User Modeling and User-Adapted Interaction **11:** 31–48, 2001.

Techniques for Plan Recognition

SANDRA CARBERRY
Department of Computer Science, University of Delaware, Newark, Delaware 19716, U.S.A.

(Received: 2 December 1999; in final form: 27 July 2000)

Abstract. Knowing a user's plans and goals can significantly improve the effectiveness of an interactive system. However, recognizing such goals and the user's intended plan for achieving them is not an easy task. Although much research has dealt with representing the knowledge necessary for plan inference and developing strategies that hypothesize the user's evolving plans, a number of serious problems still impede the use of plan recognition in large-scale, real-world applications. This paper describes the various approaches that have been taken to plan inference, along with techniques for dealing with ambiguity, robustness, and representation of requisite domain knowledge, and discusses areas for further research.

Keywords: plan inference, goals, plans, intentions

1. Introduction

Suppose that someone asked you for the location of the Federal Express office and subsequently asked about the availability of delivery outside the country. You might reasonably infer that he or she wanted to quickly get an item to someone in another country and intended to do this using Federal Express for delivery. In doing so, you had inferred the goals of the other person and a portion of that person's plan for achieving those goals. This is often referred to as *plan recognition*.

Plan recognition has been used extensively in a wide variety of computer systems. Applications include language understanding and response generation (Allen and Perrault, 1980; Perrault and Allen, 1980; Litman and Allen, 1987; Carberry, 1990b; Sider and Burger, 1992; Smith et al., 1995; Haller and Shapiro, 1996), speech-to-speech translation (Alexandersson, 1995), interfaces for computer-aided design (Goodman and Litman, 1992), UNIX help systems (Mayfield, 1992), collaborative problem-solving (Lesh et al., 1999), and automated descriptions of image sequences (Retz-Schmidt, 1991). The article in this special issue by Zukerman and Litman (2001) discusses the role of plan recognition in natural language processing. Two well-known, large-scale projects in which plan recognition plays a major role are VERBMOBIL (Alexandersson, 1995; Alexandersson et al., 1997) (a speech-to-speech translation system whose first application domain was appointment scheduling) and TRAINS (Allen et al., 1996; Ferguson et al., 1996) (a spoken dialogue system for interactive route planning) that has evolved into TRIPS (Ferguson and Allen, 1998) (a collaborative system for solving logistics problems).

One might ask whether plan recognition is performed by humans or is just an artifact of use in computer systems. Research by Schmidt et al. (1978) provides evidence that humans do in fact infer the plans of other agents. They performed experiments in which human subjects were presented with sequences of actions by another agent and were asked to summarize their observations and predict subsequent actions. The subjects' summarizations generally included one or more goals attributed to the observed agent, with the agent's actions explained in terms of how they contributed to achieving that goal. Moreover, the summaries indicated that the subjects had used the inferred plan to predict the agent's subsequent actions. This psychological research, along with an experiment by Cohen et al. (1981), support the contention that humans engage in plan recognition and use their hypotheses in subsequent reasoning.

Plan recognition can be characterized according to the role of the agent whose plan is being inferred. *Keyhole recognition* (Cohen et al., 1981) is the recognition of an agent's plan through unobtrusive observation – that is, the agent does not attempt to impact the recognition process. This kind of recognition is performed by help systems that generate unsolicited advice (Shrager and Finn, 1982; Finin, 1983). But might an agent attempt to perform actions that will aid or hinder the recognition of his plan? The former situation has been actively studied in language understanding, where the system (in order to respond appropriately to a speaker's utterances) must recognize the plan that the speaker intends to convey; this is referred to as *intended recognition* (Cohen et al., 1981). The latter situation is representative of adversarial settings, such as warfare, where an agent might actively attempt to thwart recognition of his plan (Pollack, 1986b; Azarewicz et al., 1986). However, little published research has examined situations where deception must be taken into account.

Although early work on plan recognition offered much promise and the contribution of plan recognition to robust adaptive systems has been widely recognized, a number of serious problems have hampered the use of plan recognition in realistic large-scale applications. This paper discusses the general plan inference paradigm, various approaches to plan recognition, techniques developed to address issues of plan ambiguity, robustness, and efficiency, and prospects for the future.

2. The Basic Plan Inference Process

Models of plan inference start with a set of goals that an agent might be expected to pursue in the domain and an observed action by the agent. The plan inference system then has the task of inferring the agent's goal and determining how the observed action contributes to that goal. To accomplish this, the system traditionally is provided with a set of actions that the agent might execute in the domain and a set of *recipes* that encode how an agent might go about performing these actions. These recipes constitute a *plan library* and include each action's preconditions, the subgoals that comprise performing the action, and the effects or goals of executing the action.

The reasoning performed by the system uses this domain-dependent knowledge but is itself largely domain-independent.

