Plan-based models of dialogue

Håkan Sundblad

Abstract This paper reviews plan-based models of dialogue. The original model is thoroughly presented. Important developments of the original model are also presented. Finally, benefits and drawbacks of plan-based approaches to dialogue modelling are discussed.

Introduction

Plan-based dialogue models have since long earned their place in works intended to present computational linguistics in general and dialogue systems in particular (Allen, 1995; Jurafsky and Martin, 2000; McTear, 2002). Plan-based models of dialogue have accomplished to offer a generalization where dialogue can be treated as a special case of rational non-communicative behaviour (Cohen, 1996).

This paper intends to review plan-based models from when they first appeared and onward. The starting point is taken to be the seminal work of Cohen, Perrault, and Allen (Cohen and Perrault, 1979; Perrault and Allen, 1980; Allen and Perrault, 1980).

Plan-based Models of Dialogue

Plan-based models of dialogue can be claimed to originate from three classic papers: Cohen and Perrault (1979), Perrault and Allen (1980), and Allen and Perrault (1980). Perhaps, the easiest way to introduce this class of models is to start with an intuitive description of what happens when an agent *A* asks another agent *B* a question, to which the latter then responds. The presentation is taken from Allen and Perrault (1980).

> *A* has a goal to acquire certain information. This causes him to create a plan that involves asking *B* a question. *B* will hopefully possess the sought information. *A* then executes the plan, and thereby asks *B* the question. *B* will now receive the question and attempt to infer *A's* plan. In the plan there might be goals that *A* cannot achieve without assistance. *B* can accept some of these obstacles as his own goals and create a plan to achieve them. *B* will then execute his plan and thereby respond to *A's* question.

The above formulation contains many core concepts of the plan-based approaches: goals, actions, plan construction, and plan inference.

Important Concepts

There are several concepts underlying the planning approach to dialogue modelling. Therefore, a few words must be said about these concepts before we go into detail about the approach proper.

Beliefs, Knowledge, and Wants

Mental attitudes are omnipresent in plan-based dialogue modelling. The plan-based framework is often referred to as the Beliefs, desires, and intentions (BDI) framework (Allen, 1995). Hence, formalisations of such mental attitudes are of utmost importance. To a large extent, these formalisations are virtually identical, or draw heavily, on the work by Hintikka (1963, 1969).

Belief is typically represented with the logical operator B(A, P), which should be read 'agent *A* believes that *P*'. The most important property of this operator is that what one agent believes another agent to believe has no logical relation to what the former agents believes.

Perrault and Allen (1980) distinguish between three different kinds of knowledge. The first is used to represent that an agent *A* knows that something is true:

$$\mathrm{KNOW}(A, P) = P \wedge B(A, P)$$

The second is a KNOWIF operator that involves 'know' as in '*A* knows whether *P* is true' and is represented as:

$$\mathrm{KNOWIF}(A, P) = (P \wedge B(A, P)) \vee (\neg P \wedge B(A, \neg P))$$

The final use of know corresponds to utterances such as '*A* knows where the *D* is' and is represented by quantifying over the belief operator:

$$\mathrm{KNOWREF}(A, D) = \exists y((\text{the } x : D(x) = y) \wedge B(A, \text{the } x : D(x) = y))$$

The final mental attitude that has been made into an operator is 'want', and is also referred to as 'goals' and 'plans'. This sense of 'plans' differs slightly from the notion of 'plans' in the AI sense. Operationally it is represented as:

$$W(A, P) = A \text{ has the goal to achieve } P$$

Note that by this definition *A* actually intends to achieve *P*, not just that *A* would find *P* a desirable state of affairs.

Plans and Actions

Plan-based dialogue models rests heavily on the work on planning in artificial intelligence. The formulations of actions and plans are more or less directly taken from STRIPS (Fikes and Nilsson, 1993). This system models the world by means of logical propositions. The world's state is changed by means of actions, which can be viewed as parameterized procedures. An action is described by *preconditions*, *effects*, and a *body*. This is referred to as an action schema:

Preconditions: Conditions that must be true if the action's execution is to succeed.

Effects: Conditions that becomes true after the action is executed.

Body: A set of partially ordered goal states that must be achieved in performing the action.

