Generating Interactive Explanations

Alison Cawsey*
Computer Laboratory, University of Cambridge
New Museums Site, Pembroke St, Cambridge, England
Alison.Cawsey@uk.ac.cam.cl

Abstract

Existing approaches to text generation fail to consider how interactions with the user may be managed *within* a coherent explanation or description. This paper presents an approach to generating such interactive explanations based on two levels of discourse planning – content planning and dialogue planning. The system developed allows aspects of the changing context to be monitored with an explanation, and the developing explanation to depend on this changing context. Interruptions from the user are allowed and dealt with (and resumed from) within the context of that explanation.

Introduction

Complex explanations and descriptions are required in many existing computer applications. Tutorial, advisory and help systems may all sometimes need to present some complex piece of information, which cannot be reasonably be presented in one 'chunk'. In human discourse this may result in complex 'explanatory dialogues', where an expert attempts to explain something, checking the novice's understanding as the explanation progresses, and allowing interruptions and clarifications from the novice.

This type of dialogue presents new problems for text planners. It is no longer possible to decide beforehand all the details of what is going to be said, as the interactions with the user may mean that the context (such as the system's assumptions about the user's domain knowledge) may change as the dialogue progresses. The evolving explanation should reflect that changing context. At the same time it is important to preserve the global coherence of the discourse, as each exchange with the user contributes to the same overall communicative goal.

One way to maintain that overall coherence while allowing the changing context to influence the details of the explanation is to incrementally plan the explanation, interleaving planning with execution. At any point the future explanation plan will be represented by a number of high level sub-goals still to be satisfied if the explanation is to be completed, but the order in which those goals are satisfied and the way they are realised will depend on the changing context. This paper will show how interactive explanations may be planned in this way using two levels of discourse planning: content planning and dialogue planning. The content planning level is concerned with determining what to include in an explanation, while the dialogue planning level is concerned with the overall organisation of the dialogue and with managing the interactions with the user. Interactions with the user may cause the context to change, and therefore influence future detailed content planning.

The approach taken is based on an initial analysis of human explanatory dialogues, and on work on text planning, discourse analysis and user modelling. It is described in more detail in (Cawsey, 1989). The system developed (the EDGE system) generates tutorial explanatory dialogues in the domain of electronic circuits, though the basic approach may be applied to other domains and types of discourse.

Related Research: Previous work on text planning has been concerned with generating coherent paragraph length texts given some communicative goal or pool of knowledge to convey (McKeown, 1985; Hovy, 1988; Paris, 1988; Moore & Paris, 1988). McKeown showed how common discourse strategies or schemata could be used in conjunction with focus rules to produce coherent descriptions given some initial pool of knowledge. Paris extended this approach to show how the strategies selected should depend on the expertise of the user. More recently, a general theory of text coherence, Rhetorical Structure Theory (Mann & Thompson, 1987), has been used as the basis for text planners. Hovy, for example, uses the theory to constrain the organisation of a text given what to say, while Moore uses the theory in a more goal directed fashion to also determine what to say given some communicative goal. Moore's approach shares several features with that described here, as she is concerned with interactive discourse, and how follow-up questions may be answered in context. However, she is not concerned with the global coherence and representation of that continuing interactive discourse, or how checks and interruptions may be managed.

*This research was carried out while the author was at the Department of Artificial Intelligence, University of Edinburgh, supported by grants from the Science and Engineering Research Council

Other relevant research includes work on goal-oriented dialogues (e.g., Grosz, 1977; Grosz & Sidner, 1986) and on plan inference in continuing dialogue (Carberry, 1989; Litman & Allen, 1987). This work accepts that a complex dialogue may be coherent by the fact of being related to a single dominant overall communicative goal or intention. This paper will show how interactive discourse may be *planned* given such a communicative goal, taking into account both the domain (and user) dependent organisation of the explanation content, and the content independent conventional organisation of the dialogue. The dialogue representation used is similar to that proposed by Ferrari and Reilly (1986) who show how a dialogue has two levels of organisation – the goal structure discussed above, and a level describing how different types of dialogue units may be organised.

