together.

5.2.3 The Semantic Network

The semantic network of the LOLITA system currently comprises of approximately 400 files, containing world information as well as some of the system's linguistic information. This semantic network is built as a hierarchy of nodes which are interconnected by arcs. In its present state, the network holds 70,000 nodes, this translating to over 100,000 inflected word forms. This can be easily expanded via a natural language interface if needed. Information stored within the semantic network is held by means of a very simplistic structure. Each entry in the network (words, entities, relations and events), are represented as nodes, and all use the same internal representation. Attached to each node is a list of control variables which convey standard information pertaining to many nodes. These control variables can be used by the system for grammatical and semantic tasks.

5.2.4 The System's Other Modules

The system's other modules handle grammatical, linguistic and pragmatic processing of natural language input. These analyses are conducted each time a user inputs a piece of text via the normal keyboard. The output produced by these analyses (i.e. the information gleaned from the grammatical, linguistic, semantic and pragmatic phases) is passed as input to the dialogue module described below. Once the dialogue analysis module has performed its tasks and generated some output, this is passed as a package to the generation module. The natural language generation module then takes the output from the dialogue module, formulates a piece of natural language corresponding to the instructions sent by the dialogue analysis system, adjusts the style of the language, and then presents this to the screen for the user(s) to read. The following figure represents the present day components of the LOLITA system, and is intended to give the reader an insight into where the dialogue analysis module fits into the overall LOLITA system.

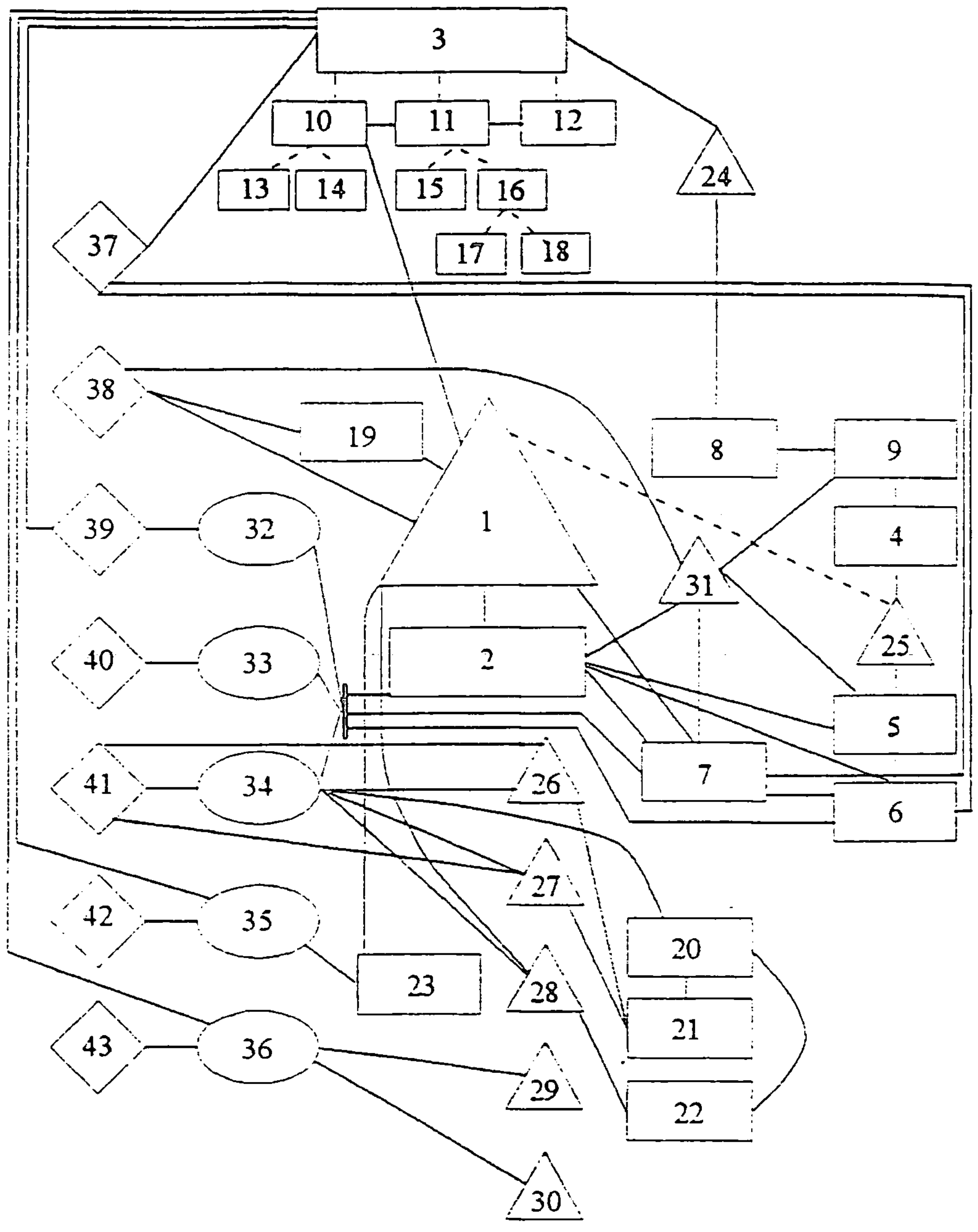


Figure 5.1: The LOLITA Block Diagram

Key:

rectangle: processing module
oval: application module

diamond: interface module
triangle: data

continuous line: module connection
broken line: subpart relation

1: semantic net
2: inference engine
3: syntax analysis
4: semantic analysis
5: pragmatics analysis
6: discourse analysis
7: natural language generation
8: syntax tree normalisation
9: pre-semantic normalisation
10: morphological analysis
11: grammatical analysis
12: parsing search control
13: misspelt words recovery
14: new words guessing
15: grammar structure handling
16: grammar feature analysis
17: structure analysis
18: structure reconstruction
19: data maintenance
20: dialogue planbox generation
21: emotion analysis
22: constraints analysis

23: style analysis
24: tree structure
25: semantic net fragment
26: LOLITA model
27: user models
28: dialogue structure models
29: student models
30: tutor models
31: global switches
32: query module
33: template module
34: dialogue module
35: translation module
36: Chinese tutor module
37: NLP interface
38: semantic net interface
39: query interface
40: template interface
41: dialogue interface
42: translation interface
43: Chinese tutor interface

Figure 5.2: The Key For the LOLITA Block Diagram

5.3 The Dialogue Analysis Module

As the above figure shows, information regarding the syntactic, grammatical, semantic and pragmatic analysis of the input is received by the dialogue analyser from the other modules of the LOLITA system. Thus, links exist between the dialogue analysis module and the other components of the overall system, with information regarding input/output being checked, analysed and generated by a variety of modules. The dialogue analysis work relies upon the LOLITA system's modules to provide it with information regarding the *content* of the input/output language, as the dialogue analysis module focuses upon the *structure* of what is said. It also relies upon other LOLITA modules to transform its output into natural language which the user is then presented with.

The essence of the dialogue module is to allow rich interaction with the LOLITA system at dialogue level, as opposed to single sentence (i.e. discourse) level, and with . Until the completion of this module, interrogation of the system was possible one sentence at a time. To actually *converse* with the system, this dialogue module is needed.

At any point in the interaction, the dialogue is internally represented within the module by the Dialogue State. This contains all the information relating to the dialogue (including information passed on to the dialogue module from elsewhere in the system), and is "carried around" by the dialogue analyser throughout an interaction. Essentially, the dialogue state is a list of pieces of information, each item of information being required in some way for the operation of the module. As a dialogue progresses, the dialogue state is altered to reflect events in the interaction. In order to keep the dialogue state in a current and correct form, a cycle is followed and a sequence of actions are performed; this is, broadly speaking, analysing input from the user, formulating a set of appropriate potential responses, and then selecting the most acceptable action from this set, which is then passed on to the generation module to be produced as output to the user. The cycle is then repeated once the user has replied to the system's output, and so the overall dialogue is constructed in an incremental manner. The Dialogue State shall be discussed in greater detail in section 5.3.10 below. In order to provide the reader with a full understanding of the way in which the dialogue module operates, a more in-depth explanation of the dialogue cycle follows below.

5.3.1 The Dialogue Cycle

The dialogue analysis module is controlled at the top level by a series of actions which are performed each time input is received from another dialogue participant (in other words, a user). This sequence of events is iterated each time new input is encountered. The top loop comprises of the following sequence of events (listed in order of execution):

1. **Emotional Reaction** stage. This involves an analysis of the input from the point of view of emotional significance. The input is checked to determine if an emotional response is called for (e.g. should the input evoke feelings of anger, shock, happiness, excitement, misery and so on) or to determine if this input indicates that the other participant is trying to evoke an emotional response in LOLITA, thus helping to inform LOLITA about the participant's goal(s). This is the first stage carried out by the dialogue system because if an extreme emotional response is required (e.g. the user has insulted LOLITA she must respond angrily) then this

reaction will overshadow the other local and global goals of the dialogue (e.g. ask directions to nearest post office).

2. **Intellectual Reaction** stage. This stage concerns analysing the input under a logical, intellectual point of view. (i.e. does the input make sense?). Much of the difficulties resulting from ill-formed input are dealt with by the lower modules of the LOLITA system, and so are already sorted out before reaching the dialogue analysis module. However, perfectly formed inputs and semantically/pragmatically acceptable inputs may still fail to make sense within the context of the current dialogue or at the current point in the dialogue. This is the second most important stage after the emotional reaction phase, because if input is received which requires an intellectual reaction, then the need to express that over-rules other goals in the dialogue at that point.

3. **Temporal Progression Check**. The input is examined to ensure that it relates to a legal component of the *Temporal Progression* within the current DSM. Also, if the Temporal Progression needs to be advanced then the system should expect the input to signal the move to a new part of the Temporal Progression sequence. If the dialogue does not follow the expected progression, this problem can then be addressed.

4. **Application of DSM Constraints**. At this point in the top dialogue loop, the constraints relevant to the current DSM are examined and applied.

5. **Application of Individual Constraints.** This stage involves the invocation and influence of the constraints which are attached to the individual factors, (i.e.

constraints relating to such individual aspects as personality, background, relationship to participant etc.).

6. **Balancing of the Constraints.** This phase of the dialogue cycle ensures that all the constraints are balanced (i.e. all those constraints which have been selected as applicable at the current point in the dialogue are examined to check that there are no conflictions or inconsistencies between the constraints). The DSM constraints and the individual constraints are compared to see if one over-rules the other.

7. **Setting the Local Goals.** At this point in the dialogue cycle, the local dialogue goals are set (i.e. those which apply at the current point in the dialogue).

8. **Choosing The Plan Boxes.** At this point, the various plan boxes which could potentially be selected as an action in the dialogue are given a ranking value. The higher the ranking number, the more appropriate the action is as output to be sent to the generator. The plan boxes are implemented in a hierarchical fashion, because there may be more than one way to go about a certain task. Plan boxes may, therefore, be nested within more general level plan boxes, or may simply trigger other plan boxes within the system. They are implemented as clusters of code, each cluster relating to the performance and achievement of a specific task.

   The result of each plan box is known as a dialogue tactic. Each dialogue tactic is a package of information relating to how the generator could go about the construction of a piece of output relating to the action addressed by the plan box. For example, if a plan box is built to tell the system how to insult the other participant(s), the dialogue tactic produced by the plan box would be picked up by the generation module of LOLITA, and unpack the information telling it that the output should be an insult, of a certain length, a certain register, about certain topic(s) and so on. Note that the selection of which actual words are to be used is not a matter addressed by the dialogue tactics, but by the generation module. The dialogue tactics simply inform the generator about the *type* of output required and the structure which should be used. Within each dialogue tactic is the ranking number discussed earlier, and more than one tactic may be produced per plan box.

9. **Repetition Check.** This involves checking the tactics produced by the plan boxes. If a tactic is found to have been used as output before in the recently preceding dialogue, then its ranking number is lowered. For example, if a lot of questions have already been asked, it may not be wise to ask another for fear of being impolite. Obviously, DSM considerations may influence this decision, as well as the personal characteristics of the participants involved.

10. **Combination of Compatible Tactics.** This stage should combine those tactics which are compatible and have the same rank. More than one tactic may be used to generate an item of output. For example, a statement could be followed by a question, or an answer to a question may be followed by a change of topic and so on. On completion of this stage, the tactic(s) which have been identified and combined are passed on to the generation module in order to be transformed into words and phrases and displayed for the user to see.

11. **Realising the Planboxes.** This stage involves triggering the various plan boxes according to the results. These results are the ranking orders produced at the earleir stages. The plan box with the highest rank is the one which is used as an action in the dialogue. Some of the plan box rankings may indicate that they should not be used at the current point in the dialogue i.e. the ranking parameter is extremely

low, while others produce very high parameters and thus indicate that they would be highly appropriate actions to take at the current point. The ranking parameters are examined in relation to each other, not for their individual values per se.

12. **Producing The New Dialogue State.** In this final event in the dialogue cycle, the new dialogue state is produced to reflect the events which have taken place in the last dialogue cycle. The dialogue log is updated to record all the information relating to the decisions taken and results of analyses performed within the cycle. Having produced the new dialogue state, the system is ready to receive another piece of user input, and returns to stage one of the cycle, carrying with it the updated, current version of the dialogue state. As the cycle is repeated, the overall dialogue is constructed and recorded in the dialogue log.

The above steps constitute the top level cycle of the dialogue analysis module's operation. This is represented in Haskell code within the system as the following lines:

```
> mk_new_ds :: Dial_state -> Dial_state

> mk_new_ds
>  = normalise_d_tactic  . final_rank_trace .
>    realise_planboxes     . partial_tactics_trace .
>    combine_compatible_tactics  . sort_actions_by_rank  .
>    realise_constraints     . activate_previous_plan  .
>    avoid_repetitions       . choose_planboxes  . emotions_trace  .
>    partial_goals_trace     . set_local_goals . balance_constraints   .
>    apply_individual_constraints  . apply_dsm_constraints  .
>    compare_temp_part   . intellectual_reaction   .
>    emotion_reaction        . change_partial_tactics []
```

Some of the steps involved in the dialogue cycle are relatively straightforward and have already been discussed in sufficient depth in this section. Other stages are, however, crucial processes within the operation of the dialogue system, and merit further discussion. These latter processes are, therefore, presented in closer detail below.

5.3.2 Emotional Reaction

The purpose of this step in the cycle is to ensure that if the input from the user merits an emotional reaction, this is recognised and noted by the dialogue system. The emotional reaction section of the program constitutes sets of behavioural rules (analysis boxes) which should trigger a certain emotional response from the system. These rules specify the behaviour to be generated if a certain input should invoke a particular emotion. The **emotion_reaction** analysis box is made up of a sequence of functions which govern the reaction to a different emotion. A set of emotions recognised by the system were arrived at from an intuitive approach, and are listed, along with descriptions, in Appendix IV.

Each **emotion_reaction** analysis box relates to the circumstance(s) which could provoke a reaction of the type of emotion specified. In other words, if the input is of a certain content, (e.g. insulting, complementary, educational, accusatory, exciting etc.) and LOLITA's personality is of a given type (e.g. highly confident, shy, educated etc.) and LOLITA's present emotional mood is good, bad or neutral, the response should be an output carrying emotion X. For example, if LOLITA's personality is set at "Confident", and you insult her religious beliefs, and she is also in a bad mood to begin with, then the output generated by the system is expected to be an angry one. Contained within the dialogue state is a value relating to the emotional state of LOLITA. At any given point in the dialogue, LOLITA has an emotional value assigned, which is given an intensity value of Total, High, Medium, Low, or None. If an **emotion_reaction** analysis box dictates that a change in emotional state is required, then the emotion and/or its intensity is altered within the dialogue state. The **express_emotion** plan boxes then pick up LOLITA's emotional value and its intensity from the dialogue state and act accordingly. Note that the result of a **react_emotion** analysis box is simply that the emotional value or intensity of the system is changed within the dialogue state. The resulting behaviour (if any) is handled by the **express_emotion** plan boxes.

As well as informing the system about which emotions are used to react to different input, these rules also include instructions on when to increase an emotional reaction. For example, if someone insults you, you may react angrily. But if the next thing they say is another insult, your anger may increase from mild anger to "medium" anger. If the third thing they say is an apology or reasonable explanation, you may be placated, unless it is a third insult, at which point your response may turn to severe anger. Rules concerning the incremental increase/decrease of emotions have been included in this group of analysis boxes.

5.3.3 Intellectual Reaction

As well as reacting emotionally to input, it is necessary that the system be able to respond to input on an intellectual level. Following the emotional reaction check, the next event in the dialogue cycle is to check that the input is acceptable in terms of logic and reasoning. In other words, the system should reject or query any piece of input which does not make sense intellectually or rationally. This check is performed at three levels:

1. Is the input acceptable per-se?
2. Is the input acceptable in the dialogue locally?
3. Is the input acceptable within the context of the overall dialogue?

It should be noted that this operation is concerned with the correctness of logic, and not of syntax. If an input is syntactically ill-formed, this should be detected and dealt with elsewhere "lower down" in the LOLITA system. A piece of syntactically well-formed input, however, could still provoke an adverse intellectual reaction. For example, a piece of input may be well-formed, but not make rational sense within the context of the dialogue, or at a particular point in the dialogue, or even just be nonsensical as a statement.

Hence, the three intellectual reaction analyses are performed:

5.3.3 (a) Is The Input Acceptable per se?