An instance of an action is constructed from an action schema, instantiating its parameters and specifying a time. This predicate is only true if the action was or will be executed at time *t*:

$$\mathrm{ACT}(\mathrm{A})\langle \mathrm{t} \rangle$$

By combining the 'want' operator with this action definition, intentional actions can be formalised. Thus, if *ACT* is an intentional action by an agent *A* at a time *t*:

$$\mathrm{ACT}(\mathrm{A})\langle \mathrm{t} \rangle \rightarrow \mathrm{AW}(\mathrm{ACT}(\mathrm{A})\langle \mathrm{t} \rangle)$$

The time specifications of actions are often omitted, unless confusion may arise.

Given an initial world state and a goal state, a plan is a sequence of actions that transforms the world from the initial to the goal state.

Speech Acts

Speech acts, as proposed by Austin (1962) and later developed by Searle (1975), plays a central role in plan-based dialogue models. The idea is that utterances are also actions at different levels. Every utterance is said to have a locutionary, illocutionary, and perlocutionary force. For example, if a customer utters "Are there any flights to Boston?", the locutionary force is simply the words that constitute the utterance, the illocutionary force is a request, and the perlocutionary force is (perhaps) that the clerk provides the requested information.

Below, some common speech acts are accounted for in a STRIPS-like fashion. More detailed accounts can be found in Cohen and Perrault (1979) and Perrault and Allen (1980).

The speech act of informing the hearer of some proposition is simply called INFORM, and corresponds to Austin (1962) and Searle's (1975) *Assertive*:

INFORM(speaker(S), hearer(H), proposition(P))
precondition: KNOW(S, P) ∧ W(S, INFORM(S, P, H))
effect: KNOW(H, P)
body: B(H, W(S, KNOW(H, P)))

The definition of INFORM is based on Grice's (1957) idea that the speaker informs the hearer of something merely by causing the hearer to believe that the speaker wants them to know something.

The speech act INFORMIF is used to inform a hearer whether a given proposition is true or false.

> **INFORMIF**(speaker(S), hearer(H), proposition(P))
> *precondition*: KNOWIF(S, P) $\wedge$ W(S, INFORM(S, H, P))
> *effect*: KNOWIF(H, P)
> *body*: B(H, W(S, KNOWIF(H, P)))

A REQUEST is simply a speech act for requesting the hearer to perform some action.

> **REQUEST**(speaker(S), hearer(H), action(ACT))
> *constraint*: H is agent of ACT
> *precondition*: W(S, ACT(H))
> *effect*: W(H, ACT(H))
> *body*: B(H, W(S, ACT(H)))

Apart from these speech acts, surface speech acts are also identified (Perrault and Allen, 1980; Allen and Perrault, 1980). These correspond to the literal meanings of, on the one hand indicative mood utterances (S.INFORM), and on the other hand imperative utterances (S.REQUEST). These acts have no preconditions, and serve only to signal the immediate intention of the speaker. Thus, they are the starting point for all the inferences made by the hearer. S.REQUEST is formalized as:

> **S.REQUEST**(speaker(S), hearer(H), action(ACT))
> *effect*: B(H, W(S, ACT(H)))

Plan Inference

Plan inference is the process of one agent *A* attempting to infer another agent *B's* plan, based on observed actions performed by *B*. Usually, this process starts with an incomplete plan, containing only a single observed action or an expected goal. Below, some inference rules are presented. These are taken from Allen and Perrault (1980). The inference rules are applied according to heuristics specifying how likely it is that the new plan specification it produces is the actual plan.

In Allen and Perrault's notation

$$B(S, W(A, X)) =i\rightarrow B(S, W(A, Y))$$

means that if *S* believes *A's* plan contains *X*, then *S* may infer that *A's* plan also contains *Y*. These inference rules can be divided into three different types: those that concern actions, those that concern knowledge, and those that concern planning by others.

Rules Concerning Actions

Allen and Perrault identify four different inference rules concerning actions: Precondition-Action Rule, Body-Action Rule, Action-Effect Rule, and Want-Action Rule.

Precondition-Action Rule: If *P* is a precondition of an action *ACT*, and an agent *S* believes that another agent *A* wants to achieve *P*, then we can probably infer that *S* believes *A* wants *ACT* to be executed.

Formally, this would be expressed as:

$$B(S, W(A, P)) = i \rightarrow B(S, W(A, ACT))$$

We will leave out most of the other formalisations for brevity. The interested reader should consult Allen and Perrault (1980).

Body-Action Rule: If *B* is part of the body of *ACT*, and if *S* believes that *A* wants *B* to be executed, it is likely that *S* believes that *A* may want to execute *ACT*.

Action-Effect Rule: If *E* is an effect of an action *ACT*, and *S* believes *A* wants to execute *ACT*, then it is plausible that *S* believes that *A* wants the effect of that action.

