

PLAN REVISION IN PERSON-MACHINE DIALOGUE

Cléo JULLIEN
Jean-Charles MARTY
Grenoble Research Center of CAP SESA INNOVATION
33, Chemin du Vieux Chêne
38240 Meylan France
e-mail: jullien@capsogeti.fr, marty@capsogeti.fr

Abstract

Dialogue is a cooperative process in which each speech act of the participants contributes to the overall purpose of the dialogue. Participating in a full dialogue implies understanding at each point of the dialogue session the role of each speech act with respect to the rest of the dialogue. We concentrate in this paper on speech acts that diverge from the straightforward unfolding of the dialogue. Such speech acts represent dialogue deviations. We analyze the representation of different types and degrees of deviations and present a plan revision mechanism for dialogue management that permits their treatment in the context of advice giving dialogues between a novice user and an expert problem solver.

1 A Sample Dialogue

In ESPRIT project ESTEAM-316¹, we are designing a dialoguer that manages the interaction between a novice user and

¹ESTEAM-316: "An Architecture for Interactive Problem Solving by Cooperating Data and Knowledge Bases" supported in part by the CEC.

advice-giving expert systems. The tested application is personal financial investment [5]. As an introduction to our paper, we present a dialogue in which the system obtains from the user the information needed to activate the financial expert system.

01,S: *Hello. Tell me about your investment problem.*

02,U: *I'd like to buy a car in two years.*

03,S: *Ok. Normally an investment plan comprises four subparts: an emergency fund, for unforeseen expenses; a cash need when you know you'll need a certain sum at a certain date, as is the case with your car; a fixed-term investment to earn the maximum in a specified time period; and a long-term investment for an indefinite amount of time which can be designed either for capital growth or for a regular income. How much do you want to invest?*

04,U: *\$20000.*

05,S: *How much do you want to put into an emergency fund?*

06,U: *Let's talk about my car first.*

07,S: *Ok, how much do you think your car will cost?*

08,U: *Oh, let's say \$9000, I guess.*

- 09,S: *And you plan to buy it in two years?*
- 10,U: *Yes.*
- 11,S: *Ok, if you put \$7,500 in a money-market fund earning 9.9% p.a., in two years you'll have \$9,000. Let's return to the emergency fund, now. How much do you want to put in it?*
- 12,U: *\$5000.*
- 13,S: *Tell me also the delay.*
- 14,U: *I don't understand what is the delay of an emergency fund.*
- 15,S: *The delay is how soon you can have your money back after asking for it. So, for your emergency fund, what is this delay?*
- 16,U: *2 days.*
- 17,S: *Do you want a fixed term investment?*
- 18,U: *No.*
- 19,S: *So, it leaves \$7,500 for your long-term investment. Is it ok?*
- 20,U: *I wanted \$10000.*
- 21,S: *If you want an investment plan for \$20,000 with an emergency fund of \$5,000 and a cash-need of \$7,500, then it only leaves \$7,500 for a long-term investment. If you want \$10,000 for a long-term investment, then you should modify something else in your plan.*
- 22,U: *Could I put \$2500 into the emergency fund?*
- 23,S: *Ok. So, to sum up, you want to invest \$20,000, with \$2,500 in an emergency fund with a delay of 2 days, \$7,500 in a cash need to buy a \$9000 car in 2 years, no fixed-term investment, and \$10,000 in a long-term investment. Is that right?*
- 24,U: *That's fine.*

The above sample dialogue has a quite simple structure: presentation of the main sub-investment types and collection of information from the user concerning each sub-investment. We notice, however, that at some points the user does not react directly as the system would expect. For instance, in sentence 6, the user is asked about saving money in an emergency fund, but she wants to talk about buying her car first. Later, in sentence 14, the user is unable to answer a question about the delay of the emergency fund because she does not understand what such a delay is.