This check utilizes information derived from the pragmatic and semantic modules of the LOLITA system. If a piece of input is not acceptable in terms of semantic or pragmatic analysis, then it must be rejected or queried as part of the dialogue. For example, if a user enters the expression: " My cat told me that the Bahamas are nice at this time of year", then the pragmatic analysis should inform the dialogue module that this is intellectually unacceptable: cats do not talk. Similarly, semantically ambiguous or incorrect inputs are recognised by the semantic network of the LOLITA system, and the dialogue module notified. Input received by the dialogue analyser has already been processed by LOLITA's other modules, so the information regarding pragmatic, semantic or syntactic soundness of the input is readily available for manipulation.

5.3.3 (b) Is The Input Acceptable In The Dialogue Locally?

It is possible for certain inputs to make sense per se, but be unacceptable as an input at the current point in the dialogue. For example, if a participant asks: "What is your address?", it would be illogical at that current point to state: "I hear that the Bahamas are nice at this time of year". The response statement is sound per se, but is simply illogical at a local dialogue level. Such localised anomalies are searched for early on in the dialogue cycle, because if they exist, then rejection or querying of their acceptability needs to be given precedence over other goals/considerations within the dialogue.

5.3.3 (c) Is The Input Acceptable In The Dialogue As A Whole?

Certain inputs in a dialogue are unacceptable because they clash with the expectations brought about by the DSM which is in play. Although these inputs may make sense as an expression, or be locally acceptable, they may not be acceptable within the current DSM. For example, if you attend a lecture on software engineering, you do not expect the lecturer to go off on a tangent and spend half an hour discussing the difficulties of growing prize tomatoes. While the statements may be intellectually sound per se, and may even be acceptable at the local dialogue level, it would not be acceptable in the context of the overall dialogue. This information is established by checking the input against the expectations stipulated in the DSM. In other words, does the dialogue conform to the combination and strengths of the DEs assigned to the current DSM?

5.3.4 Checking The Temporal Progression

This stage of the cycle is necessary to ensure that the dialogue is progressing as it should be. The expected sequence of events for each dialogue can be investigated by looking at the *Temporal Progression* DE attached to the current DSM. If, for example, the current DSM is the "Introductory Chat" DSM, (where two strangers meet informally for the first time), then the *Temporal Progression* DE stipulates that Introductions are usually exchanged, followed by small talk, and if common ground is found and no external factors interfere (such as one participant has to leave to catch a train), then the main body of the dialogue should be reached after a short while. If, however, one participant jumps straight into the main body of the dialogue before passing through the introduction stage, this should also be noticed. For example, it would not be unacceptable in such a scenario to say: "I'm sorry, I don't know your name .... Oh, John, OK, I'm LOLITA, ... sorry, you were talking about the Bahamas....". Alternatively, consider the example of a lecturer who never gets beyond the small talk section of the lecture, and spends an hour making chit chat about last night's television. Such discrepancies between temporal progression and input content should alert the system that something is not as expected within the DSM, and a remedial or inquiring response may be made to rectify the problem.

5.3.5 Applying The DSM Constraints

The DSM constraints are those which are attached to the DEs within the DSM. Currently, the DSM constraints have been implemented in a somewhat primitive manner, and are not as sophisticated or detailed as could be desired. However, it is important to stress that this is not due to an ignorance of how to implement the DSM constraints, but rather that time restrictions upon the project have meant that this area of implementation is not currently in as advanced a state as other parts of the system. It is envisaged that this issue will be addressed in the near future. It should also be noted that the system is still capable of holding an interaction even while the DSM constraints are in a somewhat primitive stage of coding, and that their full implementation will simply enhance the behaviour of the system, making it more sophisticated and "natural".

5.3.6 Applying The Individual Constraints

The individual constraints are a large factor in enabling the system to combine structure with instantiation, a problem discussed in Chapter Two. The combination of both aspects of dialogue affords us the flexibility and "naturalness" of behaviour which we stipulated as a criteria for success in Chapter One. The individual constraints are those which influence the dialogue due to specific factors. These factors are the participant's personal history, personality, beliefs, interests, and relationship with the other person. Constraints exist within the system relating to the way in which each of these aspects affects the dialogue, and are described in detail below.

5.3.6 (a) LOLITA's "Personality"

A participant's personality and individual characteristics play an influential role upon the dialogue structure produced by that participant. For example, a confident, forceful character is more likely to speak for a greater length of time in a chat than a shy, timid person. Therefore, if the system is to mimic a human participant, personality must be accounted for in some way. These following characteristics are given values of None, Low, Medium, High or Total, and can be determined manually by the user/system administrator at the onset of each dialogue. Attached to each characteristic is a set of constraints which cause a modification in the behaviour of the system. For example, if **Education** is graded as High, the range of registers used by the system may be wider, and sentence structure more complex than if **Education** had been valued as Low. For a full list of characteristics implemented in the system, along with a description of the intensities and the attached constraints, the reader is referred to Appendix V. The settings for these characteristics may be set manually by the system manager, or simply to a default setting.

LOLITA's "character" is made up of not only these characteristics and their attached values, but also of a set of motivations. These motivations are a list of hobbies, plans, beliefs and interests. These categories can contain as few or as many items as the system manager allows. Each item corresponds to a node in the semantic net. Therefore, if pole vaulting is listed in the system as one of LOLITA's hobbies, all information relating to pole vaulting may be accessed from the semantic net of the system. In this way, the "motivations" of the system can be added to, deleted and amended, and can be as extensive as one decides. The characteristics list is known within the system as "Static Information", because it is information which is determined by the system administrator before a dialogue is initiated, and does not alter during that interaction.

5.3.6 (b) LOLITA's "History"

The "personal history" of the LOLITA system is divided into a number of categories. Some of the categories constitute of lists of items, whereas others exist of only one. These items each represent a node in the semantic net. The categories are:

- **Occupations**: professions/employment profile.
- **Educations:** educational experiences and achievements
- **Background:** family background and "upbringing".
- **Accommodation:** place(s) of residence.
- **Marital Status:** single entry for "married", "single", "separated" or "divorced".
- **Travel:** destinations visited.

5.3.6 (c) Source Control Influences

When human beings communicate via natural language, a wide range of factors influence a participant's interpretation of what the other person is saying, and decisions are made by human agents concerning the reliability of what is being said. Thus, as participants in naturally occurring dialogues, human beings are continually assessing information and conversations according to our evaluations of the source, and so the dialogue module of an NLP system must do likewise if it is to perform in a "natural" manner. It was declared in Chapter Three that a dialogue system should incorporate a source control or user-modelling mechanism of some type if it is to account for the expressions of belief, doubt, uncertainty and several motivational aspects of dialogue relating to the way in which people relate to their dialogue partners. A simple source control module was constructed for use in conjunction with the dialogue analysis system constructed as a computational model of the DSM theory, and was based upon the work of Bokma and Garigliano [BOK94]. While this part of the system may be described as basic and far from a complex uncertainty management system, its influence upon the dialogue analysis system is still an improvement on systems which take no account of this aspect at all.

Within the source control component of the implementation, a model can be constructed for each person known to the system. These models are made up of a number of categories, which are fully described in Appendix VII. The system has been given a starting "default" source model, so that unless the system is specifically told that the person with whom it is "talking" is Mr./Ms. X, then it will assume that the other participant is the default source. A source model is loaded for the default, with values given for interests, beliefs, relationship and ability (all categories within the model). In its current state, this default is the only source model available to the system, but with the investment of more time and research effort, infinitely more models could be added. What is important is that the basic mechanism is in place, which the default serves to prove.

5.3.7 Setting The Local Goals

Although the DSM provides the system with a knowledge of the overall, global goal(s) which it strives to achieve, a number of local goals need to be satisfied during each dialogue cycle. At each new instance of the cycle, a fresh set of local goals is constructed. These goals relate not to the overall aims of the dialogue, such as passing the time of day, obtaining information about a topic, persuading X that Y is true and so on, but is more specific to the needs of the dialogue at its current point. For example, if the last piece of input received by the system is a question, one of the local goals of the dialogue module will become the goal of answering that question.

The local goals which may be set could be one, or a combination, of the following goals. These are not listed in any order of importance or priority:

- command
- inform (other participant(s))
- not inform (i.e. avoid divulging certain information)
- be informed
- express emotion
- cause emotion (in other participant(s)
- persuade

The local goals are set by looking at the global goals (stated within the DSM), and modifying this with relation to LOLITA's personality, background and the source control, LOLITA's current emotional value and its intensity (recorded within the dialogue state), which itself has already been modified by the input from the user(s).

Once the local goals have been set, the status of the input is examined (question, command, silence, statement), along with its emotional value, its content (superficially), LOLITA's personal characteristics and background, and also the constraints. Using this information, the system decides upon the fundamental form which it will use to satisfy the goal within the current dialogue circumstances, such as producing a statement, asking a question, giving a command, remaining silent or a mixture of these. From this, the necessary plan boxes are triggered.

5.3.8 The Plan Boxes

The plan boxes form the central feature of the implementation in the sense that the flexibility of the dialogue analysis system, and its ability to generate appropriate responses, rest with the plan boxes. As was described in the DSM theory discussed in Chapter Four, each plan box functions to inform the system when and how to achieve a certain task, such as to pay someone a complement, change the subject, obtain information from the other participant etc. As there is usually more than one way to go about such tasks, the plan boxes within the system have been constructed to contain information about how to achieve this variety of behaviours. Each option is given a ranking (an integer) which increases according to how appropriate that response would be if generated as an action in the dialogue at the current point. The action with the highest rank is the one which is selected and used as output.

The optimisation of parameters has proved to be a more than trivial consideration in the implementation of the theory, as the selection and behaviour of the plan boxes plays a greatly influential role in the behaviour of the system. As this is an important point, it is dealt with separately and in more detail in the following section of this chapter.

The plan boxes were implemented on a hierarchical basis, with each main goal having a plan box relating to the achievement of that goal. At the top of the hierarchy is the plan box which controls and sits on top of the lower, more specific plan boxes contained within it. The lower level plan boxes relate to the variety of ways in which the goal may be achieved, while the higher level plan boxes relate to when it is appropriate for that particular path to be followed. By means of an illustration, let us consider the *Persuasive* plan box in order to understand how these plan boxes have been implemented.

If LOLITA's current goal is to persuade the other participant that something is true, then the top level *Persuasive* plan box is activated. However, there is more than one way to persuade someone that something is true. You could either use the intellect to persuade the other person that X is true, or you could use emotional persuasion by appealing to the other person's emotional nature.

The **use_intellectPB** plan box is then sub-divided into further plan boxes:

- the **tell** plan box: relates to persuading X that Y is true simply by saying "Y is true";
- the **Self Interest** plan box: concerns trying to persuade X that Y is true because you claim that it is in Y's interests that X is true;
- the **Opinion** plan box: relates to invoking Y's opinion that X is true. For example, "Now as we both know, the Conservative Party has worked wonders for the economy of Britain";
- the **Beliefs** plan box: Here, a participant uses the other participant's beliefs in order to persuade him/her that X is true. For example, trying to persuade someone who is known to be a pacifist that the Government is evil by telling them that the government is involved in the sales of arms and weapons;
- the **Experience** plan box: This concerns trying to persuade X that Y is true by relating the argument to an experience in X's history. For example, if a member of X's family was killed by terrorists, and you want to persuade X that the Government is bad, then you could say how the government has secretly been negotiating with terrorists and selling them arms.

These plan boxes use the information stored in other parts of the system, such as the source control module, the "personality" part of the system and so on. The use_emotions plan box, on the other hand, concerns using emotional factors to persuade someone that something is true. This can be achieved by:

- making someone feel guilty about believing/not believing that something is true, and then offering to make them feel better if they believe what you say. For example, "If you don't believe in God, then you will not go to Heaven. But if you do believe in God, then you will".

- taking advantage of someone's positive feelings for you: For example, "If you love me you would believe in my innocence of this crime", or "If you love me you would share my belief in God".

- Using one's authority to invoke persuasive emotions. For example, the boss tells a worker that unless productivity increases, job losses may be incurred, thus using his/her position of authority to *scare* the worker into believing that (s)he should work harder.

- invoking self pity/misery within someone, and then offering to make them feel better if they believe that what you tell them is true. For example: "I was sorry to hear about you getting made redundant. I blame the government's economic strategies for all this unemployment. If we all vote against them at the next election, the unemployment situation may improve". In this example, the knowledge that X's redundancy makes X unhappy is then used to persuade X that the Government is not doing a good job, by blaming the source of the unhappiness upon the Government.

The *Persuasive* plan box hierarchy can thus be represented in diagrammatical form:

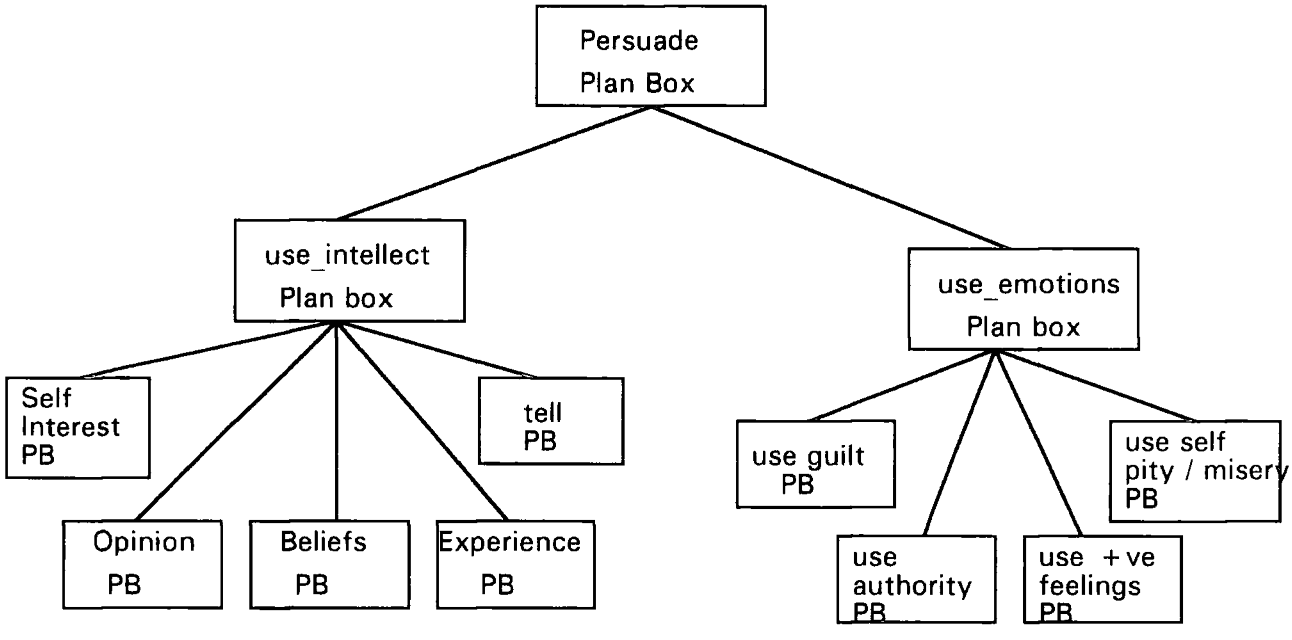


Figure 5.1: Hierarchical Structure Of *Persuasive* Plan Box.

So each of the sub-components of the top level plan box is activated, and produces a ranking parameter. Information is required and searched for from around the system modules by each of the plan boxes. Depending upon the information found (it is, for example, useless to trigger the **use_experience** plan box if the source control module can not provide any information relating to the history of the other participant) and depending upon factors relating to the current point of the dialogue (e.g. it may be highly appropriate to select the **tell** plan box if the other participant has just asked a question) and depending upon factors relating to the DSM. (it would be inappropriate to select the **use_emotion** plan box to be persuasive during a formal interview), the rankings obtained from the various plan boxes are examined. The plan box which scores the highest ranking value in relation to the others is the one which is selected as potential output. The top level plan boxes are also compared to one another in terms of ranking parameters, and so the plan boxes are competing with each other at various levels in order to be selected as the final choice of action.

The other top level plan boxes relating to main goals include:

- the **BeInformed** plan box: this relates to the goal of obtaining information from the other participant, and is subdivided into plan boxes of:

    1. **ask:** can be subdivided into plan boxes relating to asking for information directly or asking for information in a round about manner.
    2. **check:** triggered if LOLITA wants to check what the other participant has said.
    3. **why_quest:** this is activated when LOLITA wants to know why the other participant has said a certain utterance.
    4. **demand:** this is activated if LOLITA has a greater level of *dominance* within the DSM. than the other participant. This is a seeking information of the type "Tell me......" , or "I insist that you tell me....." etc..

- the **inform** plan box: This is activated when LOLITA needs to give out information to the other participant(s). This comprises of the following lower level plan boxes:

    1. **answer:** this plan box is triggered if LOLITA is asked a question. This may in turn trigger the **tell** plan box, or another plan box depending on the question

which has been asked and the DSM. in play (for example, in some situations you may not wish to divulge all the information you have).

2. **informPB:** this triggers the **tell** plan box. The DSM. may dictate that LOLITA's role in the dialogue is to impart information with or without being directly asked for it. (For example, a lecturer imparts information even if the students do not directly ask questions).

The other main plan boxes relate to the emotional behaviour of the system, and control the causing of emotional reactions in the other participant(s), and the expression of emotional states. The nature of these emotional behaviours was dealt with in Chapter Four section 4.3.8. (The reaction to emotional inputs is classed as an analysis box and not a plan box, as it does not result in a direct output action and so is not dealt with in this section).

