A TWO-LEVEL DIALOGUE REPRESENTATION[1]

Giacomo Ferrari
Department of Linguistics
University of Pisa

Ronan Reilly
Educational Research Centre
St Patrick's College, Dublin 9

ABSTRACT

In this paper a two level dialogue representation system is presented. It is intended to recognize the structure of a large range of dialogues including some nonverbal communicative acts which may be involved in an interaction. It provides a syntactic description of a dialogue which can be expressed in terms of re-writing rules. The semantic level of the proposed representation system is given by the goal and subgoal structure underlying the dialogue syntactic units. Two types of goal are identified; goals which relate to the content of the dialogue, and those which relate to communicating the content.

1. INTRODUCTION

Research on computational modelling of discourse has highlighted some important aspects of human dialogic communication. In some cases (Reichman, 1984), a structural description of linguistic communication has been attempted although not for proper dialogue. What is required is a structural description identifying a corresponding set of communicative acts which can be combined in a fixed pattern, to form higher level communication act categories or dialogue constituents.

The importance of such a structural description, if attained, is that it would make possible an axiomatic theory of dialogue, embedding rhetorical patterns, focusing, and focus shifting.

A possible basis for such a structural description is Burton's (1981) taxonomy of communication acts, with some modifications.

Such a formalization of dialogue, however, is a fully syntactic one which needs to be augmented with some semantics. Our assumption is that dialogue constituents have a semantic interpretation in terms of goals and subgoals.

A theoretical frame for a dialogue classification system based on these assumptions is being developed with the aim of providing a coherent basis for the computational modelling of dialogue.

The final aim of our project is, in fact the design and implementation of a computational dialogic interaction system, with the ability to recover from communication failure.

2. STRUCTURAL DESCRIPTION

The syntactic level of a dialogue description system has to meet three requirements:

(a) It must consist of a model whose generality makes it applicable to both person-person and person-machine dialogues, irrespective of the socio-institutional disposition of any speaker/user.

(b) It must be capable of describing in terms of communication concepts the interactive nature of dialogue exchanges.

(c) It must be capable of "naming" the units which compose the dialogue exchanges, i.e. identifying an utterance as both of a particular type and playing a certain role in a particular exchange.

The dialogue units taxonomy proposed by Burton (1981) seems to satisfy these requirements. According to her analysis system we distinguish four hierarchically related levels: the interaction, which is the largest unit; the transaction; the exchange; the move; and the act. An interaction consists of a number of transactions, a transaction consists of a number of exchanges, and so on. The smallest interactive unit in this system is the move. A move can consist of a number of linguistic and non-linguistic "acts" which are realised either as utterances or physical actions.

2.1 ACTS

An act is a way of classifying utterances which occur in a conversational move. The act is a level above that of the utterance, and a move can consist of a number of acts. For example, the following is an opening move by a teacher in a classroom dialogue consisting of two acts (in brackets):

T: those letters have a special name (Starter)
   do you know what it is (Elicitation)

Most acts are realised by a wide range of utterance types. For example, the act called 'starter' can be realised by either a statement, question, command, or a moodless item.

We distinguish the following 15 acts, marker, summons, silent stress, starter, metastatements, conclusion, informative, elicitation, directive, comment, accept, reply, react, acknowledge, preface, prompt, null. Their definition is always given in terms of both linguistic structure and function. Thus, for instance, the starter act, whose structure has been described above, has the function of providing information about, directing attention to, or thought towards an area, in order to increase the likelihood of making a correct response to the subsequent initiation. An informative act has the structure of a statement and serves the function of providing information. Some of these acts are more directly relevant to the problem of analysing person machine dialogues than others. However, together they provide a broad framework, incorporating both verbal and non-verbal aspects of dialogue, within which we can situate a detailed analysis of the more relevant acts.

42

2.2 MOVES

At a higher level seven types of move are recognized: delineating, sketching, opening, supporting, challenging, bound-opening, and re-opening.

Moves are the basic units of a dyadic exchange and can consist of a number of acts. We will define the various moves in a semi-formal manner. The formalism is a type of context-sensitive grammar. If an element is enclosed in { } it indicates that the item is optional. The symbol | indicates that the elements it separates are alternatives. If an element is enclosed in < > it indicates that the symbol is non-terminal and requires further expansion. The following is the formal description of a sketching move:

```
<sketching   move>   := {<signal>} {<pre-head>} <head>
<signal>             := marker | summons
<pre-head>           := starter
<head>               := metastatement | conclusion
```

Similar descriptions are also given for the other moves.