Example Human Dialogues: In order to illustrate some of the problems involved, consider the example human explanatory dialogue fragments given in figure 1. These are taken from a corpus of human expert-novice explanatory dialogues of circuit behaviour which were analysed in this research. The first two dialogue fragments illustrate how dialogues (and topics within them) have characteristic opening and closing sequences. In the example given, the opening sequence involves a discourse marker and *meta-comments* on the future discourse, and the closing sequence checks that the participants are ready to finish with the topic. However, in general the opening and closing sequences will depend on the particular type of discourse (such as tutorial or advisory) and the associated roles of the participants.

The next two dialogue fragments (fig. 1, ex. 3-4) illustrate how an explanatory dialogue may include both checks on the novice's understanding initiated by the expert, and interrupting clarification questions initiated by the novice. Both of these are important in generating an explanation which is understandable by and acceptable to the novice and both may result in changing assumptions about the user's level of expertise. Following such an interaction the remaining explanation may then continue more or less simply depending on these changing assumptions.

The Problem: The system described in this paper aims to show how explanatory dialogues with the features illustrated above may be generated, concentrating on the following interrelated issues:

- How should interactive discourse be planned, given a communicative goal, taking into account both the domain dependent goal structure, and the domain independent conventional organisation of the dialogue.
- How should the representation of the discourse context (including the user model) be updated?
- How should the user's understanding be checked, and interruptions dealt with within that dialogue.
- How should the remaining discourse be modified as the perceived context changes?

(1) Opening Sequence
E: Right, What I'm going to do is to get you to explain this last circuit to me. Before I do that I better say briefly what a comparator is.

(2) Closing Sequence
N: OK
E: Is that sufficient?
N: I think I know what's going on with it.

(3) Checking Level of Understanding
E: OK, do you remember anything about transistors?

(4) Interrupting Clarification Question
E: These components here, you might consider them as being both resistors. Two variable resistors. I can write down a relation for resistance..
N: You'll have to tell me what a resistance is.
E: A resistor is just ...
(clarification sub-dialogue describing resistance)
N: I see, or at least, I think I see.
E: Well, in this circuit here there are just two resistors, ...
(explanation continues more simply)

Figure 1: Example Human Explanatory Dialogue Fragments (E=Expert, N=Novice)

The following four sections attempt to address each of these issues.

Planning Interactive Discourse

In order to plan coherent interactive explanations we need consider two things: how the content of the explanation is organised, and how the dialogue with the user is organised. These two types of structure are largely independent, and so should be addressed separately. The EDGE system therefore has separate content and dialogue 'planning' rules, which act together to generate interactive explanations. These are discussed below.

Planning Explanation Content

If we are to generate both coherent and understandable content it is useful to consider two types of relation which exist between sections of an explanation. *Coherence* relations (considered in the work on text planning mentioned above) exist between text sections, while *pre-requisite* and *subskill* relations may exist between the topics explained in the text. These latter relations are discussed widely in work on curriculum planning within the field of Intelligent Tutoring Systems (e.g., Murray, 1989).

In the kinds of complex explanations considered here the latter types of relation appear to be more important than coherence relations (though the latter obviously exist). Prerequisite and subskill relations between topics also begin to explain some of the empirically derived discourse schemata suggested by McKeown (1985) and particularly by Paris (1988). For example, prerequisites to understanding how an object works (the causal trace) may include understanding

```
c-plan how-it-works (device)
   preconditions: know-user identity (device)
                  know-user function (device)
                  know-user structure (device)
   subgoals: c-goal process (device)
             c-goal behaviour (device)

c-plan function (device)
  constraints: know-user ('function
                          (device-analog device))
  subgoals: c-goal compare-contrast
                 ('function device
                          (device-analog device))

c-plan identity (device)
   constraints: device-parent (device)
   subgoals: d-goal teaching.exchange
                  ((list 'isa device
                          (device-parent device))
```

Figure 2: Example Content Planning Rules

what sort of object it is (identification) and understanding its structure (constituency).