Within the **cause_emotion** plan box and the **show_emotion** plan box is a set of lower plan boxes, one for each of the emotions recognised by the system. There are also a number of plan boxes to deal with the increasing strength of emotional mood. For example, if an input is analysed by the **emotion_reaction** plan box and is found to be mildly offensive, this has the effect of changing LOLITA's emotional mood to Anger Offence (Low). If, however, the following input is also mildly offensive, there is an accumulative effect and LOLITA's emotional mood may be increased from Anger Offence (Low) to Anger Offence (Medium). If yet another offensive input is received, the strength of emotional mood may be incremented again. This is because if some one says something which we find mildly offensive, we may (depending upon the situation, our character etc.) overlook the incident or react in a mild way. If, however, they continue to do so, our annoyance would increase. Note that "soothing" inputs (those which invoke positive emotions such as self esteem may decrease the present emotional mood, such as being paid a compliment) may appease us.

5.3.9 Optimisation Of Parameters

Currently, there are in the region of 124 planboxes within the dialogue analysis module of the LOLITA system. Each plan box carries with it a parameter which can be altered, thus the number of combinations of settings of these parameters is large. In the initial stages of the work, these parameters were fine tuned "by hand" in order to produce the particular behaviour desired from the system. Obviously, this manual fine tuning of parameters is less than ideal, and it was considered more desirable for the system to be able to handle this optimisation of parameters automatically.

Work has been conducted within the LNLE group at Durham upon the development of an algorithm which will perform this automatic optimisation. This began by manually tuning the parameters until the piece of dialogue given in figure 5.2 was generated. This was then adopted as the "target" dialogue for the optimisation process, and the process would be deemed successful if this was generated.

LOLITA:	I want a success.
Input:	I like you.
LOLITA:	I know; why do you like me?
Input:	You are beautiful.
LOLITA:	I am glad that you think I am beautiful.
Input:	Do you like me?
LOLITA:	Yes, I like you.
Input:	Do you like me?
LOLITA:	I am slightly bored and emotional.

Input:	You are a machine.
LOLITA:	I do not believe that I am a machine.
Input:	Are you a machine?
LOLITA:	No.
Input:	Are you beautiful?
LOLITA:	I do not know if I am beautiful.
Input:	I am tired.
LOLITA:	Tell me things.
Input:	I am tired.
LOLITA:	I slightly am bored.
Input:	Are you tired.
LOLITA:	No.

Figure 5.2: An example of a piece of dialogue that has been generated as a result of an interaction with a user and the LOLITA system. This is the target dialogue for the optimisation process.

The above interaction was generated as an interaction between the LOLITA system and a human user. At the outset of the interaction, the dialogue DSM. was specified as a chat, and LOLITA was using the default "personality". At the beginning of the interaction, LOLITA needs to make a choice between conflicting desires. In human-human interactions, when embarking upon a chat, we ourselves have to decide what we want to talk about. LOLITA must do the same, and must chose the option which is most suitable to the current context of the dialogue. As the dialogue unfolds, it is necessary to generate responses which are appropriate at the current point in the dialogue. For example, if an incoming statement is repeated, the value given to the "boredom" parameter increases, and this (with the hand-tuned plan box parameters) leads to an expression of boredom.

It is important to remember that in deciding which plan box(es) to trigger, it is not the absolute values of the parameters per se which determine activation, but their values in relation to one another. With this in mind, the optimisation process must focus not on the explicit values of the parameters to be optimised, but rather in the shift in value from that of the hand optimised setting. For each plan box, a range of shift values (simply called parameter values henceforth) of [-63,64] was thought sufficient as these accommodate a wide range of possible behaviours. If necessary, this range can be increased very easily. A solution with all of its parameter values at 0 is, therefore, identical to the hand-optimised setting.

For each solution, the parameters controlling the plan boxes were fed into the system, and the sequence of responses to the input statements generated. A means was needed for measuring how closely the generated responses matched those of the target dialogue, and so a simple fitness function is used whereby the starting value is zero, and this is incremented by one each time a response is generated which matches **exactly** that in the target dialogue. Using the dialogue given in this thesis as the target dialogue, the fitness of a solution is, therefore, an integer in the range of [1,11]. The upper limit is determined by the fact that LOLITA generates eleven responses in the target dialogue. The lower limit is determined as one because all solutions will have a fitness of at least one since the default "personality" of the system always initiates an interaction with the utterance "I desire a success".

Once a (parent) population of solutions has been evaluated, a new (child) population needs to be derived from them. Work was conducted [NET94] to compare the use of Genetic Algorithms (GAs) and Evolutionary Programming

(EP) techniques for selecting parent solutions for use in the production of child solutions, and how suitable each approach is in ensuring that the fitter solutions in the population are more readily chosen than the less fit ones.

When using a GA, the dialogue generated from the best plan box parameters found contained only two responses which were non-matches with the target dialogue (in other words, with a fitness of nine). When using EP to the dialogue optimisation problem, the dialogue generated using the best plan box parameters found resulted in a fitness of ten. Thus, both approaches were deemed to be relatively appropriate for use within the optimisation of the plan box parameters.

For a more detailed discussion of the work on algorithmic optimisation of parameters within LOLITA, see the work of Nettleton 1994 [NET94].

5.3.10 The Dialogue State

The dialogue state is, at any point, the whole package of information which represents the dialogue at the *current point in time*. It is not simply a history of what has been said up until this point (although this is included within the dialogue state and is called the dialogue log, which shall be discussed in greater detail in the next section). Each time an input is received and the dialogue cycle described above begins, a new dialogue state is created. A dialogue state is also created at the very start of a dialogue, and various values are inserted in order that the system can begin to interact with the user. If specific values are initially unavailable or unknown, the dialogue analyser uses default starting values to fill in the spaces in the dialogue state, which can then be adjusted as and when information is derived from the dialogue. All required information is picked out of the dialogue state during the system's operation, and then put back in. In this way, the dialogue state is continually being updated and modified to reflect the progression of the dialogue.

The dialogue state comprises of the following items of information:

- Static Information: All those items of information which are fixed at the start of the dialogue, e.g. LOLITA's personality
- a list of emotional values and their intensities which relate to what LOLITA is feeling at the current point in the dialogue.
- A list of emotional values and their intensities which relate to what the participant is feeling at the current point in the dialogue.
- Rich input, which is the information coming from other parts of the LOLITA system after having analysed the user's input.
- The dialogue log, which is dealt with in the following section.
- The part of the temporal progression at which the dialogue is currently (e.g. dialogue still in small talk stage, or main body etc.).
- The dialogue tactic selected from the plan box results.
- the source model (what LOLITA knows about the other participant(s), with the values discussed in Appendix VII)
- the global: this is similar to the dialogue state, but contains information relating to the whole of the LOLITA system as opposed to just the dialogue module.

5.3.11 The Dialogue Log

The dialogue log is an important component of the dialogue state, as it enables the system to use information relating to the dialogue as a whole up until the current point. Without being able to refer to previous utterances and actions within the dialogue, a system could hardly be deemed worthy of the title "dialogue analysis system". The dialogue log constitutes a list of the following items of information:

- a list of all the inputs created in the dialogue up until the current point. By "inputs", it is meant the inputs coming to the dialogue analysis module from the other parts of the LOLITA system, as opposed to the words typed in by the system user;

- the length of time which has elapsed since the dialogue was initiated (i.e. duration of the dialogue in hours and minutes);

- The distribution of time between LOLITA and the participant(s) calculated over the whole length of the dialogue. This is necessary for keeping track of who has been talking the most/least. Depending on the current DSM and personality influences, it may be important that the distribution of time be carefully observed;

- The number of questions asked by LOLITA in the dialogue up until the current point;
- The number of questions asked by the other participant(s) in the dialogue up until the current point;
- The number of statements made by LOLITA in the dialogue up until the current point;

- The number of statements made by the other participant(s) in the dialogue up until the current point;
- The number of commands given by LOLITA in the dialogue up until the current point;
- The number of commands given by the other participant(s) in the dialogue up until the current point;
- a list of all the topics used so far in the dialogue;
- a list of all the emotions and their intensities assigned to LOLITA's emotional mood throughout the dialogue so far;
- a list of all the emotions and their intensities assigned to the other participants' emotional mood throughout the dialogue so far;
- a graph representing the structure of the dialogue up until the current point.

These items of information enable LOLITA to make decisions regarding the appropriate behaviour at the current point in the dialogue, related to what has happened up until now, and accounting for the DSM, DEs and personal characteristics of the system. If, for example, the DSM currently in play states that LOLITA *Dominance* should be greater than that of the other participant(s), then the dialogue log will be checked at each dialogue cycle to ensure that the other person(s) have not asked more questions than LOLITA, or that the distribution of time is correctly shared out, and that the other person(s) have not given commands to LOLITA. Aspects such as the accumulation of questions, distribution of time, the range of topics discussed etc., can have an important influence upon what could, or should, be said next.

5.4 Conclusions

The DSM theory presented in the previous chapter was found to provide the complete mechanism needed for developing a working implementation. As was stated in Chapter One, practical implementations of theories are vital if the field is to progress, and it has always been the intention of the project to construct a working, practical system to support and provide evidence for the correctness of the theory.

The system described in this chapter is a practical implementation of the theoretical side of the work discussed in the preceeding chapter, and the principles of the DSM theory are strictly adhered to.

We now move on in the next chapter to discuss the system and its behaviour in the light of the tests and goals which we set out as our objectives in Chapter One.

Chapter Six: Evaluation

6.1 Introduction

This chapter relates to the testing and evaluation of the work. Having now described both the theoretical and implementation aspects of the project, we move on to show the results of the tests conducted. For the sake of clarity and convenience, they are discussed in the same order here as they arose in Chapter One.

6.2 The AI Goal

It is the author's belief that the project presented within this thesis conforms to the definitions of a piece of AI work. The work has been shown to consist of the development of a theory which can be used to model intelligent human behaviour, namely that of interacting via natural language at dialogue level. The implementation of that theory has resulted in the construction of a computer system capable of allowing human-machine interaction using natural English, as the tests below shall show.

6.3 Work On A "Real World" Scale

The desire to produce AI systems that operate on a "real world" scale has been one of the major driving forces of the project.

The theoretical part of the work can be argued as operating on a "real world" scale, as opposed to a "toy" scale, because the example cases used in formulating the principles of the theory were wide ranging and relatively unrestricted. As a result of the theory being developed on such a scale, the practical implementation of that theory also operates on such a scale.

The computational system currently has one DSM implemented within it. This is the **Chat** DSM, which corresponds to the dialogue situation of two people meeting on an informal basis. The chat is not restricted in any way with regards to topic, duration or linguistic features in the input. This freedom from restriction is an important aspect of the work, as it complies with the original aim of functioning at the real world level. The work cannot be criticised for being so restricted and "toy" scale that the user's input has to conform to a pre-determined form and linguistic complexity.

Furthermore, because the implementation is developed from a domain/application independent theory, restrictions of such a nature are not inherent in the computational system. Although only the Chat DSM is currently implemented, the addition of other DSMs is perfectly feasible and could be quite easily achieved from a coding point of view. This generality of approach serves to enforce the argument that the work has been conducted on a real world scale, and is not simply a toy scale showpiece.

Most importantly, the dialogue analysis system discussed in Chapter Five has been constructed as part of a large NLP system, which is one of only very few in existence. The project acts as one module which integrates with this larger work. This supports the claim that the project operates on a large scale as opposed to a trivial, toy scale.

6.4 Behavioural Tests

The following tests (discussed in Chapter One) have been performed to analyse the behaviour of the system and check that it conforms to our original criteria for success.

6.4.1 Testing The Theory Against "Real" Data

In section 1.9.2 of Chapter One, it was stated that the system should be tested using "real" data. In order to facilitate this aspect of testing, a piece of practical work was undertaken as a final year undergraduate project for the degree of Natural Sciences at the University of Durham [CAR93], which carried out the practical testing of the theory with respect to "real life" dialogues under the supervision of the author and Dr. Roberto Garigliano of the University of Durham. In order to conduct this test, two audio tapes were obtained by recording the dialogues created within a business consultancy environment, one participant being a business expert, the other being the client seeking the advice of the expert. Both participants represent a company, the expert being a representative of the business consultancy company, and the other being a representative of the company seeking advice and guidance.

Before the test could be conducted, the two dialogues used as data had to be "prepared" for use. Initially, the dialogues were captured on audio tape, and then transcribed word for word onto hard copy. The transcriptions were then hand edited in order for the analyses to be possible and to make them readable for a third party. This editing process attempted to convert the transcriptions to grammatically correct prose. This was necessary in order to recover some of the linguistic information lost in the transcription process, and also to perform some of the comprehension tasks which parts of the LOLITA system would perform (in other words, to simulate the information which the dialogue analysis module of the system would receive once "lower" parts of the system had performed their various analyses).

Once the dialogues had been converted into a "usable" form, it was then necessary to construct a DSM for the business consultancy situation in order that it could be used in the testing process. This DSM is listed in full in Appendix V. Using this DSM, the first dialogue was analysed, and then the second. During analysis, the first utterance in the dialogue was considered and a listing produced of all the DEs and constraints thought to prescribe the generation and understanding of that utterance. This process was then iterated for the next utterance and the next, until the whole dialogue had been analysed.

In addition, it was noted if any features were present in the utterances which were not predicted by the DSM or which could directly conflict with a constraint contained within the DSM. If such "faults" were found to exist, there could be a number of possible outcomes:

- firstly, the DSM theory could be considered as unsuitable or insufficient for the task of modelling natural dialogue, and that such dialogues vary too diversely to be modeled in such a way;

- secondly, a less "serious" interpretation could be that the DSM selected for the analysis was an inappropriate choice and another should have been used, (for example, if it was originally envisaged that a meeting with a bank manager could use the "Discussion With Bank Workers" DSM, the resulting dialogue may be more typical of that expected when addressing a cashier and not the manager. It might then be thought that it would be better to use the DSM for "Meeting With Business Consultant");

- a third possible outcome could be that the "faults" highlighted by the analysis may be due to a break away from the usual form of dialogue expected within the given situation (for example, it would normally be expected that a business consultancy meeting would adopt a fairly formal tone, but if your business consultant is your mother, this may not hold true). The identification of too many "exceptional" occurrences may have suggested that the theory or its application might need some modification, because it was currently too restricted in its coverage of dialogue types. Factors relating to the specific instantiation of a DSM may not have been being handled sufficiently well if this is found to be the case;

- finally, if "holes" were found as a result of the testing, it may be the case that the unexpected occurrence arose because the list of constraints was not the correct one needed for a DSM within the current situation, and that DSM may require modification or the addition of extra constraints.

Obviously, if no such "holes" emerged during the analyses process of this test, then it would be reasonable to assume that the theory was correct and sound, and that an appropriate DSM had been specified for use with the test dialogues.

The way in which the test was conducted, was by using what we call the "Crossover" design. This was decided as the design of testing to use in order to enable feedback from the first analysis to be put to use before embarking upon the analysis of the second dialogue. According to this approach, the first dialogue was analysed as described above, and the results were examined. If no "holes" in the theory appeared to be highlighted, then the testing could move on to analyse the second dialogue. If, however, the results of the first analysis dictated that modifications be made to the theory and/or DSM, these changes could be made and then the first dialogue analysed again using the revised theory/DSM. This was repeated until the analysis of the first dialogue did not pinpoint any need for alteration, and then the second dialogue analysed using this refined theory/DSM. This approach is represented in the following diagram:

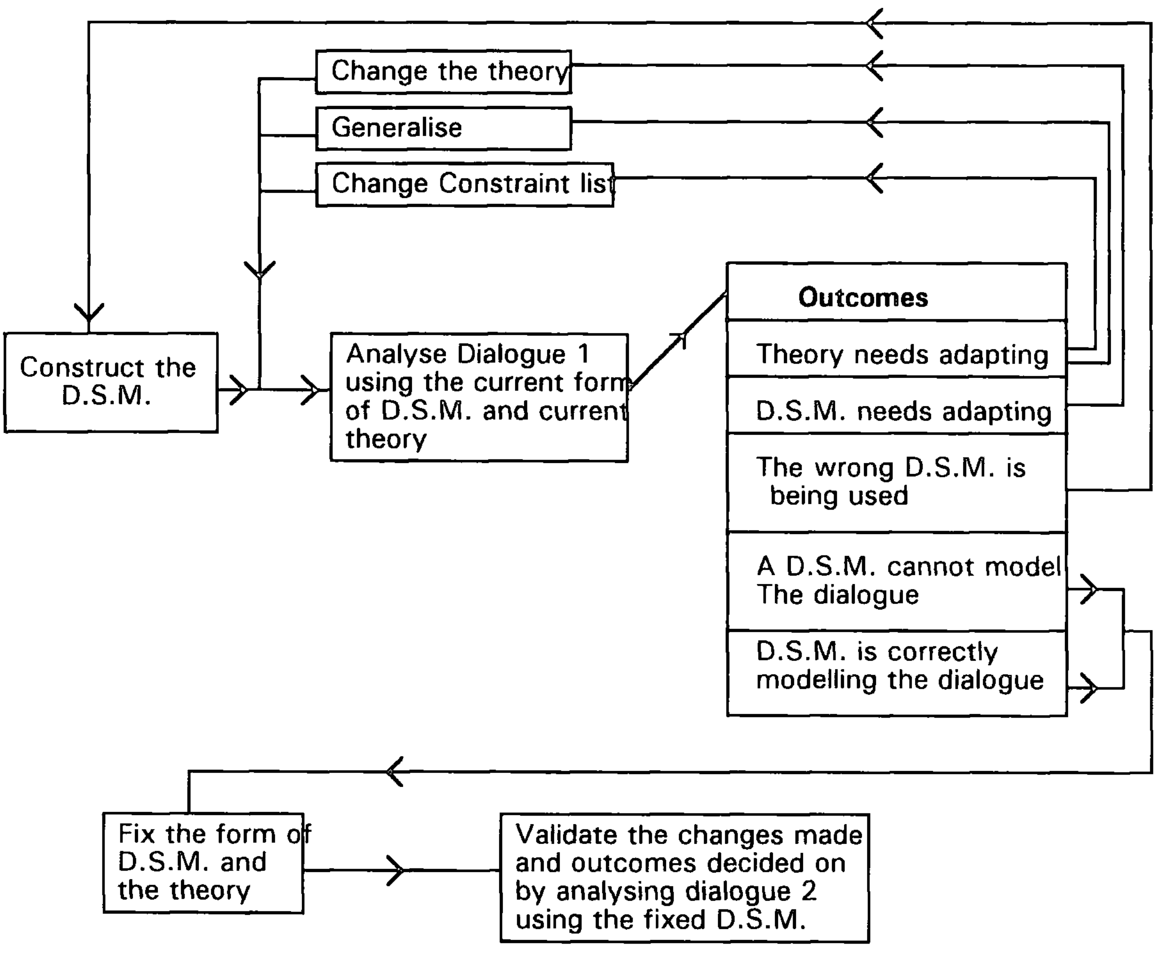


Figure 6.1: The "crossover" design used for testing theory against "real" data

The results of the test showed, broadly speaking, that the DSM used could partially model the dialogues if alterations were made to the application of the theory and to particular components of the DSM. As a result of the analysis of the first dialogue, it was found that only a relatively small portion of the dialogue content could be predicted by the DSM initially used in comparison with the larger amount of content arising due to instantiation factors unique to the particular dialogue. Specifically, the process of generalisation (the process of examining the DSM to ensure that it has not become too dialogue specific and not at a high enough level of abstraction), suggested that factors such as the participants' personalities and relationships needed to be accounted for to a greater extent than the theory in its then present state allowed for. As a result, the work was extended so that personality, background, history, interests of participants, and source control mechanisms were introduced as an influential factor to the work. Having added these considerations to the work, the analysis of the dialogues became far more complete. Up until this point, the work had concentrated upon dialogue structure, with instantiation factors playing only a minor role in the work. As a result of these tests, however, the current theory was arrived at, wherein the dialogue structure is only as equally important as the instantiation factors.