Such reactions illustrate dialogue deviations. Dialogue deviations are sentences that diverge from the straightforward unfolding of the dialogue, while having a coherent interpretation with respect to the whole dialogue. These unexpected reactions are inevitable in a dialogue where the participants are independent agents with their own goals and differing degrees of knowledge about each other and the subject under discussion.

Before describing our dialogue manager and its mechanism for handling such deviations, we present in the next section the framework we adopted to model flexible dialogue management.

2 Deviations in Dialogue

Dialogue is considered to be a cooperative activity where the goals and actions of each participant contribute to the overall purpose of the dialogue [3]. In task-oriented dialogues, we distinguish between *task level* goals and plans (e.g., investing, traveling), and *communicative level* intentions and speech acts (e.g., explaining, requesting information) [2,6,11]. We call

these aspects the intentional dimension. Each step in the dialogue concerns a particular topic. Intuitively the notion of topic might be described by a subset of objects of the problem under discussion. In fact, the boundary of this "subset" is not strict: one can only say that some objects are more salient than others. The attentional state is hence better represented by different levels in the focus of attention, corresponding to embedded subsets of objects [9].

It is important to note that in the context of dialogue, the term *deviation* is not used with its strict boolean meaning: it is a complex and a relative notion.

Deviation are complex because they involve both intentional and attentional dimensions. Deviations in this cooperative process arise from inconsistencies between the observed speech act and the dialogue considered as a coherent plan [4]. They can be classified according to the type of interactions among of interactions among intentions of the participants and and changes on the focus of attention. The sample session above illustrates several types of such deviations: at the communicative level, the novice user requests for explanations before giving a requested information; at the task level, the user gives an inconsistent value or does not want a given action in the task plan. Deviations are generally combined with changes of subjects.

Within each dimension there exist different degrees of deviations: speech acts may have indirect effects, changes in the focus of attention may be be more or less abrupt, deviations depend on the expectations each participant has concerning the reactions of the other.

Therefore, we adopted models of the dialogue structure where the relationship be-

tween the intentions and the evolution of the focus of attention are made explicit [10].

The detection and analysis of deviation after a user speech act relies on expectations and predictions in both the intentional and attentional dimensions. The system, having produced a speech act and waiting for a reaction of the user, expects in the first place a reaction corresponding exactly to the effect it intended to produce. If the user reacts differently, the system will use knowledge about possible types of deviation and the state of dialogue to analyze the nature of the deviation.

Once a deviation has been identified, the system may need to modify more or less deeply the planned course of the dialogue: from local adaptation like embedding a small clarification subdialogue (sentences 14-15), to more global revision like re-ordering entire sub-topics (sentence 6). Hence to interpret the influence of an unexpected speech act at a certain point in the dialogue the representation of the state of dialogue should reflect the structure of the whole dialogue: keeping track of past exchanges and anticipating the remainder of the dialogue.

3 Overview of the Dialogue Manager

The general organization of the Esteam-316 Dialogue System [1] is depicted in figure 1.

A Natural Language Front-End (NLF) transforms natural language utterances into literal meaning and vice-versa. The literal meaning corresponds to an isolated surface speech act. The Recognizer takes a literal meaning from the Front-End and determines whether the corresponding

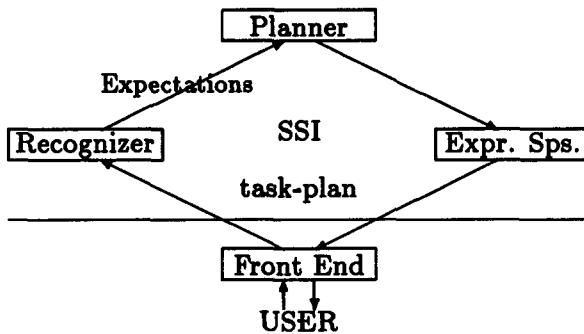


Figure 1: Overview of the Dialogue Manager

surface act of the user could be an example of, or a part of, one of the communicative actions that the system expects from the user in the context of the current dialogue. The expectations are controlled by the Planner which conducts the dialogue and maintains a structure of the system's intentions (SSI), while reacting to user's intentions detected by the Recognizer. The Planner interacts with the Expression Specialists for constructing from the communicative acts it intends to perform (*what* to communicate) an appropriate literal meaning (*how* to say it).