The content planning operators used therefore aim to capture these relations between sub-topics, while allowing for alternative discourse strategies to be used to explain these sub-topics in different ways. Planning rules may have a name, arguments, constraints, preconditions, and subgoals. Three example content planning rules are given in figure 2. The name and arguments of a rule (e.g., `how-it-works (device)`) describe the goal that the rule is used to achieve – for content plans this is the concept that the rule is used to explain. Preconditions are goals which will only be satisfied if not already true, and will normally refer to the user's knowledge. Subgoals may be dialogue or content goals (and are labelled as such). Constraints are used to select between alternative ways of satisfying the same goal, and may refer to the user's knowledge, the domain knowledge base or to the state of or type of the discourse. Expressions in brackets are evaluated given the current bindings of the plan arguments, using the normal lisp evaluator.

The second rule, for example, is used to select a *compare-contrast* strategy for describing the function of a device, applicable if the user knows the function of an analogous device. The third rule is used to describe the *identity* of a device by setting a discourse goal to have a teaching exchange about the proposition that the device is an instance of some other parent device. It will apply if the relevant knowledge (`device-parent (device)`) is in the knowledge base.

The content planning operators as they stand have a number of problems, failing to make distinct different types of relations and discourse strategies. For example, domain independent relations in the text such as *background* or *conclusion* are only implicit, while alternative strategies for explaining the same thing (e.g., *compare-contrast*) are not treated in a truly domain independent manner, being incorporated in the topic relation based planning operators. However, the approach exploits relations which are de-emphasised in other systems, and is effective in generating extended explanatory texts. The rest of the paper is independent of the particular content planning method chosen, assuming only some incremental, hierarchical decomposition of goals (as used in Moore and Paris' work (1989) for example).

Dialogue Planning

Work in the field of discourse analysis (e.g., Sinclair and Coulthard, 1975) has shown how relatively formal types of dialogues (e.g., classroom, court) have a regular hierarchical structure. The details of that structure depend on the type of dialogue, but are largely independent of the domain content.

In order to capture the hierarchical structure of explanatory discourse we define dialogue planning rules in a similar manner to content planning rules. These are based on the levels of description given in Sinclair and Coulthard's work (and related work on discourse analysis), with the details of the rules adapted from this work to explanatory discourse. In particular there are four main rule types corresponding to the main categories used in Sinclair and Coulthard's work – the *transaction* on some topic, the *exchange*, the *move* within that exchange and the *act*. A transaction will normally be composed of particular sequences of exchanges, exchanges of moves and moves of linguistic acts. These levels of description have been used in a number of recent dialogue systems (e.g., Wachtel, 1986; Ferrari & Reilly, 1986), but not for generating and controlling a dialogue given a communicative goal. The framework allows the content-independent conventional organisation of a dialogue to be defined and used to guide the selection of interactions with the user.

Example high level planning rules are given in figure 3. The first rule constrains a transaction, or discussion of some major topic to consist of an opening exchange, some exchanges on the topic and a closing exchange. The second rule captures the characteristic opening sequence for topics in tutorial discourse – the teacher provides a framing move (Sinclair and Coulthard's term for a topic opening discourse marker) and a focussing move (a meta-comment about the future discourse).

Dialogue planning rules may take content level goals as their arguments. For example, the argument `c-goal` in the example above will be a topic to be explained, such as `how-it-works (heat-detector)`. The `teaching.exchanges` rule will cause that content level goal to be set as a goal to be satisfied, so content planning rules may be used to decide how to make that topic known to the user.

Planning an Explanation

The dialogue and content planning operators are used in conjunction to plan an explanation. In general high level dialogue planning rules such as the *informing*

```
d-plan informing.transaction (c-goal)
   subgoals: boundary.exchange ('open c-goal)
             teaching.exchanges (c-goal)
             boundary.exchange ('close c-goal)

d-plan boundary.exchange (exchange-type c-goal)
   constraints: equal (exchange-type 'open)
                equal (discourse-type 'tutorial)
   subgoals: frame.move (exchange-type)
             focussing.move (exchange-type c-goal)
```

Figure 3: Example Dialogue Planning Rules

transaction given above are used to plan the overall organisation of the dialogue. However, as mentioned above, subgoals may be posted to make some particular topic known to the user. Given such a subgoal, control switches to the content planning rules, which are used to determine the content of the sequence of exchanges used to explain that topic. The content planning rules may in turn set lower level dialogue goals, such as to have some kind of exchange with the user about some proposition (as in the last rule in figure 2). Dialogue planning rules are then used to control how that exchange with the user is realised.