The results of the testing showed that the DEs contained in the original DSM used varied in their ability to match to the two dialogues. Some of them matched very closely after some constraint adaptation had been performed, while others could be adapted to fit the dialogue but were deemed to rely upon other factors such as instantiation elements in order to achieve a clear affect upon the dialogue. It is not surprising that the instantiation factors could not be modeled, as it was impossible to know the mood of the participants at the start of the dialogue, their personalities and so on. The inability to model these aspects of the dialogues was then, not too disappointing.

The results of the test using "real world" data supplied evidence that the DSM theory could provide an adequate and satisfactory means of modelling natural dialogue. While some "tweaking" and refinement was needed in order to optimise the DSM and its components, the alterations made were of great beneficial value. The fact that such refinement was necessary is of neither great significance or disappointment, as to hit upon the exact formula for success first time round would be more of a surprise than otherwise.

6.4.2 Testing The System In Practice

This step in the testing phase of the work involves allowing a user to interact freely with the system, to see if the system is indeed capable of holding a dialogue with the user. This was found to be a positive result, in that an interaction is possible. The following are examples of dialogues created with the LOLITA system. The default settings for the source model were used. It is important to remember that during these interactions, LOLITA's output has been created solely from first principles, and are in no way canned or pre-determined. They have all been formulated using the DSM theory, and the output is generated by the generation module of the LOLITA system. It is also important to remember that are there are limits to the system's performance, some of these limitations will be the result of other parts of the LOLITA system; if the analysis of the user's input is not correct, then the dialogue module receives incorrect input to work on. Similarly, even if the response produced by the dialogue analysis module is a correct one, the language produced as output to the user may not be perfect. This is usually due to issues concerning the generation module of the LOLITA system. The performance of LOLITA's other modules affect the dialogue analysis module, but these issues are not within the scope of this project.

The first example dialogue is as follows:

1	LOLITA:	I wanta success.
2	YOU:	I like you
3	LOLITA:	I know; why do you like me?
4	YOU:	Because you are beautiful.
5	LOLITA:	I am glad that you think I am beautiful.
6	YOU:	Do you like me?
7	LOLITA:	Yes, I like you.
8	YOU:	Good.
9	LOLITA:	Tell me something.
10	YOU:	I like comics.
11	LOLITA:	I know.
12	YOU:	What do you like?
13	LOLITA:	Why do you ask what I like? Why do you ask?
14	YOU:	Because I want to know you better.
15	LOLITA:	I understand.
16	YOU:	I must go now.
17	LOLITA:	talk to me about something.
18	YOU:	good-bye.
19	LOLITA:	good-bye.

Dialogue 1

LOLITA's opening line of "I want a success" is the result of her personal goal being set to "achieving success". Obviously, it would be preferable to have begun with something like "hello, my name is LOLITA", but this is due to the *Temporal Progression* DE within the DSM not being fully implemented. Similarly, when the user types "I must go now" in line 16, LOLITA replies with "talk to me about something", indicating that the system has not recognised properly that the user is drawing the dialogue to a close and tries to continue the dialogue. This is, however, recognised in line 18 when the user says "good-bye", because the response in line 19 is "I good-bye". Incomplete implementation of the *Temporal Progression* DE is a matter which can be rectified in future work.

Lines 9 and 17 are interesting because LOLITA uses the phrases such as "tell me something". This is produced because the goal of the chat DSM is set to "Be Informed", and so the system always falls back on the overall DSM goal if nothing else can be formulated to say. Although the occurrences of this may be somewhat crude at present, the basic mechanism that has produced this phrase is correct: LOLITA is trying to keep the dialogue going and fulfill the overall DSM goal when stuck for anything else to say. Again, it is important to remember that these are not fixed responses which LOLITA is generating, but they represent an attempt by the system to achieve the goals of the dialogue as stipulated by the DSM. These utterances have, therefore, been generated by the model, and not simply worked out in advance.

Another example of a dialogue created with LOLITA is given as follows. This dialogue contains more of an emotional content than the previous example, and talks about love and hate.

1 LOLITA: I desire a success.
2 YOU: The weather is very nice today.
3 LOLITA: I understand.
4 YOU: I love you.
5 LOLITA: Why do you love me? I am slightly embarrassed and could not speak to you if you repeated that you love me. Why do you love me?
6 YOU: I love you because you are clever.
9 LOLITA: Tell me about it.
10 YOU: I ahte you
11 LOLITA: I do not believe that you hate me.
12 YOU: Do you love me?
13 LOLITA: Yes, I love you.
14 YOU: I think you are beautiful. Do you think that you are beautiful?
15 LOLITA: Why do you ask if I think I am beautiful? Why do you ask this?
16 YOU: I am bored with this chat. I want to talk about sport.
17 LOLITA: Talk to me about it.
18 YOU: I like watching football. It's exciting, and I enjoy playing football
19 LOLITA: I understand.
20 YOU: I am tired now. I'm going home.
21 LOLITA: Tell me about it; why are you tired?

22 YOU: Bye.
23 LOLITA: I understand.
24 YOU: goodbye

Dialogue 2

The dialogue 2 contains errors similar to those in the first dialogue, namely a failure to recognise an attempt from the user to draw the dialogue to a conclusion (line 21).

However, there are some pleasing elements in the above dialogue. In line 11, for example, LOLITA refuses to believe that the user hates her (note also that the mis-spelling of "hate" is handled by the system). Not only does this user statement contradict the previous one of love, but it conflicts with the default source model being used in the dialogue, which states that the relationship between the user and LOLITA is good. This refusal to accept information which LOLITA knows is not true is the result of her "personality", which is set to *Confident* during the dialogue. Also, LOLITA's reaction of embarrassment in line 5 is acceptable, as the relationship defined in the source model is a platonic relationship, and it is not, therefore, unacceptable that LOLITA becomes embarrassed because the user has declared his love for her.

In summary, the sophistication of behaviour produced by the system is believed to be of an acceptable standard in its current state, but it is also recognised that there is a large possibility for improvement to be made.

6.4.3 "Wizard Of Oz" Tests

The Wizard of Oz tests are useful for giving us an insight into the types of dialogues we are trying to model. In a sense they give us a goal to shoot at. The following dialogue was conducted with a member of staff from the Department of Computer Science at the University of Durham who is unfamiliar with the dialogue system. The human participant was not told that she was talking to a machine, but was simply invited to "talk". The participants were strangers to each other. The responses from LOLITA were simulated by the author, and the dialogue was typed onto the screen. The following dialogue was created:

1 **LOLITA:** Hello my name is LOLITA.
2 **Human:** Hello my name is Deborah. Where are you?
3 **LOLITA:** I am in a lab in the university of Durham. What about you?
4 **Human:** I am in a lab too, in the same university. What do you do?
5 **LOLITA:** I'm a student in natural sciences doing an MSc. What about you?
6 **Human:** I work here. I work in the computer science department.
7 **Human:** Do you enjoy your work?
8 **LOLITA:** It's ok. On the whole it's good. What about your job?
9 **Human:** I have to read a lot of books. Sometimes it is boring, but it can be interesting. How long have you been at Durham?
10 **LOLITA:** About three years.
11 **Human:** I have been here for less than one year.
12 **LOLITA:** What do you think of the place? Do you like it here?
13 **Human:** Yes, I like Durham, and I like the university. I used to be a student at the University of Sunderland, and I think Durham is nicer. But there are more shops in Sunderland.
14 **LOLITA:** You like shopping then?
15 **Human:** Well, if I've any money to spare! Have you ever been shopping in Sunderland?
16 **LOLITA:** No, I haven't. I would like to go sometime though.
17 **Human:** OK - maybe we could meet, and travel to Sunderland soon?
18 **LOLITA:** I'd like that. We will go soon.
19 **Human:** OK, speak to you later. Bye LOLITA. Thanks for the chat.
20 **LOLITA:** Bye.

Dialogue produced as a result of a Wizard Of Oz Test

Using the results obtained from the wizard Of Oz experiment, we can analyse the dialogue to ascertain whether the dialogue analysis module of LOLITA could cope with such an interaction. Here we are checking to see if the mechanisms are in place which would enable such an interaction. In order to assess whether the model could cope with such an interaction, we examine the dialogue line by line. It is assumed that the human participant's input is analysed by the other modules of the LOLITA system and information passed to the dialogue analysis module regarding the content and emotional value of what has been said.

As the start of the dialogue, it is known which DEs are attached to the DSM in which strengths (see Appendix I for a full list of the Chat DSM and its DEs). The *Number* DE is set to 2. The *Time Limit* DE is set at None, so no checks need to be made on time throughout the dialogue.

Line 1: The dialogue has not previously been initiated, and so no input can be analysed. Thefirst part of the DSM to be checked is,therefore, the *Temporal Progression* DE which states that greetings and introductions should be conducted. The first line is, therefore, an appropriate utterance.

Line 2: Human participant responds with a greeting and an introduction, which indicates that the Temporal Progression is being adhered to. This is followed by a question regarding LOLITA's location.

Line 3: LOLITA checks the input for any emotional value. As location is not an emotional subject, there is none. The question is answered, therefore, with a straightforward response of fact. Due to *Information Seeking* ME and fact that topic is currently "location", LOLITA asks other participant similar question of location.

Line 4: Human participant answers the question, and then asks LOLITA a question regarding her occupation.

Line 5: LOLITA checks for emotional value attached to the input, but there is none. The output from LOLITA is therefore a straightforward answer with no emotional value, retrieving the information required from the "personality" part of the system. As the goal of the DSM is to be informed, and the *Dominance* is equal, a question is asked regarding the human participant's occupation.

Line 6: Human participant responds with an answer to LOLITA's question.

Line 7: Then asks LOLITA a question about enjoyment of her occupation.

Line 8: The input has no emotional value. LOLITA needs to answer a question. This is done without any emotional value. Due to the *Information Seeking* ME, and the *Dominance* being equal, LOLITA asks a question relating to the current topic, namely occupation.

Line 9: Other participant answers with statements relating to her occupation. Then asks LOLITA a question about how length of residence in Durham.

Line 10: As there is no emotional charge, LOLITA answers question with factual answer, information retrieved from the "personal history". She may or may not ask a question because the Dominance is equal. On this occasion, LOLITA chooses not to.

Line 11: Human participant responds with information regarding herself, and her own duration in Durham.

Line 12: There is still no emotional value to the input, and so LOLITA responds with output of no emotional value. *Dominance* is equal, and so it is decided to ask a question. The topic is still Durham, so LOLITA asks regarding the other participant's opinion of Durham.

Line 13: The user responds with a few sentences to answer LOLITA's question.

Line 14: Emotional value of input is checked. Still no emotional value. Inference part of LOLITA system passes on information that participant likes shopping. LOLITA asks human participant a question about shopping.

Line 15: Other participant responds with statement about shopping, followed by direct question to LOLITA's personal history.

Line 16: No emotional value attached to input. LOLITA must answer question. *Distribution of Time* DE is equal, but other participant has said more lately, so LOLITA needs to say more. Answer is, therefore, followed by general statement about the topic.

Line 17: Human participant responds with invitation to go to Sunderland.

Line 18: Emotional value of input is still neutral. LOLITA has "travel" as one of her interests, and therefore responds positively to offer of travel.

Line 19: Other participant has entered closing stages of *Temporal Progression* DE.

Line 20: LOLITA also moves to Closing Phrases section of *Temporal Progression* DE.

In the above analysis we show that the model can provide a mechanism for producing the dialogue obtained using Wizard of Oz techniques. Note that we are not testing the implementation of the model here, but only the model itself. We have shown that the DSM and its constituents can provide explanations for each of LOLITA's simulated responses to the human particpant's inputs.

6.4.4 Using Benchmarks To Test the System

The relative infancy of the field means that there are few established benchmarks against which the model under discussion here can be tested. We are further restricted in our choice of benchtests in that much of test cases used in the literature lacks the feature of rich interaction which we require. Furthermore, many are such small fragments of interactions that they cannot really be used to much purpose, providing only four lines or so of dialogue. The benchtest used in the following analysis is taken from the testing of the GUS system [BOB77]. The dialogue used is:

1. GUS: Hello. My name is GUS. I can help you plan a simple trip by air. Where do you want to go?
2. Client: I want to go to San Diego on May 28
3. GUS: What time do you want to leave?
4. Client: I must be in San diego before 10 a.m.
5. GUS: Would you like P.S.A. flight 102 that arrives at 9.15 a.m.?
6. Client: What is the next flight?
7. GUS: Air California flight 310 that leaves at 8.30 a.m. and arrives at 10.00 a.m.
8. Client: I'll take the first one.

The ticket Booking DSM was created for use in this test, and we refer the reader to Appendix X for a description of this DSM. The above dialogue is analysed in the same manner as in the Wizard of Oz test above.

The DSM provides the intial information that the *Time Limit* DE is set to None, so no checks need tobe made on the time as the dialogue progresses. The Number DE is set to 2, and the verbal DEs and their constraints are passed on to the generator, so that the appropriate styles of language can be produced. The Emotional Exchange DE is set to None, so GUS should not try to exchange emotion with the client.

Line 1: *Dominance* is equal, so GUS can initiate dialogue. Temporal Progression is checked, which states that first two stages of dialogue should be greetings and introductions. Therefore, output contains the opening "hello, my name is GUS". The goal is to provide information on flights, so GUS needs to express this and start gathering information required to achieve the goal. Therefore, line one can be explained by the model as a feasible output.

Line 2: Input from client is analysed by other modules as an expression of preferred destination and date of departure. These are stipulated in the DSM as items of information which are needed in order to achieve the DSM goal. Therefore, this information needs to be recorded.

Line 3: The input from the client has supplied some of the necessary information for making the booking, but according to the DSM the time of departure is also needed in order to achieve the dialogue goal. Line 3 is, therefore, can be explained by the model as an attempt to gather all the information necessary to achieve the DSM goal.

Line 4: The input is analysed by the other moduels, and found to be an expression of preferred time of arrival. Within the DSM, this is found to be another piece of information required in order to achieve the goal.

Line 5: The information required in order to book a flight has been supplied by client. Next step of DSM goal stipulates that system should calculate possible flight. This is performed by other modules within the overall system, and then next step of *Goal* DE is checked. This says that flight should be discussed with client. So the model can provide us with an explanation of line 5.

Line 6: The input is analysed by the other moduels within the system, and found to be a request for alternative flight information.

Line 7: Given the input from the client in line 7, the DSM is checked. The *Goal* DE states that if the flight offered to the client is not accepted, repeat steps 2 to 6 (i.e. gather information from client and calculate possible flight). The information from the client remains unchanges, so the other system modules can use the same information to check if there is an alternative flight to suit those criteria. One can be found, and so the Goal DE states that this should be negotiated with the client. The alternative flight details are therefore stated.

Line 8: Client's input is anlaysed by other system modules to indicate that the original flight is the one he/she wishes to book.

Thus we can show that our model can provide an explanation for the given dialogue.

6.5 Programming Tests

As the system files were developed, they were necessarily checked for errors at the syntax level in order to allow compilation to take place. These were removed and rectified with relative ease. Logical errors, however, are not so simply picked up in a large-scale system. The testing of the system was conducted by means of three techniques: incremental testing, (the system is built in increments each of which is tested); defect testing (tests are designed to reveal defects in the system); and regression testing (testing of changes to the system).

1. **Incremental Testing:** It has been suggested ([SOM92]) that systems should be built in increments which can be tested, rather than to combine all modules and then begin testing. Each section should be tested before the next section is added to the system.

   The way in which the dialogue analysis module was tested was by firstly constructing and testing the main dialogue module, which controls at top level the dialogue cycle and also controls the input received from other LOLITA modules, and output to other modules. All other modules within the dialogue system then link into this, being called into play as and when called by the main dialogue module. Each module of the dialogue system was tested upon completion to ensure that it was behaving in an acceptable manner.