An advice giving dialogue is a particular case of task-oriented dialogue. The task-level plan reflects the problem of the user. The advice giving system has only communicative intentions for constructing and refining the task-level plan. A complete advice giving session contains three phases: problem formulation, resolution and presentation of the solution. In this paper, we concentrate on first phase. During problem formulation, the system helps the user to refine, select and instantiate appropriate subplans according to the user's goals. The task plan is initialized by a stereotypical pattern of actions which could be recommended as part of a "good" solution. The result of the problem formu-

lation phase is a coherent task-level plan which can be passed to an expert problem solver.

4 Representation of the State of Dialogue

The communicative intentions of the system are stored in the Structure of System Intentions: the SSI reflects the state of the plan of dialogue from the system's point of view.

4.1 Communicative Level Plans

The Planner uses a hierarchical set of plans for constructing the SSI. Plans are associated with the various communicative level intentions of the system. We have designed two types of plans: *dialogue-plans* and *communicative-plans*.

Dialogue-Plans are the most abstract plans of the Planner. They express the strategy of the overall advice-giving session [7]. The purpose of these plans is to capture procedural knowledge for an "ideal" advice-giving session. They are used to initiate the SSI, but also include means for adaptation at execution time [8].

Basically the models of dialogue-plans express dominance and sequencing relations among the sub-parts of the dialogue session (*decomposition*), and the part of the task level plan that is in focus at a given step of the communicative level plan (*parameter*).

Communicative-Plans models contain the effects an elementary communicative intention of the system in terms of the immediate expectations about the

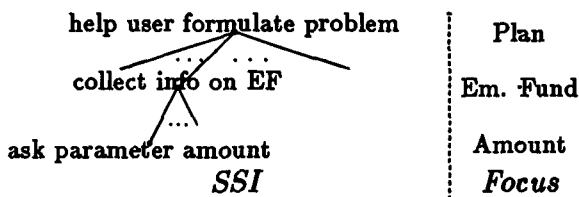


Figure 2: Structure of System's Intentions and Attentional State

communicative acts of the user. For the Planner a communicative plan is seen as a primitive action to be passed to the Expression Specialists for execution.

4.2 Expectation Stack and Structure of System Intentions

The SSI is a tree enhanced by orderings relations, in which each node represents a communicative level plan of the system, and links decomposition relations of plans into subplans. In addition each node in the SSI is related to a given subpart of the task-plan. At a given point in the dialogue the attentional state can be represented by a stack in which the bottom contains the focus associated with the most general plan and the top contains the focus of the plan currently executed. For example, when the system asks for the amount of the emergency fund, there will be three focus levels in the stack: the investment plan at the bottom, the emergency fund in the middle and the amount on top (see figure 2).

Using this attentional state the Recognizer could only analyze changes of focus. The problem is still to provide the Recognizer with expectations concerning the intentions of the user. The method of representing expectations is to attach a set user's communicative acts to each type of system's intentions in the SSI and to or-

give parm (delay, EF) request expl (delay, EF)
give parm (amount, EF) request expl (amount, EF)
request expl (EF) refuse (EF)
Acts concerning Cash Need and other subplans
give parm plan (PI) request expl plan (PI) refuse plan (PI)
Others

Figure 3: A sample State of the Expectation Stack

ganize these expectations according to the attentional state. We obtain the *Expectation Stack*.

Let us consider an example to understand better the significance of the different objects included in the Expectation Stack. Figure 3 represents the state of the Expectation Stack when the system asks for the delay of the emergency fund (sentence 13 or 15).