Want-Action Rule: If *S* believes that *A* wants *N* to want some action *ACT* to be executed, then *S* may believe that *A* wants *ACT* to be executed.

Rules Concerning Knowledge

These rules regard how goals of acquiring knowledge relate to goals and actions that use that knowledge. Allen and Perrault (1980) identify four such rules: Know-positive Rule, Know-negative Rule, Know-value Rule, and Know-term Rule.

Know-positive Rule and Know-negative Rule: These rules consider the case where *S* believes that *A* wants to know whether a proposition *P* is true, from which we may infer that *S* believes that *A* wants to achieve a goal that requires *P* to be true or false.

Know-value Rule: This rule reflects that if *S* believes *A* wants to know if a proposition *P* is true, *S* may believe that *A* wants this in order to establish the identity of one of the terms in *P*.

Know-term Rule: This rule concerns goals of finding referents of descriptions.

Rules Concerning Planning by Others

Plan construction rules can be seen as the inverses of plan inference rules. Hence, they can be defined as a set of rules that describe possible constructions. Plan construction is, like plan inference, governed by a set of heuristics. The plan construction rules are: Action-Precondition Rule, Action-Body Rule, Effect-Action Rule, Know Rule, Nested-Planning Rule, and the Recognizing Nested-Planning Rule. Because of these rules resemblance to previously mentioned rule, they are not described in further detail.

Obstacle Detection

Once the speaker's plan has been inferred by the listener, he tries to identify obstacles in that plan in order to be helpful. An obstacle is simply any goal specification that is not initially true or achieved within the plan.

An Illustrative Example

To make thing a bit more concrete, we will illustrate a possible inference chain with an example[1].

Assume that a customer utters the indirect speech act "Can you give me a list of flights from Atlanta?" This would give rise to the following inference chain:

Rule	Step	Result
	0	.REQUEST(S, H, INFORMIF(H, S, CANDO(H, GIVE(H, S, LIST))))
AE	1	B(H, W(S, INFORMIF(H, S, CANDO(H, GIVE(H, S, LIST)))))
AE/EI	2	B(H, W(S, KNOWIF(H, S, CANDO(H, GIVE(H, S, LIST)))))
KP/EI	3	B(H, W(S, CANDO(H, GIVE(H, S, LIST))))
PA/EI	4	B(H, W(S, GIVE(H, S, LIST)))
BA	5	REQUEST(H, S, GIVE(H, S, LIST))

Starting with step 0, the hearer initially interprets the speaker's speech act literally as a question. In step 1, the Action-Effect (AE) inference rule, the rationale being that since the speaker asked for something it is plausible that he wants it. Step 2 also involves the AE rule. This time suggesting that if the speaker wants an INFORMIF, and KNOWIF is an effect of that act, the speaker probably also wants KNOWIF. In step 3 an inference is made on the basis that people rarely ask questions about thing they are not interested in. The inference rule used is the

[1] This example is taken from Jurafsky and Martin (2000).

Know-Positive Rule (KP). Hence, if the speaker asks whether something is true or not, he is likely to want it to be true. Using the Precondition-Action Rule (PA) in step 4, it is inferred that if the speaker wants a precondition for an action, it is likely that the speaker also wants the action. Finally, in step 5 the Body-Action Rule (BA) is applied, since from the definition of REQUEST we can infer that if the speaker wants someone to know that the speaker wants them to do something, then the speaker is probably issuing a REQUEST.

Plan Generation

Plan construction is the process of creating a sequence of actions that will transform a world from an initial state to a goal state. The approach to planning taken in Perrault and Allen (1980) is backward chaining, where one attempts to find an operator which, if executed, would achieve the goal. If its preconditions are satisfied in the initial world state, the plan is complete. If not, the planner attempts to satisfy the preconditions. Planning can also be done at different levels of detail.

The plan construction rules proposed in Perrault and Allen (1980) are basically inverses of the plan inference rules. Hence, they propose the Effect-Action Rule, Action-Precondition Rule, and the Action-Body Rule.

Improvements and developments

The plan-based approach to dialogue modelling as first proposed by Allen and Perrault (1980) has since been developed in order to deal with problems of the original model. Some of these developments are described below.