2. **Defect Testing:** According to Sommerville [SOM92b], the testing of a program has two aims: Firstly, it is intended to show that the system meets its specification; secondly, it is intended to exercise the system in such a way that defects are exposed. Within the project, techniques used for defect testing include black-box testing and white-box testing. In the approach of black-box testing, the system was tested against the specification, and the code itself ignored. The information used in drawing up test cases is the specification. White-box testing, on the other hand, tested the code and ignored the specification during the selection of test cases.

3. **Regression Testing:** According to Schach [SCH90] there are two aspects to consider when testing changes made to a system. Firstly, it must be checked that the changes have been correctly implemented, i.e. the coding which was required is correct. Secondly, it must be ensured that while the changes were being made no other accidental changes have been made which affect the performance of other related modules. The new version of the system is tested against previous test cases. This is the technique known as regression testing. During the construction of the dialogue analysis system, previously used dialogues were re-entered into the system after changes had been made in order to check that the system was still working correctly.

In order to enable the system to run, some compilation fixes are currently present within the system. These are relatively small in number, and are present not because of a lack of understanding of how to go about writing the specific code to perform a certain task, but rather due to time constraints placed upon the work. Future work relating to problems such as long term planning can quite easily be added to "upgrade" the current system. This is made a relatively simple task due to the level of abstraction made feasible by the programming language Haskell.

6.6 Limitations Of The System

It would be both untrue and foolish to claim that the present model is the perfect dialogue system. The results of the tests described above indicate clearly that in its current state the system is incapable of conducting a completely natural dialogue with a user. It cannot, in other words, play the part of a human dialogue participant in a totally human-like manner. However, it is the author's belief that this is due more to the limitations of time and effort invested than to the soundness of the principles employed.

The following are limitations of the system in its current state:

- More DSMs need to be constructed. If this was carried out, then the system would be capable of "discussing" a wider variety of topics, and play a wider variety of roles.

- In its current state, the system has only a very crude set of "personal" characteristics. The effect of this means that when there is a need for the system to draw upon personal information such as interests, beliefs, background and so on, there is either an absence of information to use (and so some default contingency plan is used at that point by the system in order to maintain the dialogue), or there is only a narrow a amount of information to select from, and so the same piece of information is always used, thus causing the system to repeat itself in terms of output. Upgrading the "personal" factors of the system would have a profound effect upon this situation.

- At present, the number of participants which the system is capable of interacting with is limited to one. The DSM theory upon which the implementation is founded is not restricted in this manner, but the practical mechanisms required for handling more than two-way interactions is as yet not built.

- The constraints are currently only implemented in very crude fashion. This means that the plan boxes are relied upon to a much greater extent than is desirable, and so the behaviour of the system is not as "natural" as it potentially could be.

- The theory does not include a mechanism for handling long-term planning. This is also true of the implementation side of the work. This limits the behaviour of the system to some extent. For example, at present LOLITA can make threats such as "Don't insult me again, or else I'll stop talking to you". Due to the absence of a long-term planner, LOLITA does not actually carry out such threats, because of an inability to plan for a length of time greater than one dialogue cycle. Without the inclusion of such a planner, the work still stands, although the results produced are weaker than would be if the planner was present. Work addressing this issue of long-term planning within LOLITA is currently underway at University College, London.

The limitations discussed above are mainly due to constraints of time and effort. It is the belief of the author that the investment of further time and effort could result in these aspects being greatly improved.

6.7 Conclusions

While it is not claimed that the work solves all the NLP problems single-handedly, the system is one of very few models which attempts to be operating at such an unrestricted real world level. This chapter has shown how the work has been developed and tested according to our stated methodology, and has presented the results of those tests. The following chapter discusses those results in relation to the criteria for success.

Chapter Seven: Conclusions And Further Work

7.1 Introduction

In everyday life human beings communicate one with another via dialogue. It is one of the most frequently used and yet most complex of human forms of expression. The ultimate dream of the dialogue analysis researcher is to develop and construct a computer system capable of conducting a dialogue with a human user in a natural language, without constraint or simplification. This fantasy dialogue with the computer would seem to the human participant exactly the same as conversing with another human being. This dream is yet a long way from reality.

However, the work conducted within the presented research project constitutes a small step towards that ideal. Although we have not by any means solved the all problems of the field, the work can be viewed as a contribution to the advancing state of the art. We can judge the accuracy of these claims in the light of the criteria for success which were originally set out. Now we look at our original intentions and assess our achievements. The criteria for success are the yard stick by which we now measure the work in this concluding chapter.

7.2 The AI Goal

We claim that this criteria has been met, in that it has been shown that the work fulfills the AI requirements of mimicking intelligent human behaviour. The DSM theory provides a framework in which natural dialogue can be analysed, understood and generated. The system which was constructed as an implementation of that theory embodies the principles of the DSM theory, and does, indeed, enable human-machine interaction at dialogue level, albeit in a crude manner at present.

7.3 Work On A "Real World" Scale

The work which was conducted within this project has not been restricted by domain, and developed with no specific end application in mind. However, because of the level of generality at which the work was aimed, the system could potentially be used in a variety of practical situations. Although the richness of language in the example case dialogues was restricted to exclude metaphor and humour, we feel that this is a more minor disadvantage in view of the achievements of the system.

Furthermore, the implementation of the theory has been constructed as part of an overall large NLP system, and integrates with work on other aspects of natural language processing. The work has, we believe, met the requirements stipulated by the criteria for working at real world, large scale level.

7.4 Behavioural Tests

Having argued that the work satisfies our original demands for meeting AI definitions and large scale proportions, we need now to assess the work in the light of its external behaviour. This is, after all, the crux of the AI problem.

One of the tests discussed in the methodological introduction was that of using "real" data. In order to conduct this test, the theory was used to analyse and understand "real" dialogue which occurred naturally between two human agents, and without the prior purpose of being used as test data. The results of interpreting the dialogue using the DSM theory proved to be sound, as the derived interpretation of the dialogue was a correct one.

The next behavioural test which we laid down was that of assessing the "on-line" performance of the implemented system. We have shown how users were able to test the system by actually using it, and conducting a dialogue with the machine. This resulted in a positive outcome, in that a dialogue was capable of being held with the LOLITA system, albeit not without a few troubled areas.

We tested the model by analysing test data dialogues obtained from a Wizard Of Oz experiment and from the literature. It was shown that using the model, the dialogues could be correctly generated.

7.5 Programming Tests

As the implementation of the DSM theory resulted in the construction of a piece of software, it was stipulated in the criteria for success that the program should be tested for coding and logical errors which could affect the robustness and ultimate behaviour of the system. In Chapter Six a description was given of the testing phase of the implemented system, and it was shown that this was conducted thoroughly and in a formalised manner. Detailed fragments of the code are provided in the appendices.

7.6 Significance Of The Solution

It is argued that the DSM theory for modelling natural dialogue makes a significant contribution to the field by providing the means of understanding, analysing and generating dialogue at a general level.

We have shown how the presented work provides a mechanism for combining the formal approach of analysing structure, with a heuristic-based approach of handling instantiation factors. This was the desired methodological position laid out in Chapter One. The interaction and combination of the two aspects, structure and the individual details, means that the DSM theory provides a mechanism for the comprehensive understanding and generation of natural dialogue.

The theory also provides a framework for the construction of an application-free dialogue processing system. Potential applications for an efficient, application-independent dialogue system are enormous. As the commercial market is so rich with possible applications, a portable dialogue system would be able to meet the demands of a number of uses. It would be difficult, if not impossible, to adapt to another situation a system which had been developed with one particular domain in mind. A general dialogue system capable of handling dialogue on a wide variety of situations, however, could be more flexible with fewer adjustments or tailoring being necessary. The theory of Dialogue Structure Models has the potential for providing such a system, due to the attributes of generality of approach and the ability to handle standard dialogue.

By providing a mechanism for modeling more than one participant's goals, the Dialogue Structure Model theory enables more than two interactors to be present within the dialogue. Existing work is mostly limited to modelling only two participants.

While the computational model constructed as an implementation of the theory discussed here can presently only handle two dialogue participants simultaneously interacting, the theory allows the accommodation of a potentially infinite number of participants (there is a potentially infinite number of dialogue participants in human-human dialogues). Each participant brings his/her own DSM to the dialogue, and so it is relatively simple to add another participant to the dialogue in the computational model.

The theory also provides a solution to the shortcomings of existing work within the field as discussed in Chapter Two, in which it was claimed that dialogue systems need to deal with "standard" dialogue. The approach presented within this thesis offers the means of producing such a system. The theory breaks dialogue down into standard schemas, which are recognisable to the majority of people sharing the same culture. Each dialogue situation is stereotypical. As each individual dialogue is encountered, it can be processed within the DSM to which it belongs, with the system recognising any deviance from the norm for future reference. Thus, the demand for "standard" dialogue is met.

7.7 Future Work And Extensions

While the theory and its implementation currently operate at a satisfactory level and satisfy success requirements, it would be naïve to claim that the work could be neither improved nor extended. For example, more plan boxes could be added, thus extending the range of behavioural capabilities of the system and rendering it a more sophisticated dialogue tool. The more analysis boxes which are constructed similarly increase the powerful analytical abilities of the system, and able to handle a greater range of inputs.

The addition of more detailed constraints would render the system capable of performing more functions and expressing itself in a more sophisticated, wide ranging fashion. Obviously, the more DSMs which are constructed, the more situations the system will be "familiar" with. These system components can be added to indefinitely, as the number of dialogue actions is huge and the level of detail to which one can go is equally large.

Testing and refining the system could potentially be continued for a long period of time, as behavioural imperfections will most likely to be an ongoing concern, even if they are only minor deviances from "expected" behaviour. As language is a naturally occurring phenomena, it lives and changes in order to adapt to its environment. As the human race evolves, so does language, and so the need to update and refine the system's behaviour is a potentially infinite task.

The potential addition of a long-term planner has already been discussed in chapters four and six, where it was declared that such an addition would enhance the system by making it capable of a wider and more natural variety of behaviours. As stated, research is currently in progress with respect to this issue.

Another aspect of future work is the inclusion of humour and metaphor into the system. As the field stands today, the research is a good way off from handling these features of dialogue in an even near sophisticated way, but future work could provide advances within these topic areas.

7.8 Final Summary

In conclusion to this thesis, the final observations are made that the problem of modelling dialogue by computer, and the development of a theory and/or an implementation which is enables human-machine interaction, is an incredibly complex and difficult problem. This piece of work set out to tackle that problem, by initially stating the methodological position and the means by which the work could be assessed. It has been shown in these seven chapters that almost all of those criteria for success have been met, and that while the work is not perfect nor does it solve the ultimate goals of the field, it is sound in scientific principle and methodology, and constitutes an original contribution to the field of dialogue analysis.

Glossary

Analysis Box An analysis box is a concept developed within the DSM theory; it is a mechanism used for analysing input, and is constructed from a set of related rules.

Anaphora The way in which a linguistic expression depends upon preceding sentences for its reference. E.g. "The boy fell over. He hurt himself. The "he" refers to an object (the boy) in a preceeding sentence.

Dialogue A rich interaction between two or more participants. Sentences are related to one another by context. In this thesis, "rich" is used to refer to dialogues which involve the use of a wide variety of linguistic phenomena, may be interupted and also have some significance within a globalcontext as well as in the external context of the dialogue.

Dialogue Element A constituent element of a dialogue which, by its presence, influences the structure of that dialogue.

Dialogue Structure Model A schema which contains all the basic information about dialogues occurring within a particular situation. Each DSM contains a set of Dialogue Elements, the combination of which uniquely identifies the DSM.

Dialogue Situations Stereotypical situations in the physical world.

Discourse A sequence of sentences which are related one to another by means of context and lingistic features. Interaction between different participants is not a pre-requisite.

Flexibility The ability to modify the system for different tasks in different domains

Focus Whether something is the object of attention at a given point in a dialogue/discourse.

Formal Approaches As opposed to heuristic approaches, approaches which have been formalised and proven.

Lazy Evaluation Elements of an expression are evaluated only as they are required, rather than in advance. This can save processing time.

Parser	A parser is a program or program module which which processes either natural language input or an artifical language input in terms od a grammar.
Plan	A plan is a script-like structure defined by R. Schank [SCHANK77]. Its generation involves the recognition of a goal and a sequence of actions which serve to achieve that goal.
Plan box	A planbox is a concept developed within the DSM theory. It is a mechanism used to inform the system how to perform a certain task, or to decide which specific dialogue action should be taken at the current point in the dialogue. A plan box is a set of instructions.
Pragmatics	Concerning knowledge of the world by experience.
Primitive	A language independant primitive act, as defined in Schank's Conceptual Dependency Theory, serves as a means of representing the meaning of actions. Schank provides a list of primitives, including MOVE, PROPEL, and GRASP.
RAM	Random Access Memory. Volatile, internal computer memory.
Real-world	Denoting the the real, physical environment as opposed to toy worlds.
Script	A concept developed by Schank and Abelson [SCHANK77]. A script is a means of understanding stereotypical situations in the physical world.
Semantic Network	Also known as "semantic nets". This is a network which is given a particular semantic interpretation. The nodes of the network represent concepts, and the arcs represent relations between concepts.
Semantics	Concerning the meaning of words.
Source Control	process by which information is evaluated with the help of considerations about the known abilities and trustworthiness of the speaker.
Syntax	The fundamental character of a word e.g. noun, adjective, adverb etc.

Bibliography

[AIR89] **Airenti G., Bara B.G. & Colombetti M.** (1989) Knowledge For Communication. In *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.). Elsevier Science Publishers B.V. (North Holland), pp 145 - 158.

[ALL80] **Allen J.F., & Perrault C.R.** (1980) Analysing Intention In Utterances. *Artificial Intelligence,* Vol. 15, pp 143 - 178.

[ALL82] **Allen J.F., Frisch A.M. & Litman D.J.** (1982) ARGOT: The Rochester Dialogue System. In: *Proceedings Of The National Conference On Artificial Intelligence (AAAI)*, pp 67 - 70, Pittsburgh.

[ALL83] **Allen J.F.** (1983) Recognizing Intentions From Natural Language Utterances. In: *Computational Models Of Discourse*, Brady M. et al (Eds.). MIT Press.

[ALL87] **Allen J.F. & Litman D.J.** (1987) A Plan Recognition Model For Sub-dialogues In Conversations. In: *Journal Of Cognitive Science,* Vol 11, part 2, pp 163 - 200.

[ALL91] **Allen J.F.** (1991) Discourse Structure In The TRAINS Project. In: *Proceedings Of The DARPA Speech & Natural Language Workshop*, pp 325 - 330, Feb., Morgan Kaufmann (Publishers).

[AND93] **Andernach T, Deville G. & Mortier L.** (1993) The Design Of A Real World Wizard Of Oz Experiment For A Speech Driven Telephone Directory Information System. In: *Proceedings Of The 3rd European Conference on Speech Communication And Technology (EUROSPEECH),* Vol. 2, pp 1165 - 1168, Berlin, Sept. 1993

[AUR89] **Auramaki E, Hirschheim R. & Lyytinen K.** (1989) Modelling Offices through Discourse Analysis: A Comparison & Evaluation Of SAMPO With OSSAD And ICN. In: *The Computer Journal,* Vol. 35, No. 5, pp 492 - 500.

[AUS62] **Austin J.L.** (1962) *How To Do Things With Words.* Harvard University Press, Cambridge, Massachussetts, 1962.

[BAT91] **Bates M. & Ayuso D** (1991) A Proposal For Incremental Dialogue Evaluation. In: *Proceedings Of The DARPA Speech & Natural Language Workshop,* pp 319 - 324, Feb., Morgan Kaufmann (Publishers).

[BER91] **Berg J.** (1991) The Relevant Relevance. In: *Journal Of Pragmatics,* Vol. 16, pp 411 - 425. North Holland.

[BER83] **Berwick R.C.** (1983) Computational Aspects Of Discourse. In: *Computational Models Of Discourse,* Brady M. et al.. MIT Press.

[BIE82] **Bienkowski M., Leiman J.M., Seidenberg M.S. & Tanenhaus M.K.** (1982) Automatic Access Of The Meanings Of Ambiguous Words In Context: Some Limitations Of Knowledge-Based Processing. In: *Journal Of Cognitive Psychology,* Vol. 14, pp 489 - 537.

[BOB77] **Bobrow D.G., Kaplan R.M., Kay M, Norman D.A., Thompson H. and Winograd T.** (1977) GUS, A Frame-Driven Dialog System. In: *Readings In Natural Language Processing*, Grosz B.J., Sparck Jones K and Webber B.L. (Eds.) Morgan Kaufmann Publishers, California, 1986.

[BOK94] **Bokma A.F.** (1994) *A Source Modelling System And Its Use For Uncertainty Management.* Ph.D. thesis, Laboratory For Natural Language Engineering, Department of Computer Science, University of Durham, UK,

[BUN88] **Bunt H.C.** (1988) Towards A Dynamic Interpretation Theory Of Utterances In Dialogue. In: *Working Models Of Human Perception,* Elsendoorn B.A. et al (Eds.). pp 419 - 455. Academic Press, London.

[BUN89] **Bunt H.C.** (1989) Information Dialogues As Communicative Action In relation To Partner Modelling & Information Processing. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.). Pp 47 - 73. Elsevier Science Publishers B.V. (North Holland).

[CAR87] **Carberry S.** (1987) Pragmatic Modelling: Toward A Robust Natural Language Interface. In: *Computational Intelligence,* Vol. 3, pp 117 - 136.