To infer the agent's goal from the observed action, the plan inference system constructs a sequence of goals and actions that connect the observed action to one of the possible domain goals. This is accomplished by chaining from actions to goals achieved by the action, from these goals to other actions for which the goal is a precondition or subgoal, from these actions to their goals, etc. This process is depicted in Figure 1 where the A_j are actions and the G_i are goals. The inference path A_1,G_a,A_4,G_b,A_5,G_x provides an explanation for the observed action A_1 in terms of how it contributes to the goal G_x.

Perhaps the most well-known of the early plan recognition systems are those of Robert Wilensky (Wilensky, 1978, 1983) and James Allen (Allen and Perrault, 1980; Perrault and Allen, 1980). Wilensky's system was developed for story understanding, and the inference path served as an explanation for the occurrence of action A_1 in the story. In Allen's system,the plan library specified how to perform not only domain actions but also speech acts such as *Inform* and *Request*. The action A_1 was the primitive speech act associated with an utterance and the inference path served to explain the intentions underlying a speaker's utterance. Allen's system was developed to account for extra helpful information included in natural language responses to queries and also to interpret indirect speech acts. For example,by inferring the

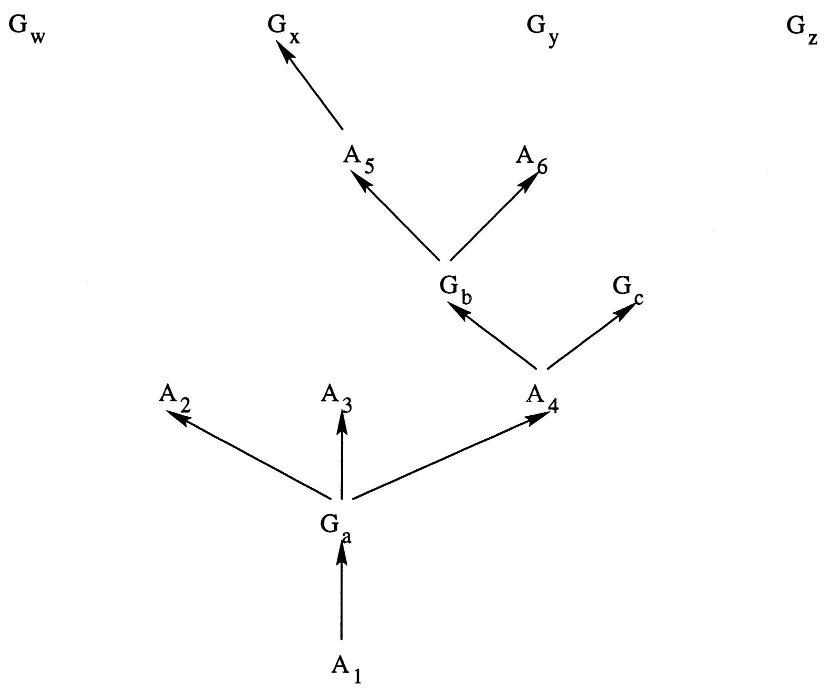

Figure 1. Chaining for Plan Inference.

plan motivating the speaker's utterance, the system could extract obstacles (knowledge that the speaker was missing) from the plan and provide augmented responses that removed these identified obstacles.

Although subsequent work has built heavily on these early systems, they had many limitations. Wilensky's system selected the shortest inference path connecting an observed action to an expected goal, without considering the current focus of attention in the story. Allen's system assumed that an agent had one of a small number of top-level goals which could be deduced from a single utterance. Carberry (1983, 1990b) built on Allen's work by providing a model of plan recognition that inferred an agent's plan incrementally as the dialogue progressed. She used a tree structure called a *context model* to represent the system's current beliefs about an agent's goals and partial plan for accomplishing these goals. From a new utterance, her system hypothesized a set of actions (called *candidate focused actions*) on which the system believed the speaker's attention might now be focused, used focusing heuristics to select the candidate focused action that was most coherently related to the existing focus of attention in the context model, and then expanded the context model to include it. Her approach enabled plan recognition to handle extended dialogues where the agent's top-level goal could not be deduced at the outset.

3. Narrowing the Hypotheses

In most realistic situations, chaining produces multiple hypotheses about an agent's plan, some of which appear more plausible than others. Thus plan recognition research has been forced to explore techniques that narrow the space of viable hypotheses.

Wu (1991) contended that plan recognition systems should not just passively work with the information provided by the user but should instead request specific information that would help the system disambiguate the user's plan sufficiently for the purposes of the current interaction. He proposed a decision-theoretic approach that took into account the utility of the different hypotheses and of potential system queries to identify the most effective query. Although Wu did not implement his proposal, the basic concept was adopted by van Beek and Cohen (1991). They developed a system that critiqued the multiple hypotheses to identify faults in the inferred plans and entered into a clarification subdialogue when the choice of plan affected the system's response – i.e., when the detected faults differed. Criteria for queries in the clarification subdialogue included coherence (resulting in a top-down approach), minimization of the dialogue, and minimization of the length of each question.