In addition to these standard conventions we adopt a special convention to illustrate the effects of context. This takes the form:

C[<arg1> <arg2>] := expansion

Where C is a function which, when evaluated, expands <arg2> subject to something being true of <arg1>. For example, we might want to restrict the expansion of <arg2> to only those situation in which <arg1> has already occurred as an act in the previous move. This can be accomplished by an appropriately defined C function.

2.3 EXCHANGES

Sequences of moves compose an exchange. According to Burton two types of exchange are distinguished: explicit boundary (EB) exchanges and conversational (C) exchanges. Explicit boundary exchanges occur, as their name suggests, at the boundary of transactions. They can include delineating and/or sketching moves which must be supported by another speaker.

Conversational exchanges consist of chains of opening, challenging and/or supporting moves. EB exchanges have the following structure:

```
<EB-exchange>   := {<delineating move>}
                   {<sketching move>}
                   <supporting move>
```

The structure of a C-exchange is as follows:

$$<\text{C-exchange}> := <\text{initiation}>$$
$$\{<\text{supporting move}>$$
$$\{<\text{bound-opening move}>$$
$$<\text{supporting move}>$$
$$\{<\text{supporting move}>\}^{l}\}^{m}\}^{n}\}$$

$$<\text{initiation}> := <\text{opening move}>|<\text{challenging move}>|$$
$$<\text{re-opening}>$$

The superscripts l, m, and n represent sets of numbers of repetitions. Where l contains m numbers, m contains n numbers, and n consists of just one number. The numbers, the repetition factors themselves, can range from 0 upwards. However, when n is 0, m and l are also 0. This arrangement allows us to generate a different bound-opening/supporting pattern for each m.

2.4 TRANSACTIONS AND INTERACTIONS

Transactions and interactions are the final two levels of the classification hierarchy. Transactions consist of patterns of exchanges, and interactions consist of unordered strings of transactions. Transactions have the following form:

$$<\text{transaction}> := \{<\text{EB-exchange}>\}$$
$$C[<\text{opening move}> <\text{C-exchange}>]$$
$$\{C[<\text{challenging move}> <\text{C-exchange}>]$$
$$\{C[<\text{re-opening move}> <\text{C-exchange}>]\}^{l}\}^{m}$$

The C function in this case expands its second argument if the first argument, a move, is the initiator of the C-exchange.

Finally interactions take the form:

$$<\text{interaction}> := <\text{transaction}>^{l}$$

3. THE COMMUNICATIVE COMPONENT

The following two assumptions form the basis for a pairing of a dialogue planning mechanism with elements of the dialogue description system:

- dialogue participants always have two cooperating types of goals, substantive real life goals (S-goals), which determine "what to say", and linguistic/communicative goals (C-goals), which determine "how to say it". No relation of necessity seems to hold between them. In fact, in most cases there are many different ways of expressing the same goal.

- it is possible to identify hierarchical relations between goals and subgoals both for substantive and communicative goals. However, in the high level dialogue description system units S-goals seem to be more important, while at the low levels C-goal seem to prevail.

The highest level of the discourse structure is the transaction. Given that the dialogue as a whole is a means of effecting the high-level goals of one or other of the participants, we can functionally define transaction as the unit of dialogue concerned with effecting these high-level goals.

At the highest level of the dialogue's goal structure the dominant goals motivating the transaction are those concerned with the substance of the dialogue, not the means by which the substance is conveyed. As we move down this hierarchy it is possible to discern a bifurcation of goals into one group concerned with the substance of the dialogue, and the other concerned with communication of this substance. These are the S-goals and C-goals (or more properly, S and C subgoals) mentioned above.

A transaction is always motivated by a general goal such as seek information, make a train journey (Allen, 1982), make a reservation. The social context, for example the relation between speaker and hearer or simply a social convention, can suggest rhetorical choices. Among these might be the direct stating of the general goal, the indirect revelation of the general goal by several related questions, the questioning of a system's general capabilities before asking, and so on.

Essentially, exchanges can be thought of as the topic-bearing elements of the dialogue. New topics are introduced by either an EB-exchange (explicit boundary), a C-exchange (conversational) with an opening move as its first move, or a C-exchange with a challenge as its first move. Topics that have been discussed prior to the most recent challenge are re-introduced by a C-exchange with a re-opening move as its first move. Topics that have been discussed less recently are re-introduced by means of an EB-exchange containing a sketching move.