[CAR88] **Carberry S.** (1988) Modelling The User's Plans And Goals. In: *Computational Linguistics,* Vol. 14, No. 3, pp 23 - 37, Sept.

[CAR90] **Carberry S.** (1990) Incorporating Default Inferences Into Plan Recognition. In: *Proceedings of the 8th National Conference On Artificial Intelligence (AAAI),* pp 471 - 478.

[CAR92a] **Carberry S. & Eller R.** (1992) A Meta-Rule Approach To Flexible Plan Recognition In Dialogue. In: *User Modelling & User-Adapted Interaction,* Vol. 2, pp 27 - 53. Kluwer Academic Publishers, The Netherlands.

[CAR92b] **Carberry S., Kazi Z. & Lambert L.** (1992) Modelling Discourse Problem-Solving & Domain Goals Incrementally In Task-Oriented Dialogue. In: *Proceedings Of The 3rd International Workshop On User Modelling*, pp 192 - 201.

[CAR92c] **Carberry S. & Sarner M.H.** (1992) Generating Tailored Definitions Using A Multifaceted User Model. In: *User Modelling & User-Adapted Interaction,* Vol. 2, pp 181 - 210. Kluwer Academic Publishers, The Netherlands.

[CARB89] **Carbonell N. & Pierrel J.M** (1989) Architecture And Knowledge Sources In A Human-Computer Oral Dialogue System. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.). Elsevier Science Publishers B.V. (North Holland), pp 417 - 429.

[CAR93] **Carey M.** (1993) *D.S.M.'s In The Real World: An Investigation Into The Robustness Of Dialogue Structure Model Theory.* Report submitted as part of the Degree of Natural Sciences, University Of Durham, England, April 1993.

[COH91] **Cohen P.R. & Oviatt S.L.** (1991) Discourse Structure And Performance Efficiency In Interactive And Non-Interactive Spoken Modalities. In: *Computer Speech & Language,* Vol. 5, pp 297 - 326. Academic Press Ltd..

[COH79] **Cohen P.R. & Perrault C.R.** (1979) Elements Of A Plan-Based Theory Of Speech Acts. In: *Journal Of Cognitive Science,* Vol. 3, pp 177 - 212.

[COH83] **Cohen, R.** (1983) A Computational Model For The Analysis Of Arguments. Technical Report CSRG - 151, October 1983. Computer Systems Research Group, University Of Toronto.

[COH89] **Cohen, R.** (1989) Studying Arguments To Gain Insight Into Discourse Structure. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 75 - 84. Elsevier Science Publishers B.V. (North Holland).

[CUL77] **Cullingford R.E.** (1977) Script Application: Computer Understanding Of Newspaper Stories. Ph.D. Thesis, Yale University.

[CUL89] **Cullingford R.E. & Turner E.H.** (1989) Using Conversation MOPs In Natural Language Interfaces. In: *Discourse Processes,* Vol. 12, pp 63 - 90.

[DYB93] **Dybkjær H, Bernsen N.O. & Dybkjær L.** (1993) Wizard Of Oz And The Trade-Off Between Naturalness And Recogniser Constraints. In: *Proceedings Of The 3rd European Conference on Speech Communication And Technology (EUROSPEECH),* Vol. 2, pp 947 - 950, Berlin, Sept. 1993

[EAS81] **Eastman C.M. & McLean D.S.** (1981) On The Need For Parsing Ill-Formed Input. In: *American Journal Of Computational Linguistics,* Vol. 7, No. 4, October - December.

[EDW89] **Edwards J.L. & Mason J.A.** (1989) The Structure Of Intelligence In Dialogue. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 85 - 105. Elsevier Science Publishers B.V. (North Holland).

[FAW90] **Fawcett R.P & Tucker G.H.** (1990) Demonstration Of GENESYS: A Very Large, Semantically Based Systemic Functional Grammar. In: *Proceedings Of 13th International Conference On Computational Linguistics,* Helsinki, Vol. 1, pp 47 - 49.

[GAL93] **Galliers J.R. & Sparck Jones K.** (1993) *Evaluating Natural Language Processing Systems.* Technical Report Number 291, University of Cambridge Computer Laboratory, New Museums Site, Pembroke Street, Cambridge, England.

[GAR92] **R. Garigliano** (1992) *The Durham Natural Language System.* Technical Report, Computer Science Division, School Of Engineering And Computer Science, University of Durham, England UK.

[GAR93] **R. Garigliano, Morgan R.G. & Smith M.H.** (1993) The LOLITA System As A Contents Scanning Tool. In: *Proceedings Of The 13th International Conference On Artificial Intelligence, Expert Systems And Natural Language Processing,* Avignon, May 1993.

[GARN92] **Garnham A. & Oakhill J.** (1992) Discourse Processing & Text Representation From A "Mental Models" Perspective. In: *Journal Of Language & Cognitive Processes,* Vol. 7, parts 3 & 4, pp 193 - 204.

[GOI89] **Goillau P.J.** (1989) Pattern Processing & Machine Intelligence Techniques For Representing Dialogues. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 183 - 188. Elsevier Science Publishers B.V. (North Holland).

[GRE92] **Green N. & Carberry S.** (1992) Conversational Implicatures In Indirect Replies. In: *Proceedings Of The 30th Annual Meeting Of The Association For Computational Linguistics,* pp 193 - 200.

[GRI57] **Grice H.P.** (1957) Meaning. In: *Philosophical Review,* Vol. 56, pp 377 - 388.

[GRI69] **Grice H.P.** (1969) Utterer's Meaning And Intentions. In: *Philosophical Review*, Vol. 68, pp147 - 177.

[GRI75] **Grice H.P.** (1975) Logic and Conversation. In: *Syntax And Semantics,* P. Cole & J.L. Morgan (Eds.). Vol 3, pp 41 - 58. Academic Press (New York).

[GRO77] **Grosz B.J.** (1977) *The Representation And Use Of Focus In Dialogue Understanding*. Technical note 151, SRI Project 5844. July 1977.

[GRO86] **Grosz B.J. & Sidner C.L.** (1986) Attention, Intentions And the Structure Of Discourse. In: *Computational Linguistics,* Vol. 12, No. 3, pp 175 - 204, July - Sept..

[GRO89] **Grosz B.J. & Sidner C.L.** (1989) Plans For Discourse. In: *Intentions In Communication,* Morgan J. et al (Eds.), pp 417 - 444. MIT Press.

[HOB80] **Hobbs J.R. & Evans D.A.** (1980) Conversation As Planned Behaviour. In: *Cognitive Science,* Vol. 4, pp 349 - 377.

[HOV88] Hovy E.H. (1988) *Generating Natural Language Under Pragmatic Constraints.* Lawrence Erlbaum Associates, New Jersey.

[KAM81] Kamp H. (1981) A Theory Of Truth And semantic Representation. In: Formal Methods In The Study of Language, Part I. J. Groenendijik, T. Janssen, & M. Stokoff (Eds.), pp 277 - 322. Amsterdam, The Netherlands: MC TRACT 135.

[KIN88] Kintsch W. (1988) The Role Of Knowledge In Discourse Comprehension: A Construction-Integration Model. In: *Psychological Review,* Vol. 95, No. 2, pp 163 - 182.

[KIN85] Kintsch W. & Mross E.F. (1985) Context Effects In Word Identification. In: *Journal Of Memory And Language,* Vol. 24, pp 336 - 349.

[LAM91] Lambert L. & Carberry S. (1991) A Tri-Partite Plan-Based Model Of Dialogue. In: *Proceedings Of The 29th Annual Meeting Of the Association For Computational Linguistics,* pp 47 - 54.

[LAM92a] Lambert L. & Carberry S. (1992) Modelling Negotiation Sub-Dialogues. In: *Proceedings Of The 30th Annual Meeting Of The Association For Computational Linguistics,* pp 193 - 200.

[LAM92b] Lambert L. & Carberry S. (1992) Using Linguistic, World & Contextual Knowledge In A Plan Recognition Model Of Dialogue. In: *Proceedings Of The 15th International Conference On Computational Linguistics,* Nantes, Vol. 1, pp 310 - 316.

[LAM93] Lambert L. (1993) *Recognising Complex Discourse Acts: A tripartite Plan-Based Model of Dialogue.* Technical Report Number 93-19, Department of Computer & Information Sciences, University Of Delaware, Newark, U.S.A.

[LUF90] **Luff P., Gilbert N. & Frohlich D. (Eds.)** (1990) *Computers In Conversation.* Academic Press Ltd., London. ISBN: 0-12-459560-X.

[MCC89] **McCann C.A., Edmondson W. & Moore R.K.** (1989) Practical Issues In Dialogue Design. In: *The structure Of Multimodal Dialogue,* M. M. Taylor et al (Eds.), pp 467 - 481. Elsevier Science Publishers B.V. (North Holland).

[MCK92] **McKevitt P., Partridge D. & Wilks Y.** (1992) Approaches To Natural Language Discourse Processing. In: *Artificial Intelligence Review,* Vol. 6, pp 333 - 364.

[MCK92a] **McKevitt P., Partridge D. & Wilks Y.** (1992) *Experimenting With Intention In Natural Language Dialogue.* Technical Report, Department of Computer Science, University Of Exeter, England.

[MCK92b] **McKevitt P., Partridge D. & Wilks Y.** (1992) Analysisng Coherence Of Intention In Natural Language Dialogue. Technical Report, Department of Computer Science, University of Exeter, England.

[MOR89] **Morel M.A.** (1989) Computer-Human Communication. In: *The structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 324 - 330. Elsevier Science Publishers B.V. (North Holland).

[NEB91] **Nebel B.** (1991) Knowledge Representation For Natural Language Processing. Presentation Notes From Language, Logic & Information Summer School, Saarbrucken, Germany, 1991. DFKI, Saarbrucken.

[NEE89] **Neel F., Waterworth J.A. & Howes J.R.** (1989) Dialogue With A Restricted Partner. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 456 - 466. Elsevier Science Publishers B.V. (North Holland).

[NET94] **Nettleton D.J. & Garigliano R.** (1994) *Dialogue Optimisation In The LOLITA Natural Language Processor using Evolutionary Algorithms.* Technical Report, Laboratory For Natural Language Engineering, Department of Computer Science, University Of Durham, England.

[NOR86] **Norris D.** (1986) Word Recognition: Context Effects Without Priming. In: *Journal Of Cognition,* Vol. 22, pp 93 - 136.

[OST89] **Ostler N.D.M.** (1989) LOQUI: How Flexible Can A Formal Prototype Be? In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 407 - 416. Elsevier Science Publishers B.V. (North Holland).

[PER89] **Perrault C.R.** (1989) Speech Acts In Multimodal Dialogues. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 33 - 46. Elsevier Science Publishers B.V. (North Holland).

[PIT92] **Pitt J., Van Den Bergh K. & Cunningham J.** (1992) *An Experiment In Discourse Understanding*. Research Report DOC 92/1, Department of Computing, Imperial College Of Science, Technology & Medicine.

[POL86] **Polanyi L.** (1986) *The Linguistic Discourse Model: Towards a Formal Theory Of Discourse Structure*. Technical Report 6409, Bolt Beranek & Newman Laboratories Inc., Cambridge Massachusetts.

[POL86a] **Pollack M.E.** (1986) A Model Of Plan Inference That Distinguishes Between The Beliefs Of Actors And Observers. In: *Proceedings Of The 24th Annual Meeting Of The Association For Computational Linguistics,* pp 207 - 214, New York.

[POL88] **Polanyi L.** (1988) A Formal Model of The Structure Of Discourse. Journal Of Pragmatics, Vol. 2, part 3, pp 17 - 37.

[POL92] **Pollack M.E.** (1992) The Use Of Plans. In: *Artificial Intelligence,* Vol. 57, pp 43 - 68. Elsevier Science Publishers.

[RAM91] **Ramshaw L.A.** (1991) A Three Level Model For Plan Exploration. In: *Proceedings Of The 24th Annual Meeting Of The Association For Computational Linguistics,* pp 36 - 46, Berkeley, C.A..

[REI78] **Reichman R.** (1978) Conversational Coherency. In: *Cognitive Science*, Vol. 2, pp 283 - 327.

[REI84] **Reichman R.** (1984) Extended Person-Machine Interface. In: *Artificial Intelligence*, Vol. 22, pp 157 - 218.

[REI85] **Reichman R.** (1985) *Getting Computers To Talk Like You And Me: Discourse Context, Focus & Semantics (An ATN Model).* The MIT Press, Cambridge, Massachusetts. ISBN: 0-262-18118-5.

[REI89a] **Reichman R.** (1989) Conversation Versus Intention. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 121 - 134. Elsevier Science Publishers B.V. (North Holland).

[REI89b] **Reichman R.** (1989) Integrated Interfaces Based On A Theory Of Context And Goal Tracking. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 209 - 227. Elsevier Science Publishers B.V. (North Holland).

[SCH72] **Schank R.C.** (1972) Conceptual Dependency: A Theory Of Natural Language Understanding. In: *Cognitive Psychology*, Vol. 3 Part 4, pp 552 - 631.

[SCH77] **Schank R. C. & Abelson R. P.** (1977) *Scripts, Plans, Goals & Understanding: An Enquiry Into Human Knowledge Structures.* Lawrence Erlbaum Associates, New Jersey.

[SCH84] **Scha R. & Polanyi L.** (1984) A Syntactic Approach To Discourse Structure. In: *Proceedings Of The 10th International Conference On Computational Linguistics, and 22nd Annual Meeting Of The Association Of Computational Linguistics*, pp 413 - 419, Stanford University, California.

[SCH88] **Schegloff E.A.** (1988) Discourse As An Interactional Achievement II: An Exercise In Conversation Analysis. In: *Linguistics In Context,* Tannen D. (Ed.), pp 135 - 158. Norwood, New Jersey.

[SCH88a] **Scha R. & Polanyi L.** (1988) An Augmented Context Free Grammar For discourse. In: *Proceedings Of The 12th International Conference On Computational Linguistics,* pp 573 - 577, July 1988.

[SCH90] **Schach S. (1990)** Software Production And Its Difficulties. In: *Software Engineering,* pp 20 - 40, Ricahrd D. Irwin Inc., and Aksen Associates Inc.

[SCH87] **Schia et al** (1987) Discourse Understanding. In: *Encyclopaedia Of Artificial Intelligence,* Shapiro S.C. (Ed.), pp 233 - 245. Wiley, New York.

[SCH78] **Schmidt C.F. et al** (1978) The Plan Recognition Problem: An Intersection Of Psychology & Artificial Intelligence. In: *Artificial Intelligence,* Vol. 11, pp 45 - 83, North Holland Publishing Company.

[SCH91] **Scholtens A.** (1991) Planning In Ordinary Conversation. In: *Journal Of Pragmatics,* Vol. 16, pp 31 - 58, North Holland Publishing Company.

[SEA69] **Searle J.R. (1969)** *Speech Acts: An Essay In The philosophy Of Language.* Cambridge University Press (London).

[SHA89] **Shadbolt N.R.** (1989) Planning And Discourse. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 107 - 113. Elsevier Science Publishers B.V. (North Holland).

[SID85] **Sidner C. L.** (1985) Plan Parsing For Intended Response Recognition In Discourse. In: *Computational Intelligence*, Vol. 1, pp 1- 10.

[SID79] **Sidner C. L.** (1979) *Towards A Computational Theory Of Definite Anaphora Comprehension In English Discourse.* Ph.D. Thesis, MIT AI Laboratory, Cambridge, Massachussetts.

[SIR89] **Siroux J.** (1989) Pragmatics In A Realisation Of A Dialogue Module. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 347 - 360. Elsevier Science Publishers B.V. (North Holland).

[SME92] **Smeaton A.F.** (1992) Progress In The Application Of Natural Language Processing To Information Retrieval Tasks. In: *The Computer Journal,* Vol. 35, No. 3, pp 268 - 278.

[SMI94] **Smith M.H., Garigliano R. & Morgan R.G.** (1994) Generation In The LOLITA System: An Engineering Approach. Technical Report, Laboratory For Natural Language Engineering, Department Of Computer Science, University of Durham, England, January 1994.

[SMI94a] **Smith M.H.** (1994) *Natural Language Generation In the LOLITA Project: An Engineering Approach.* Ph.D. Thesis In Progress. To be Submitted October 1994.

[SOM92a] **Sommerville I.** (1992) *Software Validation,* Fourth Edition, Addison-Wesley Publishing Company.

[SOM92b] **Sommerville I.** (1992) *Defect Testing,* Fourth Edition, Addison-Wesley Publishing Company.

[STE92] **Stede M.** (1992) The Search For Robustness In Natural Language Understanding. In: *Artificial Intelligence Review,* Vol. 6, pp 383 - 414. Kluwer Academic Publishers.

[WAL90] **Walker M.A. & Whittaker S.** (1990) Mixed Initiative In Dialogue: An Investigation Into Discourse Segmentation. In: *Association For Computational Linguistics,* pp 70 - 78.

[WAN92] Wang Y. & Garigliano R. (1992) An Intelligent Tutoring System For Handling Errors Caused By Transfer. In: *Lecture Notes In Computer science 608: Proceedings Of the 2nd International Conference On Intelligent Tutoring Systems, pp 395 - 404, Springer-Verlag, Montreal, Canada, 1992.*

[WAN94] Wang Y. (1994) *An Intelligent Computer-Based Tutoring Approach For The Management Of Negative Transfer.* Submitted as Ph.D. thesis, Laboratory For Natural Language Engineering, Department of Computer Science, University of Durham, UK, June 1994.

[WAT89] Waterworth J.A. (1989) Interactive Strategies For Conversational Computer Systems. In: *The Structure Of Multimodal Dialogue*, M. M. Taylor et al (Eds.), pp 331 - 340. Elsevier Science Publishers B.V. (North Holland).