However, unnecessary questions can disrupt the dialogue and appear unintelligent. Heuristics that cull the intuitively less plausible hypotheses have played a prominent role in most plan recognition systems. Many of these heuristics have been based on rationality or coherency. For example, Allen and Perrault's model of intended recognition (Perrault and Allen, 1980) downgraded a hypothesis about an agent's partial plan if the hypothesis was the result of one of several alternative inferences, since the speaker could not have intended the hearer to recognize which inference to choose. For example, if an agent wanted to know if a specified proposition was true, then the hearer might infer that the agent has a plan in which the proposition must hold or alternatively a plan in which the proposition must be false; in either case, the agent is checking to see whether the required conditions on the proposition are satisfied. Litman (1986) preferred plans that represented the most coherent discourse moves. Raskutti and Zukerman (1991) devised a probabilistic approach for assessing the likelihood of competing hypotheses during intended recognition. Their system used domain-independent heuristics to distribute probabilities among the possible domain plans inferred solely from the user's new input[1] and to estimate the probability of different relations between these plans and the potential plans inferred from the preceding dialogue. They then used Bayes rule to compute the probability of each new competing hypothesis based on the above probabilities and the probability of each previous hypothesis. These probabilities were subsequently revised based on an information-theoretic measure that estimated the extent to which a hypothesis was specified well-enough for the intended actions to be taken; the underlying motivation was that a speaker would provide sufficient information for the listener to carry out her role in the interaction.

Domain information can also contribute to the intuitive plausibility of alternative hypotheses about the user's plan. For example, the speaker who asks about the location of the Federal Express office most likely wants to send a package, but other alternatives exist, such as applying for a job or even robbing the office. Carberry (1990a) devised a strategy, motivated by psychological studies of human inference and decision-making, that could sanction rational default inferences about the user's plan but defer unwarranted conclusions until further evidence was accumulated. It was implemented using Dempster–Shafer theory (Shafer, 1976) to represent the support that individual pieces of evidence give to alternative hypotheses and to compute the combined support offered by multiple pieces of evidence.

The above efforts are concerned with methodologies that are applicable to all users. However, knowledge about the individual user or user group can impact the plausibility of competing hypotheses. Gertner (Gertner and Weber, 1996; Gertner, 1997) developed a system for inferring physician's plans from their actions during emergency center trauma care. Since the physician was presumed to be an expert in the domain and thus to be more likely to pursue suboptimal plans rather than incorrect ones, her system included what she termed a *bias* toward hypotheses that included fewer goals deemed incorrect by the system's expert reasoning. One can envision how such biases might be devised for other user groups. For example, a tutoring system that is inferring a fifth-grader's plan for solving long-division problems from his solutions to several sample problems might have

[1] Raskutti and Zukerman take into account the number of alternative propositions that might be inferred and the modality of the user's utterance.

a bias toward hypotheses about the student's plan that do not involve simple subtraction errors, since students at this level presumably have mastered subtraction.

Plan recognition systems might perform even better if they could be adapted to the individual agent whose plan is being inferred. Ardisonno and Sestero (1996) used beliefs about the user, obtained from stereotypes and heuristic analysis of the user's previous actions, to help identify implausible plans. Bauer (1994, 1995, 1996) was concerned with situations (such as handling email) in which users typically repeated the same pattern of actions when presented with a previously encountered goal. He viewed these repeated patterns as preferences, and used Dempster-Shafer theory to represent preferences deduced from previous activity (Bauer, 1995, 1996) and to develop a probabilistic mechanism (Bauer, 1994) for choosing among competing hypotheses given such user-specific evidence. Bauer's work was specifically designed for domains where repeated patterns of actions can be observed; it will be interesting to explore how a mechanism such as Bauer's might be extended to preferences that do not represent repeated patterns of actions, such as the preferences recognized by Elzer (Elzer et al., 1994; Carberry et al., 1999) from a collaborative planning dialogue.

Lesh (1997) furthered this user-tailored approach by investigating how observation of an agent's behavior might be used for more general adaptation of the plan recognition process. Lesh distinguished between potential goals (goals that might represent the user's primary intention and which the system might want to recognize) and background goals (spurious goals that are not related to the user's main goal). Adding or removing elements in the set of potential goals increased the efficiency of the recognition system by adapting it so that it considered only goals that the particular user tended to pursue. Adding or removing elements in the set of background goals adapted the recognition system to the particular user by enabling it to filter out spurious actions that the user tended to perform even though they were irrelevant to the task. Lesh's methodology employed hill climbing to select from among the set of adaptations that were appropriate based on an agent's observed behavior over time.