The planning process proceeds incrementally, so the future plan is not fully determined before the explanation begins. Planning begins by putting a goal – the principle goal/purpose of the discourse – on an agenda. It proceeds by selecting a goal from the agenda depending on its priority, selecting which planning rule to use (based on constraints on these rules), and using this rule to find new subgoals of this goal to put on the agenda . Interactions with the user may result in further goals being put on the agenda, which may be realised as clarification sub-dialogues, for example. They may also cause assumptions about the user's knowledge to be revised, which will in turn influence the future detailed explanation plan.

Updating the Discourse Context

If we are to generative interactive explanations which depend on the changing context it is important to be able to represent and update that discourse context. In this work we are concerned with three important aspects of this changing context:

- Building up a hierarchical model of the discourse so far, and using this when, for example, managing interruptions.
- Recording the currently salient objects and discourse segments, giving a simple model of focus (cf. Grosz & Sidner, 1986), and using this both in pronoun selection and in influencing content ordering.
- Updating assumptions about the user's knowledge from their interactions with the system and using this to influence the content of the continuing explanation.

These aspects of discourse context have all been discussed widely in the literature, but have not been used together in a discourse generator of this sort. The EDGE system allows them to both be updated and to influence the continuing discourse.

The first two are updated in a very simple way. The hierarchical structure of the discourse is reflected by the goal/subgoal structure of the instantiated planning rules, while a simple model of focus can be taken from the arguments of the planning rules. The different types of planning rules (content and dialogue) include in their arguments domain objects, discourse segments and propositions, so all these will be in the simple representation of focus.

The user model is updated after each exchange with the user. Dialogue planning rules may describe both the structure of different types of exchanges, and the effect of these exchanges on the systems assessment of the user's knowledge. For example, an exchange where the user is told something and asks the system to continue allows the system to assume that they probably understood what they were told. An exchange where the user is asked, and correctly answers some question allows the system to assume that they certainly know the answer. Based on these direct inferences the system may also use a number of indirect inference rules to update assumptions about other propositions, higher level concepts, or the user's general level of expertise (cf. Kass & Finin, 1987). For example, if a concept is believed known then prerequisite concepts may also be believed probably known, and if a user asks a question about something very basic the system may revised their assumed level of expertise.

Managing Interruptions

If the user interrupts in the middle of the explanation with some clarification question, the system must be able to respond to the problem (if appropriate), yet resume the previous interrupted discourse in a clear way so that the explanation may be completed.

In the EDGE system interruptions are dealt with using a special interruption dialogue planning rule, which determines how interruptions should be opened and how the previous discourse should be resumed[1]. Responses to interrupting clarifications begin with a discourse marker such as 'well' or 'OK', and end (if the discourse focus has changed) with either a repetition or meta-comment on the interrupted discourse. The meta-comment (as given in the example below) refers to the lowest level interrupted discourse segment in the discourse hierarchy, found by examining the discourse model. The example below illustrates how the system handles a very short interruption.

System: .. the output voltage is high.
User: What's this component?
System: OK, This component here is a light dependent resistor.

[1]This may be compared with Litman & Allen's use of meta-plans in dealing with clarifications in discourse interpretation (Litman & Allen, 1987).

Anyway, we were going through what the light detector circuit does.
When the light detector circuit has a high input light intensity...

Note that in this example both the user and the system may refer to 'this' component by pointing at a diagram. This aspect of the system is described in (Cawsey, 1989).

Modifying the Evolving Discourse

After any exchange with the user the discourse context will have changed, as described earlier. As the remaining explanation is only partially planned out in advance, and the planning depends heavily on the context, the details of the remaining explanation will depend on that changing context.

Three aspects of the explanation depend on the context and so will be influenced by such changes:

Discourse Strategy: Different strategies, both influencing the dialogue and content may be selected depending on context. Constraints on planning rules refer to the current context, and constrain which strategies may be selected. For example, in the second rule in figure 2 a *compare-contrast* strategy could be selected if a similar device was known, while in figure 3 the form of a topic opening exchange depended on the discourse type. If the context changes so that, for example, the user's asssumed knowledge changes, the system will respond appropriately.