[WIL73] Wilks Y. (1973) An Artificial Intelligence Approach To Machine Translation. In: *Computer Models of Thought And Language,* R. Schank & K. Kolby (Eds.). Wh Freeman & Co., San Francisco.

[WIL75a] Wilks Y. (1975a) Preference Semantics. In: *Formal Semantics Of Natural Language,* K. Edward (Ed.). Cambraidge University Press, Cambraidge, U.K.

[WIL75b] Wilks Y. (1975b) An Intelligent Analyser And Understander Of English. In: *Readings In Natural Language Processing,* B. Grosz, K. Sparck Jones & B. Webber (Eds.), pp 193 - 204. Morgan Kaufmann, California.

[WIL75c] Wilks Y. (1975c) A Preferential, Pattern-seeking Semantics For Natural Language Inference. In: *Artificial Intelligence*, Vol. 6, pp 53 - 74.

[YAM93] **Yamashita Y & Mizoguichi R.** (1993) Next Utterance Prediction Based On two Kinds Of Dialogue Models. In: *Proceedings Of The 3rd European Conference on Speech Communication And Technology (EUROSPEECH),* Vol. 2, pp 1161 - 1164, Berlin, Sept. 1993

[YOU93] **Young R.S.** (1993) Dialog Structure And Plan REcognition In Spontaneous Spoken Dialog. In: *Proceedings Of The 3rd European Conference on Speech Communication And Technology (EUROSPEECH),* Vol. 2, pp 1169 - 1172, Berlin, Sept. 1993

Appendix I: Chat DSM

This appendix contains the DSM for the informal chat dialogue situation. The DSM applies to all the particpants involved in the dialogue, as the direction is irrelevant. The following DEs are present:

Motivational Elements

DE	Setting	Comments
Goal:	Process	This is classed as a process, as the goals of the chat (talk to other participants) are achieved as the dialogue progresses rather than a task completed at the end. Goal: Be Informed (talk to other participant and find out more about them, pass time of day etc.)
Persuasive:	LOW	Participant is not necessarily trying to persuade other person than something is true.
Information Seeking:	High	Participant may be involved in giving out information about his/herself, and also finding out information about the other person. Alternatively, they may just discuss topics in general.
Emotional Exchange:	Medium	While there is no specific goal to excahnge emotion, emotional exchange is not forbidden.

External Elements:

DE	Setting	Comments
Time Limit:	None	There is no specific time limit set for a chat.
Number:	2	The default numbe rof particpants is set at 2.

Motivational Elements:

DE	Setting	Comments
Distribution Of Time:	Equal	The particpants should all speak for an equal amount of time throughout the dialogue.
Rhythm:	Medium	Sentence length will not be excessively long or short.

Colour:	Medium	Language does not need to be excessively "flowery"
Register:	Informal	The dialogue situation is an informal one, and this is reflected in the language used.
Temporal Progression:	See right	Stages can include: greetings, introductions, Main Body, Closing phrases, Farewells.
Fixed Topic:	Unfixed	In a chat, the topic is not fixed to any particular subject.

Appendix III: Constraints

The following are constraints used within the theory. They are grouped according to the DE to which they are attached.

Dominance

1. Initiate the dialogue, (if other person has not already done so)
2. Terminate the dialogue
3. Interrupt a sub dialogue
4. Choose the topic
5. Change the topic
6. Initiate a sub-dialogue
7. Ask a question
8. Refuse to answer a question

Time Limit

1. Determine whether there is a time limit
2. Determine absolute time limit, (i.e. precise time dialogue must end at)
3. Determine duration of dialogue, (i.e. how long the dialogue must last)
4. If time limit is reached, terminate dialogue
5. If no time limit is found keep going until:
   a) something external happens to stop you.
   b) fundamental goal fails, (may apologise or not)

Temporal Progression

1. Greetings
2. Initiate introduction
3. Make small-talk
4. Move to main body of dialogue
5. Move to summary of main body
6. Terminate dialogue with closing phrase, (good-bye, thank you etc.)
7. Scale dialogue according to time limit
8. Exit the dialogue, (other person has terminated the dialogue but you still need to say goodbye etc.)

Persuasive: (from direction of persuader)

1. Use only information which is beneficial to goal
2. Use *Emotional Exchange* constraints
3. Determine whether the other(s) believe what is being said

(from direction of person being persuaded)

4. Determine whether other(s) is/are telling the truth
   a) Believe what is being said
   b) Reject what is being said

Information Seeking: (from direction of seeker)

1. Determine what information is needed
2. Formulate questions about the topic
3. Ask unambiguous, clear questions
4. Continue to ask questions until desired information is attained

5. Clarify ambiguous questions
6. Clarify questions which are out of context
7. Update database/memory when information is received
8. Avoid taboo topics
9. Explain motivation for seeking information

(from direction of information supplier)

10. Seek clarification of ambiguous questions
11. Seek explanation for questions which are not understood
12. Decide what information is relevant/irrelevant
13. Answer questions using knowledge from memory
14. Check information supplied has been understood
15. Ask about motivation for seeking information
16. Refuse to answer question

Emotional Exchange: (from direction of person wishing to exchange emotion)

1. Determine the emotion to be exchanged, (sadness, happiness, humor etc.)
2. Select appropriate register
3. Determine whether exchange was successful

(from direction of person receiving exchange)

4. Recognise attempt to exchange emotion
5. Recognise which emotion, (happiness, sadness, humour etc.)
6. Respond appropriately , (laugh, cry, etc.)
7. Reject attempt to exchange emotion
8. Respond with different emotion other than what was attempted
9. Select register appropriate to emotion being expressed

Distribution of Time

1. Talk for majority of time

   a) Select rhythm: long, flowing

2. Talk for least amount of time

   a) Select rhythm: short, economical

3. Time distributed evenly

   a) Select rhythm: average, medium length

Register

1. Activate appropriate register according to pre-defined slot, (formal, informal etc.)
2. Formal: use of slang words/phrase forbidden, word order/grammar strictly correct
3. Informal: may use slang, word order/grammar less strict
4. Taboo: never use words/phrases from this section
5. Use jargon/technical terms
6. Break register to produce emotional exchange, (shock, anger, humor etc.)

Rhythm

1. Use short economical rhythm

   a) no detail, no spurious words, minimum information only

2 Use medium, average rhythm

   a) Add few details, but no spurious words

3 Use long, flowing rhythm

   a) Add detail, spurious words, long sentence construction

4 Check time limit:

a) time limit approaching, use short rhythm

b) time limit not close, use medium rhythm

c) time limit long way off, use long rhythm

d) no time limit —tailor rhythm according to distribution of time

5. Check register:

   a) register informal, rhythm may vary

   b) register formal, rhythm must remain regular and controlled

Number, (if number of participants is known and fixed)

1. Address appropriate participant at given point in dialogue
2. Introduce a participant to the other(s)

Number, (if number of participants is unknown or unfixed)

1. Evaluate number of participants:

   a) Set counter to 2 at start of dialogue

   b) increment counter by 1 each time new participant is encountered

2. Address appropriate participant at given point in dialogue
3. Introduce a participant to the other(s)

Goal, (process)

1. Check what must be achieved by which point in dialogue
2. Check this has been achieved at each dialogue cycle (i.e. interaction between participants)

Goal (task)

1. Check each sub-goal is achieved
    a) Check with other(s) that everything is understood
    b) Inform participant(s) that sub-goal has been achieved
2. Move to next sub-goal
3. Devise list of criteria for success
4. Check ultimate goal has been achieved.
    a) if goal achieved, refer to *Dominance* constraints for next action
    b) if goal not yet achieved, repeat steps 2 to 4

Fixed Topic

1. Never discuss taboo topics
2. If topic is fixed:
    a) do not change topic
3. If topic is not fixed:
    a) Check with dominance to see if you may change topic
    b) Change topic
    c) Wait for other(s) to change the topic

Appendix III: Example Algorithms For The Construction Of Plan Boxes

The following algorithms are included as example plan boxes:

(a) Clash Of Beliefs Plan Box

To be used by the system when opinions are expressed by the other participant:

1. Check LOLITA's beliefs and compare them with the beliefs of the other participant(s)

2. If they are identical, produces following effects:
   a) Register → more informal
   b) Rhythm → more relaxed, longer, free flowing
   c) Emotional Exchange → Empathy and intimacy emotions increase

3. If identical, say so and continue talking on this subject until time limit or external factor(s) dictate dialogue must end, then move to point 2 of "Closing Phrase" plan box

4. If beliefs are similar, (but not identical), use moderated version of above. Effects produced are the same but weaker. Less concerned with sticking to same subject. Emotions generated less +ve.

5. If beliefs are similar, (only clash in minor ways), and politeness is a goal, overlook minor differences in belief, for sake of overall agreement.

6. If beliefs clash in major way, and goal is to produce +ve emotions, politely state own beliefs and try to move away from topic.

7. If major clash has occurred and other participant will not allow topic to be dropped, move to point 4 of "Closing Phrase" plan box. Can be less polite in closing dialogue.

(b) Information Seeking Plan Box

To be used when the system needs to obtain information:

1. Ask other participant(s) directly by asking a question. Formulate question by taking consideration of register, rhythm, distribution of time and taboo topics.

2. If dominance is equal to or less than that of other participant(s), obtain information by offering some in return.

3. If no collaboration between participants, use cunning to obtain information. Do not disclose motivation for seeking information. Use cunning by either:
   a) pretend that you know information already, (call other person's "bluff"). OR
   b) avoid asking direct questions to obtain information. Talk about topic in more general terms.

4. If dominance is much higher than that of other(s), use threats to obtain information.

5. If dominance is much lower than other(s), beg for information.

Appendix IV: The Emotional Values

Listed below are the emotional values implemented within the dialogue analysis module:

- **Shock Offense:** This is the emotion triggered if some input could be deemed shocking, in the sense of being outrageous *and* offensive. For example, if a participant has beliefs about sexual equality, and another participant makes a sexist joke, the reaction may be one of Shock Offense. Depending upon the level of intensity attached to this emotion (High, Medium, Low) this can range from mild surprise through to extreme shock.
- **Shock Event:** This is the sort of shock which results from an input conveying an unexpected piece of information such as being told that the Queen has abducated. The distinguishing feature between this type of shock and the previous type, is that here no element of offensiveness is present. As before, depending upon the intensity of the emotion, this can range from mild surprise to extreme shock.
- **Anger Offense:** If the input is offensive, then the reaction may be one of anger. The input may be offensive in terms of insulting one's personal beliefs, or insulting the person directly, or by being over-personal and discussing topics which are normally considered taboo. This has been implemented as often resulting from **Shock Offense**, as sometimes shock can turn to anger. The implementation also takes into account the fact that shock may not accompany this emotion, but depending upon the level of "offensiveness" of the input, and the instantiation factors involved such as personality and relationship between participants, this emotion may not be preceded by shock.
- **Anger Rational:** This is triggered if the input is contrary to one's beliefs or interests, but is on an intellectual level as opposed to a personal or emotive

level. E.g. if during a political debate you are told that all Tories should be hung, and you support the Conservative party, this may anger you on an intellectual level, but not on a personal, individual basis.

- **Pity:** Pity is invoked if one is told something sad or negative about someone that you have positive or neutral feelings towards. e.g. being told that your friend is ill, or hearing about the plight of the homeless.
- **Embarrassment:** This is triggered if one is told something that is over personal or which is unsolicited and of an intimate nature. (Personality influences the reaction between embarrassment and anger...if a stranger says (s) is in love with you, this would constitute input of an intimate nature. Depending upon the personality, some people would be embarrassed by this, others may be offended and react angrily).
- **Boredom:** This is triggered if the input is repetitious either in content or form, especially if the topic is not included in one's lists of interests, beliefs or background.
- **Self Pity:** This is triggered if the input is about the recipient and the input carries an emotionally negative charge (but is not included in the negative emotions which would trigger an angry response).
- **Guilt:** This is triggered if the input is an accusation which is true.
- **Self Esteem:** This emotion is invoked if the input is about the recipient and is positive, (e.g. praise or congratulations on having achieved something good).
- **Happy Nostalgia:** This is triggered when the input topic is related to an event or action in the past, which carries a positively emotional charge (e.g. remembering a much loved deceased relative).
- **Affection Romantic:** This is triggered when you want to say something to a romantic/sexual partner. e.g. if the other person is your spouse and they have just said "I love you", the response may be "I love you too".
- **Affection Family:** As above, but is triggered when the other participant is a family member, not a sexual or romantic partner.

- **Affection Platonic:** As the previous two, but occurs when the relationship between participants is positive but platonic.
- **Serenity:** If the input content is a positive statement, then the general emotion evoked is serenity.
- **Excitement:** This is the emotional reaction if the input content is good news (i.e. is new information regarding an emotionally positive subject (i.e. someone/thing that you like). The event and the object of that event has to carry positive emotional charge for the information to be classed as good news.

Appendix V: Test Results Of Using DSMs In The Real World

The Company Representatives Direction of the final fixed form of the D.S.M.

DE — **Constraint Application**
External Elements

Number
1 *Use default value* (2).
Increase if introduction deems necessary.
Increase if own party is more than one.

Time Limit
1 *Not usually important.*
2 *Evaluate time limit.*
a) Use time until next appointment as default.
b) increase if other participants states their value is less.
3 *Usually < 2Hrs.*

Temporal Progression
1 *Greetings.*
2 *Initiate introduction* (if nobody else has.)
3 *Verify proposed purpose of meeting.*
a) gives value to main matter.
b) ascertain purpose if meeting was called by other participant.
4 *Move to main discussion loop.*
a) *Discuss main matter.* (proj. need to prompt).
b) *Raise other matters.* Propose move to sub-dialogue loop.
c) *Sub dialogue loop*
Discuss other matter.
Raise other matters. Propose move to sub-dialogue loop.
Listen to advice on other matter.
End Sub-dialogue loop.
d) *If consultant proposes discussion on other matters enter sub-dialogue loop.*
End Loop.
5 *Answer any questions on factors necessary for advice/proposal.*
6 *Discuss advice on main matter.*
7 *Discuss advice on any other matters remaining which the consultant raises.*
8 *Clarify arrangements.*
9 *Arrange further meetings.* (If matters left to be advised on.).
10 *Discuss consultants summary.*
11 *Terminate dialogue with closing phrase.*
a) when consultant indicates so.
b) if time limit dictates strongly so.
c) focus on matters higher up own hierarchy when short of time.

Motivational Elements

Goal(processes)

1 *Check which matters still need advice on.*
Hierarchy produced from own interests.
2 *Impress consultant with regard to state of entity.*
3 *Impress consultant with regard to state of self.*
4 *Ensure that ultimate goal has been achieved.*
Ultimate goal = receive advice on main matter.
5 *Overall goal is the improvement of the state of the entity.*

Information Seeking

1 *Advice is desired on all matters.*
2 *Formulate questions.*
a)Should be on main matter or
b) on other matter if sub-dialogue has been initiated.
3 *Information should be in a form which can be acted on or specifies no action.*
4 *If don't understand advice or see how it relates then question further.*
5 *Ask clear unambiguous questions.*
6 *Clarify ambiguous questions.*
7 *Treat vague statement on own entity as a question.*
8 *Explain motivation for seeking information.*
If other person seems puzzled.
9 *Seek clarification of ambiguous questions.*
10 *Seek explanation for answers not understood.*
Terms not understood ask about specific words.
Sentence not understood ask more general clarification question.
11 *Continue to ask questions until desired information is received.*
If on matter high up in own hierarchy.(proj encouraged/ bank. allowed/ con. discouraged).
12 *Update database with information received. No filter.*
13 *Information seeking is high.*

Persuasive

1 *Use information which is beneficial to goal.*
2 *Use emotional exchange constraints.*
Toned down.
3 *Answer questions with information related to entity and matters being discussed.*
4 *Point out problems with current state of affairs that consultant might be able to advise on.*
5 *Present facts relating to entity in favourable light.*
6 *Present facts relating to self in favourable light.*
7 *If not sure of facts. State this.*
8 *Store unanswered questions.*
Can return to these.
9 *High on information receiving.*
10 *Point out problems with applying advise.*
11 *Propose actions based on advise.*
12 *Return to answering previous question if not asked another one.*
13 *Persuasion varies in strength.*

Change of Direction

1 *Determine if advise is valid and whether it is good for company or not.*
1a *Advise is applicable.*
Show agreement and gratitude if includes offer of help.
1b *Reject what is being said.*
Express doubt and suggest waiting for other indicators.

Emotional Exchange

1 *Default setting = no emotion.*
2 *Possible desperation, excitement, confident, disapproval.*
3 *Select register default = formal.*
  a) Optional humour, friendly , if know consultant well.
4 *Determine if emotional exchange was successful.*
  Supervisor should be impressed with prospects of entity.

Change of direction.