Context model

One problem with the original model proposed by Allen and Perrault's (1980) is that agents are assumed to have one of a small number of top-level goals. Further, this goal is taken to be deducible from a single utterance. Carberry (1990a) realized this and provided a model that infers an agent's plan incrementally as the dialogue progresses. The structure that is built up is referred to as a context model. The context model represents the system's beliefs about an agent's goals and partial plan for accomplishing these goals. Given a new utterance, her system hypothesizes a set of candidate focus actions on which the system believes the agent's attention might be focused at the given moment. Using heuristics, the candidate focused action that is most coherently related to the existing focus of attention in the context model is chosen. Finally, the context model is expanded to include it.

This approach solves the above mentioned problem. If the agent's top level goal is not deducible from the outset, the extended dialogue is used for this deduction.

Multiple hypotheses

Plan recognition is usually accomplished using various chaining algorithms. However, chaining procedures often lead to multiple hypotheses about an agent's plan. Therefore, efforts have been made to find techniques for narrowing the hypothesis space.

Perrault and Allen (1980) realized this problem and devised different heuristics to rank hypotheses. For instance, their approach downgraded hypothesis about an agent's partial plan if it was the result of one of several alternative inferences. The logic being that the speaker could not have intended the hearer which inference to choose.

One approach developed by Wu (1991) and subsequently implemented by van Beek and Cohen (1991) is to make the dialogue behaviour actively directed towards disambiguating agent's plans. The idea is to enable the system to identify which is the most effective query for disambiguating a given set of hypothesis. The system implemented by van Beek and Cohen (1991) critiqued the multiple hypotheses in order to identify faults in the inferred plans and entered into a clarification subdialogue when the choice of plan affected the system's response. The drawback with this approach is that the dialogue can become disrupted by unnecessary questions.

A natural approach is to use information about the domain when choosing between several hypotheses. Carberry (1990b) developed a strategy that was motivated by psychological studies on human inference and decision making. Her strategy could both defer unwarranted decisions about the relationship of a new action to the agent's overall plan and sanction rational default inferences.

Non-domain goals and plans

To a large extent, research in plan-based models of dialogue has been concerned with reasoning about domain actions and the goals to which they contribute. However, an agent usually has goals of other types as well. For instance, in a collaborative planning dialogue many utterances do not concern domain goals, but rather to how the participants should approach solving the problem. For the system to be an effective collaborative partner, it therefore needs to be able to recognize problem solving strategies (Elzer, 1995).

Similarly, expressions of doubt such as "Isn't the Ferrari faster than the Skoda?" must also be handled correctly. This expression can both be a request for verification or it can actually express doubt about some proposition (Carberry and Lambert, 1999). In the former case it presents no problem, but in the latter case an appropriate response must address the implied relationship between the queried proposition and the proposition that is implicitly being doubted.

Another type of plans is plans *about* plans, so called metaplans. Wilensky (1983) was the first to address the importance of metaplans. However, he pursued them in the area of story understanding. In the context of language understanding, Litman and Allen (1987) introduced the notion of discourse metaplans to capture how an agent might extend, continue, or modify the plan being pursued during the dialogue. They also developed a plan inference system in which represented and reasoned about domain plans and metaplans in a unified framework.

Discussion

It is to some extent difficult to assess the benefits of plan-based models of dialogue compared to for instance grammar based approaches because little formal evaluation has been carried out. Instead, it seems that isolated example dialogues or dialogue phenomena are selected and argued for. Also, little evaluation has been done within the planning community to establish exactly how much the introduction of context models and metaplans contribute to improved performance (cf. Carberry, 1996).

On a similar note, it would be interesting to study in what types of dialogue and domains plan-based models are superior to other models and in what situations they are an unnecessarily cumbersome way to carry on a dialogue. This kind of models are in their nature computer intensive and the dialogue can become halting as a result.

One objection against plan-based dialogue modelling concerns the lack of theoretical base. Even though the approach provides a solid computational model, it is not entirely clear how the mental constructs postulated in the model correlates to people's actual mental states. Disregarding the obvious objections by eliminative materialists (Churchland, 1981) and behaviourists alike, one would want the model to be rooted more deeply in psychological research. However, there seems to be similarities, at least at a surface level, with for instance Clark's (1996) ideas of action ladders and joint activities.

References

Allen, J. F. 1995. Natural language understanding. The Benjamin/Cummings Publishing Company, Inc.

Allen, J. F. and C. R. Perrault. 1980. Analyzing intention in utterances. Artificial Intelligence, 15(3):143–178.

Austin, J. L. 1962. How to do things with words. Harvard University Press.

Carberry, S. 1990a. Plan recognition in natural language dialogue. MIT Press.