Discourse Content: The changing assumptions about the user's knowledge will also influence what is included and what is left out of an explanation. If prerequisite knowledge is already believed known it will be left out. If it is unknown whether it is known then the system will ask a question to check before deciding whether to include it.

Ordering of Content: If the domain objects in focus change this may influence what goals on the agenda are considered next, and hence the ordering of the explanation. This may occur following a clarification question from the user, introducing new objects into the discourse. The system uses the simple heuristic of (all things being equal) trying to maintain the current topic, though it could easily be extended to take account of more complex kinds of focus shifts (e.g., McKeown, 1985).

Example

In order to illustrate the points above we will consider how the system generates the example given in figure 4[2]. Initially the system is given the task of explaining to the user how a 'heat detector circuit' works, and puts the following goal on the agenda:

```
informing.transaction
    (know-user how-it-works (heat-detector))
```

This goal indicates that the system should initiate a particular type of dialogue – an informing transaction – which has as its (content) goal that the user should know how a heat detector circuit works. The appropriate dialogue planning rule is selected (figure 3, first rule) and after further application of lower level dialogue planning rules the opening sentence generated. Control then passes to the content planner which begins to plan a sequences of exchanges to make known to the user how the heat detector works. In this case the first and third rules in figure 2 are used in selecting to inform the user that the heat detector is a potential divider circuit, the details of this *teaching exchange* being determined by the dialogue planner.

After the system's first utterance the following goals will be on the agenda:

```
know-user function (light-detector)
know-user components (light-detector)
c-goal process (light-detector)
c-goal behaviour (light-detector)
```

This is a partial explanation plan to make sure that the user knows the function and components of the light detector, then to explain the process and behaviour. Now, after the user's clarification question in utterance 2 the system checks this agenda and notes that the user's question (about device *behaviour*) will be answered later, in (arguably) a more appropriate place. The system therefore delays answering this question and goes on to continue planning the explanation. Now, the next goal selected from the agenda is `know-user function (light-detector)`. As the system does not know whether the user already knows this it checks with the user at the end of utterance 3, and following the user's response describes the purpose of the device to the user (utterance 5).

Now, the user's next question (utterance 6) is answered as it is not part of the planned explanation. The interruption comes in the middle of the system describing the components of the heat detector, and the system selects next to describe the thermistor (changing the default order of topics to maintain some continuity with the clarification sub-dialogue). Note that if it were not possible to continue the interrupted topic in this way the system would give a meta-comment on the interrupted discourse as described earlier.

Following the description of the components, the system begins the main part of the explanation of how the heat detector circuit works, explaining the process given different input values. (Note that the labels H1, R1 etc. refer to a circuit diagram which is displayed). Eventually the system will conclude with a summary of the circuit's overall behaviour.

Different types of exchanges with the user cause the user model to be updated in different ways. For example, if the user asks a question about something (as in utterances 2 and 6) the system assumes they don't know this information. This may in turn cause a possible revision of the assigned user stereotype. The user model influences the explanation in various ways.

[2] Note that in this example user questions are selected from menus. These menus are either fixed (given a particular explanation) or obtained by clicking on a diagram. Sentence level generation of the system output is achieved using simple template-based techniques, though could be improved by interfacing the system to a full natural language sentence generator.

1 **System:** I'm going to explain how the heat detector circuit works. It is a sort of potential divider circuit.

2 **User:** What does a heat detector circuit do?

3 **System:** I'll be getting on to that in a minute. Do you know what it is for?

4 **User:** No.

5 **System:** OK, Its purpose is to provide an output voltage which depends on the input heat intensity. It has the following components: a fixed resistor and thermistor.

6 **User:** What's this component?

7 **System:** This component here is a thermistor. The thermistor is a bit like a light dependent resistor except its output resistance depends on input heat intensity and not on input light intensity. The fixed resistor always has a medium output resistance. Now, lets go through how the heat detector works given different input heat intensity values. Suppose the input heat intensity is high. When the thermistor has a high input heat intensity (H1) its output resistance (R1) will be quite low. When the potential divider circuit has one quite low input resistance (R1) and one medium input resistance (R2) then its output resistance is quite high. Suppose ...