The most expected answer is the delay of the emergency fund, but the system also foresees the possibility of a request for explanation from the user. At the next level, the system expects the user to speak about another parameter of the same plan (the amount). At the level still further below, he/she can speak about the emergency fund in general, he/she can for example refuse the emergency fund, or ask for explanation on it. And so on, until the system reaches the most unexpected reactions of the user, i.e., even things that are not related to the investment problem. (in the level "others")

5 Dialogue Management: Execution and Revision

In the previous sections, we have presented the different structures needed to reflect the state of dialogue. The aim of this part

is to show how these structures are used for handling revisions during the execution of the dialogue.

The SSI is generated by a depth-first expansion of the abstract communicative goals of the system (use of dialogue-plan models). This process stops as soon as the Planner reaches an atomic action (communicative-plan) that produces a request toward the user. At this point, the Expectation Stack is derived from the state of the SSI and reflects the precise topic of the question on top.

The input of the user is analyzed by the NLF and the Recognizer [1] is called to derive the user's intentions encoded in the answer. The Recognizer returns the speech act in the Expectation Stack it was able to match. In most cases, the returned speech act corresponds to the direct expected answer and the dialogue continues without revision.

A deviation occurs, however, when the Recognizer does not return the level corresponding to the most expected answer, or when the user tries to put inconsistent value in the task plan (an example of such an inconsistent value is given in sentence 20). In this case, the Planner must use consistency constraints attached to the task plan. Thus, the pointer returned by the Recognizer in the Expectation Stack indicates all the changes of subject or of intention by the user while the constraints attached to parameters of the task plan reveal inconsistent values.

There is a *revision* associated with each type of deviation. A revision is a structural transformation pattern for correcting the SSI after the user's answer in order to continue a coherent dialogue.

We illustrate the result of a revision on sentences 5 and 6 (see figure 4).

We can see that the transformation takes

into account a subsequent return to the emergency fund and an explicit re-introduction of this subject.

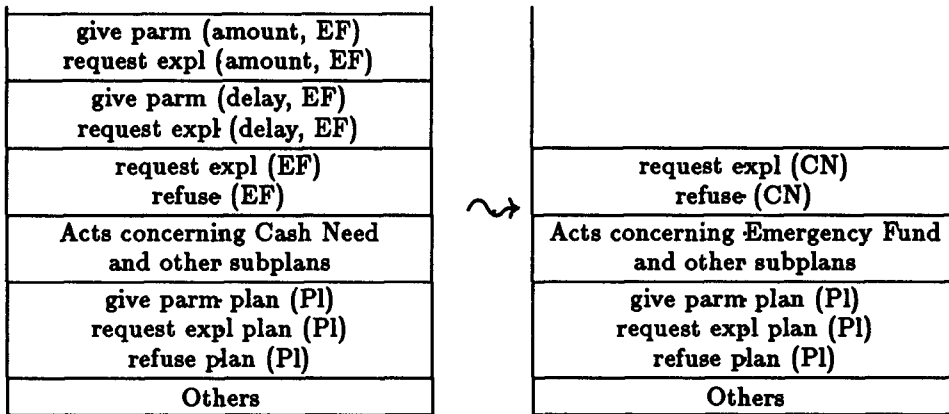
The same types of transformation are used to treat the deviations caused by the requests for explanation (sentence 14). In the case of inconsistent values (sentence 20), a subplan is inserted for explaining why the value is inconsistent and asking the user to change something in order to correct the violation (sentences 21).

It is interesting to note here that the strategy of the dialoguer can easily be modified by changing either the models of the plans or the transformations associated with the cases of deviation. We presented above a strategy in the dialogue that "follows the user" (*i.e.*, the dialoguer is very cooperative and changes the subject each time the user wants to). For instance, when the user decides to speak about his/her car, the system is cooperative and allows him/her to do so. It would also have been possible to adopt another strategy and to say "ok, we will talk about your car later, but now we need to discuss the emergency fund because people tend to forget it otherwise".