There is, therefore, a strong connection between the exchange structure of the transaction and the pattern of topic-shifts in the dialogue. These topic-shifts, in turn, are related to the conversant's shifting goal structure. This is especially true in task-oriented dialogues, where the component operations of the task are mirrored in the topic structure of the dialogue (Grosz, 1981). A general goal can be, therefore, split into a sequence of subgoals both because the task consists in reality of a sequence of subtasks (Grosz, 1981) and because of rhetorical reasons. This gives a very special status to exchanges in our classification system. A transaction is in fact, divided into several exchanges determined either by the structure of the task to be carried out, or by rhetorical considerations, or by both. In particular, we should distinguish between two types of exchange:

- a subtask exchange, which aims at reaching some substantive subgoal

- an instrumental exchange, which aims at attaining some communicative (sub)goal, such as introducing the terms of the conversation, or clarifying some unclear substantive goal or subgoal.

Every exchange is composed of moves and is, in most cases, opened by some form of topic shifting move and closed by a concluding move. Other subcategories of exchanges can be found. However, a pattern of moves is associated with every exchange. At this level different rhetorical choices (motivated by C-goals) may appear in the form of different distributions of instrumental or non-clarifying exchanges within a transaction. The S-goals of an exchange can be computed from the interpretation of the utterances comprising it, utilising some notion of general focus (Sidner, 1979).

At a lower level, the goals that motivate the moves are drawn equally from S-goals and C-goals. At the move level, the S-goal structure becomes less relevant to the sequence of moves. Moves are mostly rhetorical elements. They signal the pushing to a (new) topic (topic shifting and topic re-introducing moves), the continuation of a topic (topic

maintaining moves), and the popping of a topic (non-introduction).

Pushing and popping are the opening and closing moves of an exchange. The coherence of the topic is to be expected within an exchange and should be perhaps checked from one move to the other. A topic shift is in itself a pushing move. The notion of topic probably coincides with the notion of focused goal.

Moves serve as syntactic components of exchanges and every move is a step in a linguistic goal structure. Therefore, every exchange consists of a pattern of moves representing the communicative choices of the dialogue participants. In some cases, there is a correspondence between the S-goal associated with an exchange and some move (challenge). Also a communicative adherence between one move and the following should hold. Coherence means that the topic must be roughly the same, while adherence means that the given move can be followed only by a specific set of moves.

Finally, an act is a limited linguistic act, uttered to effect very local communicative/rhetoric goals. In a dialogue most moves consists of a single act, but this does not always hold. A distinction can be made at the act level between acts which are drawn from a limited class of utterances, and acts which are not limited in this way. We will call these closed and open classes, respectively. Most of the closed class items are associated with acts which subserve C-goals rather than S-goals. This is not too surprising since the items from the closed class do not usually convey substantive information but usually serve as go-ahead signals in a dialogue. Whereas the primary function of open class items is to convey task-relevant information.

4. SYNTACTIC STRUCTURE AND GOAL STRUCTURE

From the previous paragraph it has been clear that there is a different distribution of S and C-goals along the hierarchy of dialogue units. In particular, higher level units are more related to S-goals, while low level units are connected to C-goals.

The model presented is intended to serve mainly descriptive and representational purposes. No definition is given of the process of inference of the goal structure from the syntactic structure of dialogue. However it is possible to imagine that such a process rely, among others, on

- the functions assigned to dialogue subunits

- the actions mentioned in any specific utterance.

5. CONCLUSIONS AND PERSPECTIVES

The advantages of the proposed dialogue description system are the following:

- the design of a grammar for the description of dialogue units and subunits is made possible;

- the distinction between S- and C-goals allows the treatment of possible interrupt and clarification subdialogues in the same frame as the goal directed parts of the dialogue;

Further research will be devoted to the specification of a more detailed grammar of exchanges and moves, and to the establishing of a stricter correspondence between more types of goals. However, a still more important point to be clarified is the specification of the formal devices by which the semantic (goal) structure can be inferred from utterance and act elements and 'raised' to the higher dialogue units.

6. REFERENCES

Allen, J. (1982). ARGOT: a system overview. Technical Report 101, Department of Computer Science, University of Rochester, Rochester, New York.

Burton, D. (1981). Analysing spoken discourse. In M. Coulthard, & M. Montgomery (Eds.), Studies in discourse analysis (pp. 61-81). London, England: Routledge & Kegan Paul.

Grosz, B. (1981). Focusing and description in natural language dialogues. In A. Joshi, B. Webber, & I. Sag (Eds.), Elements of discourse understanding (pp. 84-105). Cambridge, England: Cambridge University Press.

Reichman, R. (1984). Extended person-machine interface. Artificial Intelligence, 22, 157-218

Sidner, C. L. (1979). Toward a computational theory of definite anaphora comprehension in English discourse. Technical Report 537, MIT Artificial Intelligence Laboratory, Cambridge, MA.

Sinclair, J., & Coulthard, M. (1975). Towards an analysis of discourse. London, England: Oxford University Press.