Figure 4: Example EDGE Explanatory Dialogue Fragment

Here the system has been told that the user is a relative novice, so may not know what the purpose of the circuit is, for example. If the system believed that the user probably knew this it would be omitted from the explanation. The user is assumed (from a previous session) to know what a light dependent resistor is like, so a simple comparison is drawn in utterance 7.

Conclusion

This paper describes an approach to generating interactive explanations. The system described has been fully implemented, and used to generate interactive graphical/textual explanations of circuit behaviour. The system has been evaluated with a small number of naive users who found the approach helpful and the explanations coherent.

The EDGE system shows how dialogue planning rules and content planning rules may be used in conjunction to generate interactive discourse. The user may interrupt after any utterance to ask clarification questions, and the system may check the user's understanding as the explanation progresses. As the context changes (given these different exchanges with the user) so will the development of the remaining explanation.

The simple hierarchical decomposition used in planning the explanation allows for a practical, efficient system. At any point the future explanation plan will be only partially determined, so the changing context will naturally influence the details of the plan while retaining global coherence.

Future work should concentrate on combining this approach with other similar explanation planning systems which do not allow for interruptions and checking moves (e.g., Moore & Paris, 1989). We also need to consider further, for example, cases where the partial future explanation plan becomes completely inappropriate and needs to be abandoned.

References

Carberry, S. 1989. Plan Recognition and its Use in Understanding Dialogue. In Wahlster, W. and Kobsa, A. (eds) User Modelling in Dialogue Systems, pp. 163-162. Springer Verlag.

Cawsey, A. 1989. Generating Explanatory Discourse: A Plan-based Interactive Approach. Ph.D. diss., Department of Artificial Intelligence, University of Edinburgh. To be published by MIT press.

Ferrari, G. and Reilly, R. 1986. A Two-Level Dialogue Representation. In Proceedings of 11th International Conference on Computational Linguistics, Bonn.

Grosz, B.J. 1977. The Representation and Use of Focus in Dialogue Understanding, Technical Report, 151, SRI International.

Grosz, B.J. and Sidner, C.L. 1986. Attention, Intentions and the Structure of Discourse. *Computational Linguistics* 12(3):175-204.

Hovy, E.H. 1988. Planning Coherent Multisentential Text. In Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics, Buffalo, New York.

Kass, R. and Finin, T. 1987. Rules for the Implicit Aquisition of Knowledge about the User. In Proceedings of the Sixth National Conference on Artificial Intelligence.

Litman, D.L. and Allen, J.F. 1987. A Plan Recognition Model for Subdialogues in Conversations. *Cognitive Science* 11:163-200.

Mann, W.C. and Thomson, S.A. 1987. Rhetorical Structure Theory: A Theory of Text Organisation. In L. Polanyi, ed. The Structure of Discourse. Ablex Publishing Corporation, Norwood, N.J.

McKeown, K.R. 1985. *Text Generation.* Cambridge University Press, Cambridge, England.

Moore, J.D. and Paris, C. 1989. Planning Text for Advisory Dialogues. In Proceedings of the Twenty-Seventh Annual Meeting of the Association for Computational Linguistics, Vancouver, B.C., Canada.

Murray, W. 1989. Control for Intelligent Tutoring Systems: A Blackboard-Based Dynamic Instructional Planner. In *Artificial Intelligence and Education* ed. Bierman, D., Brueker, J. and Sandberg, J. pp. 150-168. IOS.

Paris, C. 1988. Tailoring Object Descriptions to a User's Level of Expertise. *Computational Linguistics* 14:(3):64-78.

Sinclair, J. McH. and Coulthard, R.M. 1975. *Towards and Analysis of Dsicourse: The English Used by Teachers and Pupils.* Oxford University Press, Oxford, England.

Wachtel, T. 1986. Pragmatic Sensitivity in Natural Language Interfaces and the Structure of Conversation. In Proceedings of 11th International Conference on Computational Linguistics, Bonn.