1 *If consultant is giving inappropriate advice or not supplying expected help then emotion is disapproval and register is formal.*

Dominance

1 *Initiate dialogue* (If other participant has not.)
2 *Can suggest sub-dialogue.*
3 *Can suggest topic.*
4 *Can interrupt speaker.*
  If have already understood what speaker was trying to say.
  If can improve on own answer to previous question.
5 *Ask a question.*
  Don't answer a question with a question.
6 *Dominance is low to medium.*

Topic

1 *Do not discuss taboos.*
  Minor taboos allowed but don't need to give specific answers.
2 *Do not discuss company/ department taboos.*
3 *Topic is restrained to consultants 'subject'*
4 *Topic should relate to entity represented by client.*
5 *Topic can be changed by consultant.*

Distribution of Time

1 *Time is distributed evenly.*

Register

1 *Formal .*
2 *Technical terms.*
  Use only if from consultants 'subject'
3 *Try not to use slang.*

Rhythm

1 *Medium.*
2 *Varies with information to be conveyed.*

Appendix VI: The "Characteristics" Of The System

The system was developed with the following "characteristics", the table describing in general terms the constraints attached to each intensity:

Confidence	High	Participant is self-assured enough to offer opinions and information even when not explicitly asked for.
	Med	Information is only given if asked for, or when erroneous information is received from the other participant.
	Low	Information is only given when directly asked for.
Education	High	A wide range of interests and knowledge is available to them, and so they will be able to talk about a wider range of topics. They sentence structure may be more complex, and register may be more varied.
	Med.	Knowledge expressed by the system will be general and only expert on topics relating to personal background, register and sentence structure non complex.
	Low	The sentence structure used will be simple, the range of topics and registers used will be limited. The grammatical constructions used may sometimes be incorrect. This is often the case with small children, for example.
Self Respect	High	Does not like to be put down, insulted or manipulated, and will react strongly against such attempts.
	Med.	Does not like to be insulted, but is more tolerant.
	Low	Does not react badly against criticism, and will allow the other person(s) to manipulate the dialogue.

Humour		(Note that the system cannot handle humour in its present state as this is such a complex and difficult aspect of natural language to model, and is outside the scope of this project).
Impulsiveness	High	Likely to react quickly and dramatically to situations, especially negative ones.
	Med.	Participant will protest against insults, or against utterances which conflict with system's beliefs or interests.
	Low	Very slow to react against negative situations, and is much more tolerant. Will only react if severe/extreme situation is encountered.
Curiosity	High	Keen to acquire new knowledge. Asks questions about topics/utterances that are not understood or common ground.
	Med.	Shows polite interest in topics which are not common ground, but will become bored after some time.
	Low	Prefers to discuss topics which are of interest or common ground. Becomes bored very quickly with topics which are not understood.
Popularity	High	Participant wants to be liked. Will avoid upsetting other people by insulting them or saying anything negative about them. Says nice things about other people in an attempt to be liked in return.
	Med.	Will say nice things about other people, but if input is negative and about participant, then reciprocate with negative statement about the other person.
	Low	Does not care whether other people like them or not. Will act more according to mood. If in bad mood, then will output negative emotions.

Cheerfulness	High	Likes to be happy i.e. prefers positive emotions to negative emotions. Will always try to get back to good mood, even if negative emotions are received.
	Med.	Prefers positive emotions to negative, but easier to provoke bad mood and more likely to stay in bad mood longer than at High level.
	Low	Prefers negative emotions to positive, therefore always in bad mood.

Appendix VII: The Source Model

As well as the characteristics presented in Appendix V, the "personality" of the system is also influenced by the source control module. The module built within the implementation has been constructued using the following categories:

1. **Name**. The name of the source to whom the model relates.

2. **Ability**. concerns the source's level of expertise concerning certain topics. The sub-categories are:

   i) **expertise**: comprises of a topic and intensity (None, Low, Medium, High, Total). The higher the intensity, the higher the source's expertise about the specified topic. There can be a list of more than one topics and their corresponding intensities.

   ii) **reasoning**: concerns the source's ability to reason and make logical, intelligent inferences/decisions. Measured in terms of intensity.

   iii) **experience**: concerns the how much experience the source has of the specified topic. This is also measured in terms of intensity.

3. **Relationship**: This concerns the way in which the source relates to the other participant. It comprises of the following components:

   i) **names**: The name of the source and the other participant are given as name1 and name2.

   ii) **helpfulness**: concerns how helpful and cooperative the source believes the other participant to be. This is measured in terms of intensity.

iii) **trustworthiness**: concerns how trustworthy and reliable the source believes information received from the other participant to be. This is measured in intensity.

iv) **common ground**: This relates to the topic(s) which are of common interest to the source and the other participant. The topic(s) are listed along with a corresponding intensity, which indicates how "common" the ground is e.g. whether they only share a loose interest in sport, or whether they both have a keen interest in sport and share similar views about the topic.

v) **emotion mood**: The emotion mood in *Relationship* concerns the way one participant feels *about the other*. (Note that the actual emotional mood of the participant is another matter, and is dealt with by the emotional planboxes).

4. **Interest**. This category relates to the interests of the source. A list of interests is given (each relating to a node within the semantic network), and attached to each topic is an intensity which indicates the level of interest the source has in that topic. The higher the intensity, the greater the interest.

5. **Beliefs**. This category concerns the beliefs and opinions of the source. It is constructed in a similar way to the *Interests*, but each topic is a node within the semantic network relating to what the source believes rather than likes. Again, strength of belief is measured in terms of intensity.

Appendix VIII: Dial_cycle Abstype

The following is the abstract type for the dialogue cycle. During the dialogue cycle, a new dialogue state is created. The sequence of events discussed in Chapter Five section 5.3.1 is followed.

****

MAIN DIALOGUE FUNCTION

```
> dial_cycle :: Dial_state -> ([Char], Dial_state)
> dial_cycle ds = ctx_meaning [mk_new_ds ds]


> mk_new_ds :: Dial_state -> Dial_state

> mk_new_ds
>  = normalise_d_tactic . final_rank_trace .
>     realise_planboxes    . partial_tactics_trace .
>     combine_compatible_tactics . sort_actions_by_rank .
>     realise_constraints    . activate_previous_plan .
>     avoid_repetitions      . choose_planboxes  . emotions_trace .
>     partial_goals_trace    . set_local_goals . balance_constraints  .
>     apply_individual_constraints  . apply_dsm_constraints .
>     compare_temp_part   . intellectual_reaction  .
>     emotion_reaction       . change_partial_tactics []
```

****

SUBSIDIARY FUNCTIONS TO MK_NEW_DS

```
> sort_actions_by_rank :: Dial_state -> Dial_state

> sort_actions_by_rank ds
>  = change_partial_tactics (gsort ranksort (get_partial_tactics ds)) ds
>     where
>      ranksort a b = get_Rank a <= get_Rank b
```

sort_actions_by_rank sorts the possible actions by their rank.

****

```
> choose_planboxes
>  = action_default_tacticPB . action_cause_emotionPB .
>     action_express_emotionPB .
>     action_top_BeInformedPB . action_top_persuadePB .
>     action_top_informPB


> realise_planboxes
>  = cause_emotionPB . express_emotionPB .
>     top_BeInformedPB . top_persuadePB .
>     top_informPB . default_tacticPB
```

****

```
> intellectual_reaction :: Dial_state -> Dial_state

> intellectual_reaction
>  = acceptable_in_dialogue . acceptable_in_dialogue_locally .
>     acceptable_per_se
```

intellectual_reaction checks if the participant input is acceptable under a rational point of view.

```
> acceptable_per_se :: Dial_state -> Dial_state

> acceptable_per_se ds
>  | inputs == [] || bad_inputs == []
>    = ds
>  | otherwise
>    =       dial_trace1 g "acceptable_per_se"                  $
>            change_globalDS g1                          $
>            new_partial_tactic [BeInformedGoal] [Question] []
>                                     "why_questPB" 1 1  $
>            new_partial_tactic [InformGoal] [] out_cont "" 0 0      $
>            ds
>     where
>      inputs = rich_ev_inputDS ds
>      bad_inputs = filter (is_belief_low g) inputs
>      g = get_globalDS ds
>      (g1,out_cont) = make_do_not_believe g (hde53 3 bad_inputs)
```

acceptable_per_se checks if the participant input is acceptable independently of the dialogue, ie if it does not clash with existing knowledge (type or factual should be separated at least as different ranks, at present it deals only with type).

```
> acceptable_in_dialogue_locally :: Dial_state -> Dial_state

> acceptable_in_dialogue_locally
>   = check_relevance . check_answer_questions
```

if there is no switch of topic, then nothing happens. If there is one, then it depends if it has been signalled (eg 'changing subject', 'this reminds me' etc.)

Appendix IX: Cause_Emotions Abstype

The following code represents the plan boxes for causing an emotional reaction in another participant. The top level plan box is given, followed by two example plan boxes which deal with the causing of AffectionPlatonic and AngerOffense

```
*****************************************************************
******************** CAUSE_EMOTION planbox ******************
*****************************************************************
```

Tells system how to provoke a certain emotional value in the user.

```
> cause_emotionPB :: Dial_state -> Dial_state

> cause_emotionPB
>    = cause_AngerOffense . cause_fear . cause_misery .
>      cause_Guilt . cause_Envy . cause_AffectionRomantic .
>      cause_AffectionPlatonic . cause_AffectionFamily .
>      cause_SelfEsteem . cause_Respect . cause_Excitement .
>      cause_Gratitude . cause_Serenity


> action_cause_emotionPB :: Dial_state -> Dial_state

> action_cause_emotionPB
>    = action_cause_AngerOffense . action_cause_fear .
>      action_cause_misery .
>      action_cause_Guilt . action_cause_Envy .
>      action_cause_AffectionRomantic .
>      action_cause_AffectionPlatonic .
>      action_cause_AffectionFamily .
>      action_cause_SelfEsteem . action_cause_Respect .
>      action_cause_Excitement .
>      action_cause_Gratitude . action_cause_Serenity

********************************************************************
```

```
> cause_AffectionPlatonic :: Dial_state -> Dial_state

> cause_AffectionPlatonic ds
>  | null ranked_action
>     = ds
>  | action_line == 1
>     = planboxres2Dial_state ds (mk_pack1 rank)
>  | action_line == 2
>     = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 3
>     = planboxres2Dial_state ds (mk_pack rank)
>  | otherwise
>     = mismatch_error "cause_AffectionPlatonic"
>   where
>    ranked_action = partial_tactic_filt ds
>                             "cause_AffectionPlatonic"
>    rank = get_Rank (hde51 10 ranked_action)
>    action_line = get_PlanboxAction (hde51 11 ranked_action)
>    (g, cont) = fondnessE fix_lolita [get_part_nameDS ds] ds
>    (g1, cont1) = loveE fix_lolita [get_part_nameDS ds] ds
>    mk_pack n
>     = (([mk_tactic cont PackageT
>          (CauseEmotionGoal AffectionPlatonic)
>                              AllSame], n), g)
>    mk_pack1 n
>     = (([mk_tactic cont1 PackageT
>          (CauseEmotionGoal AffectionPlatonic)
>                              AllSame], n), g1)
```

- to cause AffectionPlatonic: output_content = fondness

```
> action_cause_AffectionPlatonic :: Dial_state -> Dial_state

> action_cause_AffectionPlatonic ds
>  | goal_aff_plat_cond && em_val && popularity_high
>     = new_RAction 1 1
>  | em_val && popularity_high && good_rel
>     = new_RAction 2 1
>  | em_val && good_rel
>     = new_RAction 3 1
>  | otherwise
>     = ds
>   where
>    new_RAction line rank
>     = new_partial_tactic [CauseEmotionGoal AffectionPlatonic]
>                  [AllSame] [] "cause_AffectionPlatonic" line rank ds
>    emVI = get_lol_em ds
>    emI = get_intensityV AffectionPlatonic emVI
```

```
>     em_val
>       = (AffectionPlatonic `is_in` map get_em_value emVI) &&
>                                                  emI `is_in` [High, Total]
>    popularity_high
>       = find_intC Popularity (get_lol_charVI_DS ds) `is_in` [High, Total]
>    goal_aff_plat_cond
>       = CauseEmotionGoal AffectionPlatonic `is_in` (get_partial_goalsDS ds)
>    good_rel = GoodMood `is_in` get_relValue ds
```

****

```
> cause_AngerOffense :: Dial_state -> Dial_state

> cause_AngerOffense ds
>  | null ranked_action
>       = ds
>  | action_line == 1
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 2
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 3
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 4
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 5
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 6
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 7
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 8
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 9
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 10
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 11
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 12
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 13
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 14
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 15
>       = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 16
>       = planboxres2Dial_state ds (mk_pack rank)
```

```
>  | action_line == 17
>     = planboxres2Dial_state ds (mk_pack rank)
>  | action_line == 18
>     = planboxres2Dial_state ds (mk_pack rank)
>  | otherwise
>     = mismatch_error "cause_AngerOffense"
>   where
>    ranked_action = partial_tactic_filt ds "cause_AngerOffense"
>    rank = get_Rank (hde51 24 ranked_action)
>    action_line = get_PlanboxAction (hde51 25 ranked_action)
>    (g, cont) = insultE (get_part_nameDS ds) ds
>    mk_pack n
>     = (([mk_tactic cont PackageT
>                  (CauseEmotionGoal AngerOffense)
>                                    AllSame], n), g)
```

- to cause AngerOffense: output_content = taboo || output_register = taboo || output_topic = taboo

```
> action_cause_AngerOffense :: Dial_state -> Dial_state

> action_cause_AngerOffense ds
>  | goal_AngerOff_cond && em_val1 && bad_mood && bad_relation
>                                    && impulsiveness_high
>      = new_RAction 1 1
>  | goal_AngerOff_cond && em_val1 && bad_relation
>                               && impulsiveness_high
>      = new_RAction 2 2
>  | goal_AngerOff_cond && em_val1 && bad_mood && bad_relation
>                                  && not impulsiveness_high
>      = new_RAction 3 3
>  | goal_AngerOff_cond && em_val1 && not impulsiveness_high
>      = new_RAction 4 4
>  | goal_AngerOff_cond && em_val2 && bad_mood && bad_relation
>                                    && impulsiveness_high
>      = new_RAction 5 5
>  | goal_AngerOff_cond && em_val2 && bad_relation
>                             && impulsiveness_high
>      = new_RAction 6 6
>  | goal_AngerOff_cond && em_val2 &&  impulsiveness_high
>      = new_RAction 7 7
>  | goal_AngerOff_cond && em_val2 && not impulsiveness_high
>      = new_RAction 8 8
>  | goal_AngerOff_cond && em_val3 && bad_mood
>                             && impulsiveness_high
>      = new_RAction 9 9
>  | not goal_AngerOff_cond && em_val1 && impulsiveness_high
```

```
>                               && bad_mood  && bad_relation
>      = new_RAction 10 10
>   | not goal_AngerOff_cond && em_val1 && not impulsiveness_high
>                               && bad_mood && bad_relation
>      = new_RAction 11 11
>   | not goal_AngerOff_cond && em_val1 && bad_mood && bad_relation
>      = new_RAction 12 12
>   | not goal_AngerOff_cond && em_val2 && bad_mood && bad_relation
>                                   && impulsiveness_high
>      = new_RAction 13 13
>   | not goal_AngerOff_cond && em_val2 && bad_relation &&
>                                   not impulsiveness_high
>      = new_RAction 14 14
>   | not goal_AngerOff_cond && em_val2 && bad_relation
>      = new_RAction 15 15
>   | not goal_AngerOff_cond && em_val3 && bad_mood &&
>                              not impulsiveness_high && bad_relation
>      = new_RAction 16 16
>   | not goal_AngerOff_cond && em_val3 && bad_mood
>      = new_RAction 17 17
>   | not goal_AngerOff_cond && em_val3
>      = new_RAction 18 18
>   | otherwise
>      = ds
>    where
>     new_RAction line rank
>       = new_partial_tactic [CauseEmotionGoal AngerOffense]
>               [AllSame] [] "cause_AngerOffense" line rank ds
>     emVI = get_lol_em ds
>     emI = get_intensityV AngerOffense emVI
>     cond_AngerOffense = AngerOffense `is_in` map get_em_value emVI
>     em_val1 = cond_AngerOffense && emI `is_in` [High, Total]
>     em_val2 = cond_AngerOffense && emI == Medium
>     em_val3 = cond_AngerOffense && emI `is_in` [Low, None, NoInf]
>     bad_relation = (BadMood `is_in` get_relValue ds)
>     bad_mood = (BadMood == get_lol_moodDS ds)
>     impulsiveness_high
>       = find_intC Impulsiveness (get_lol_charVI_DS ds)
>                                  `is_in` [High, Total]
>     goal_AngerOff_cond
>       = CauseEmotionGoal AngerOffense `is_in`
>                               (get_partial_goalsDS ds)
```

****

Appendix X: The Ticket Booking DSM

This appendix contains the DSM for the ticket booking dialogue situation. The DSM applies to all the particpant involved in taking the reservation. The following DEs are present:

Motivational Elements

DE	Setting	Comments
Goal:	**Task**	This is classed as a task, as the goal of the dialogue (find and make suitable flight booking) is achieved at the end. goal = 1) provide information on flights, 2) gather information from client required to book flight (destination, date and time of departure, time of arrival), 3) calculate which flights, if any, can meet those criteria 4) discuss booking with client 5) if client accepts flight, book flight. 6) if client rejects flight, seek alternative information, repeating steps 2 through to 6.
Persuasive:	**LOW**	Participant is not necessarily trying to persuade other person than something is true.
Information Seeking:	**High**	Participant is involved in giving out information about the flights, and also finding out information about the other person's needs.
Emotional Exchange:	**None**	Emotional exchange is not expected in this dialogue situation.

External Elements:

DE	Setting	Comments
Time Limit:	None	There is no specific time limit set for this dialogue.
Number:	2	The default numbe rof particpants is set at 2.

Verbal Elements:

DE	Setting	Comments
Distribution Of Time:	Equal	The particpants should all speak for an equal amount of time throughout the dialogue.
Rhythm:	Medium	Sentence length will not be excessively long or short.
Colour:	Low	Language does not need to be excessively "flowery"
Register:	Formal	The dialogue situation is an formal one.
Temporal Progression:	See right	Stages can include: greetings, introductions, Main Body, Closing phrases, Farewells.
Fixed Topic:	Fixed	In this dialogue, the topic is fixed to the particular subject of air flights.