Carberry, S. 1990b. Incorporating default inferences into plan recognition. In Proceedings of the 8th National Conference on Artificial Intelligence, pages 471-478.

Carberry, S. 1996. Plan recognition: achievements, problems, and prospects.

Carberry, S. and L. Lambert. 1999. A process model for recognizing communicative acts and modelling negotiation dialogue. Computational Linguistics, 25(1):1-53.

Chu-Carroll, J. and S. Carberry. 1998. Collaborative response generation in planning dialogues. Computational Linguistics, 24(3):355–400.

Churchland, P. M. 1981. Eliminative materialism and the propositional attitudes. Journal of Philosophy, 78:67–90.

Clark, H. H. 1996. Using language. Cambridge University Press.

Cohen, P. R. 1996. Dialogue modelling. In Ronald A. Cole, editor, Survey of the State of the Art in Human Language Technology. chapter 6.3, pages 233–240.

Cohen, P. R., J. L. Morgan, and M. E. Pollack, editors. 1990. Intentions in communication. MIT Press.

Cohen, P. R. and C. R. Perrault. 1979. Elements of a plan based theory of speech acts. Cognitive Science, 3(3):177–212.

Elzer, S. 1995. The role of user preferences and problem-solving knowledge in plan recognition for expert consultation systems. In Proceedings of the IJCAI Workshop on the Next Generation of Plan Recognition Systems, pages 37-41.

Fikes, R. E. and N. J. Nilsson. 1993. STRIPS: a retrospective. Artificial Intelligence, 59(1–2):227–232.

Gordon, D. and G. Lakoff. 1971/1975. Conversational postulates. In P. Cole and J. L. Morgan, editors, Syntax and semantics, vol. 3: speech acts. Academic Press, New York.

Grice, H. P. 1957/1994. Meaning. In R. M. Harnish, editor, Basic topics in the philosophy of language. Harvester Wheatsheaf, pages 21–29.

Grosz, B. J. and C. L. Sidner. 1986. Attention, intentions and the structure of discourse. Computational Linguistics, 12(3):175–204, July-September.

Hintikka, J. 1969a. Knowledge and belief. Cornell University Press, Ithaca, NY.

Hintikka, J. 1969b. Semantics for propositional attitudes. In J. W. Davis, D. J. Hockney, and W. K. Wilson, editors, Philosophical logic. Reidel, Dordrecht, Holland, pages 21–45.

Hirschberg, J. and J. Pierrehumbert. 1986. The intonational structuring of discourse. In Proceedings of ACL-86, pages 136–144.

Hirschberg, J. and D. J. Litman. 1993. Empirical studies on the disambiguation of cue phrases. Computational Linguistics, 19(3):501–530.
Jurafsky, D.. S. and J. H. Martin. 2000. Speech and language processing. Prentice-Hall, Inc.
Litman, D. J. 1985. Plan recognition and discourse analysis: an integrated approach for understanding dialogues. Ph.D. thesis, University of Rochester.
Litman, D. J. and J. F. Allen. 1987. A plan recognition model for subdialogues in conversation. Cognitive Science, 11:163-200.
McTear, M. F. 2002. Spoken dialogue technology: enabling the conversational user interface. ACM Computing Surveys (CSUR), 34(1):60–169, March.
Passonneau, R. J. and D. J. Litman. 1993. Intention-based segmentation: human reliability and correlation with linguistic cues. In Proceedings of ACL-93, pages 148–155, Columbus, Ohio. ACL.
Perrault, C. R. and J. F. Allen. 1980. A plan-based analysis of indirect speech acts. American Journal of Computational Linguistics, 6(3–4):167–182.
Pierrehumbert, J. and J. Hirschberg, 1990. The meaning of intonational contours in the interpretation of discourse, chapter 14. MIT Press.
Pollack, M. E., 1990. Plans as complex mental attitudes, chapter 5. MIT Press.
Searle, J. R. 1975. Indirect speech acts. In P. Cole and J. L. Morgan, editors, Syntax and semantics, vol. 3: speech acts. Academic Press, New York.
van Beek, P. and P. Cohen. 1991. Resolving plan ambiguity for cooperative response generation. In Proceedings of the 25th annual meeting of the Association for Computational Linguistics, pages 938-944.
Wilensky, R. 1983. Planning and understanding. Addison-Wesley.
Wu, D. 1991. Active acquisition of user models: implications for decision-theoretic dialog planning and plan recognition. User Models and User-adapted Interaction, 1(2):149-172