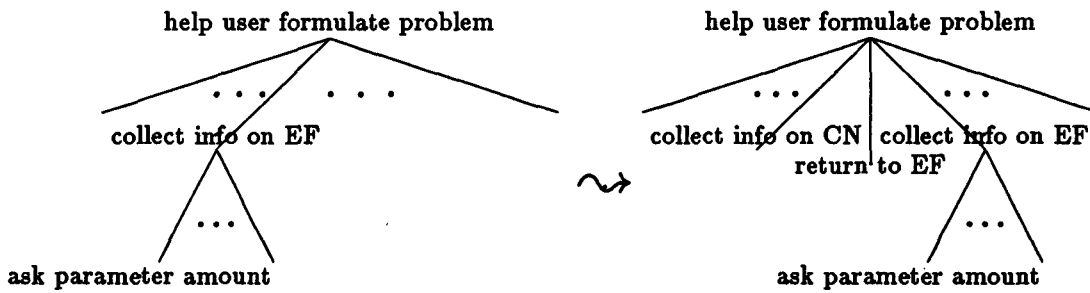
6 Conclusion

The adaptive planning method presented above takes into account both analyses of intentions and changes of topic. Recognition of intentions is organized around the attentional state in order to delimit the scope of the revision. The revision mechanism is currently implemented in an experimental prototype of the ESTEAM-316 Dialoguer (written in PROLOG and running on SUN Workstations).

This approach seems appropriate for the type of dialogues where the "expert" advice giving system has a quite directive



Transformation of the Expectation Stack



Transformation of the SSI

Figure 4: A Example of revision: Changing Subject

control of the dialogue. Revisions give some degrees of flexibility to the "novice" user who is unfamiliar with the domain and the progression in the consultation. Future work will extend the set of revision strategies to take into account deviations that might arise the final phase of an advice giving session, presentation and negotiation of the solution.

References

- [1] ESTEAM 316. *An Architecture for Interactive Problem Solving by Cooperating Data and Knowledge Bases*. Technical Report Deliverable 3, ES-PRIT Program, 1987.
- [2] J. F. Allen and C. R. Perrault. Analyzing intentions in utterances. *Artificial Intelligence*, 3(15):143-178, 1980.
- [3] James F. Allen. Plans, goals and natural language. *Research Review of Computer Science Department, University of Rochester*, 4-12, 1986.
- [4] Carol A. Broverman and W. Bruce Croft. Reasoning about exceptions during plan execution monitoring. *Proc. of AAAI*, 190-195, 1987.
- [5] A. Bruffaerts, E. Henin, and V. Marlair. *An Expert System Prototype for Financial Counseling*. Technical Report Research Report 507, Philips Research Laboratory, 1986.
- [6] Philip R. Cohen and C. Raymond Perrault. Elements of a plan-based theory of speech acts. *Cognitive Science*, (3):177-212, 1979.
- [7] P. Decitre, T. Grossi, C. Jullien, and JP. Solvay. Planning for problem formulation in advice-giving dialogue. *Proc. of ACL*, 1987.
- [8] M. P. Georgeff and A. L. Lansky. *Procedural Knowledge*. Technical Report Technical Note 411, SRI International, January 1987.
- [9] Barbara J. Grosz. *The Representation and Use of Focus in Dialogue Understanding*. Technical Report TR 151, Artificial Intelligence Center, SRI International, 1977.
- [10] Barbara J. Grosz and Candace L. Sidner. Attention, intentions and the structure of discourse. *Computational Linguistics*, 12(3):175-205, 1986.
- [11] Diane J. Litman. *Discourse and Problem Solving*. Technical Report TR 130, Computer Science Dpt., University of Rochester, October 1983.