The above efforts have considered a wide variety of mechanisms for narrowing the space of viable hypotheses about an agent's plan. The heuristics based on rationality on the part of the planning agent are probably appropriate for any system. But what about the large number of other heuristics and methods that have been explored? Unfortunately, as with much of the research on plan recognition, there has been little evaluation of the individual heuristics to determine their effectiveness. Bauer demonstrated that his adaptive mechanism improved performance of plan recognition in the domain of electronic mail. Lesh's experiments (Lesh, 1997) show that an adaptive system can improve performance. However, his experimental data was provided by other planning systems or simulations rather than observations of humans in realistic scenarios. Thus although the mechanisms employed in the various systems all appear plausible, evaluation studies are needed to identify their impact on plan recognition in large scale real-world situations.

4. Other Models of Plan Recognition

Most plan recognition systems have followed the basic model outlined in Section 2. However, other formalisms have been proposed for dealing with the uncertainty inherent in plan inference, most notably formal argumentation models and approaches based on probabilistic reasoning.

Several researchers have captured plan recognition in a formal model of argumentation (Konolige and Pollack, 1989) or abduction (Appelt and Pollack, 1992; Waern, 1994). Appelt and Pollock (1992) used weighted abduction in which weights were assigned to the premises of each rule. The cost of proving a conclusion C was the sum of the costs of proving the premises in a rule whose consequent was C. The cost associated with a premise depended on whether it was true, proven from other rules, or assumed; in the latter case, the premise's cost was affected by the weight assigned to the premise. The best hypothesis about the agent's plan was the one given by the lowest cost proof. As noted by Appelt and Pollock, weighted abduction captures domain-dependent knowledge about the likelihood that a premise is true.

Approaches based on formal models of probability have gained increased prominence during the past decade. Section 3 discusses Bauer's use of Dempster–Shafer theory for rating hypotheses about an agent's plan. One of the arguments given by proponents of Dempster–Shafer as a method for reasoning under uncertainty is that it distinguishes lack of evidence for a proposition from evidence against the proposition. Other researchers favor Bayesian reasoning, generally captured in Bayesian belief networks (Pearl, 1988). Although Raskutti and Zukerman used Bayes rule to compute the final probabilities of competing hypotheses (see Section 3), their system implemented a heuristic approach.

Charniak and Goldman (1991, 1993) constructed the first Bayesian plan inference system. Their system used *marker passing* (Charniak, 1986) (a form of spreading activation in a network of nodes and links) to identify potential explanations for observed actions and to identify nodes for insertion into a Bayesian belief network. In a Bayesian belief network, nodes represent random variables; arcs between nodes represent causal dependencies, captured by conditional probability distributions that for each value of the parent node give the probability of each of the various possible values of the child node. When used for plan inference, the random variables are propositions, the root nodes represent hypotheses about an agent's plan, and the probability assigned to a node represents the likelihood of a proposition given the evidence. Bayes rule is used to compute the probability of each proposition from the causal evidence provided by its parents and the diagnostic evidence provided by its children. As new evidence is added to the network, the probabilities at each node are recomputed, thereby propagating the evidence through the nodes.

Charniak and Goldman applied their plan recognition system to the problem of understanding a character's actions in a story. As with any system based on Bayesian networks, it required a large number of prior and conditional probabilities. Such systems are most appropriate for domains where these probabilities can be reliably estimated, and where the causal influence among nodes can be reliably determined. It is unclear that story understanding has these attributes. In addition, it appears that Charniak and Goldman's system is not sensitive to the order in which actions are observed, something that should affect plan recognition in longer stories.

Albrecht et al. (1997, 1998) constructed a plan inference system based on *Dynamic Belief Networks*. Dynamic Belief Networks (Dean and Wellman, 1991) capture the influence of temporal aspects by using multiple nodes to represent the status of a variable at different instances of time. Their plan inference system was used to infer an agent's plan during an adventure game. The joint probability distributions were based on data collected from actual games. With the ability to assign reliable probabilities based on collected observations, this was a good domain in which to investigate probabilistic plan inference. However, the designers were still faced with the problem of identifying the appropriate network structure. They investigated four networks of different complexity, and provided an extensive evaluation of the impact of the different networks on the quality of plan recognition.

The results of Albrecht et al. suggest that Dynamic Belief Networks offer a promising approach to keyhole plan recognition in situations where sufficient training data can be collected and the causal structure of the network clearly identified. For other applications of Dynamic Belief Networks in plan recognition, see Pynadath and Wellman (1995) and Forbes et al. (1995).

5. Extending Plan Recognition to Non-Domain Plans

Most plan inference research has been concerned with domain actions and the domain goals to which they contribute. However, in many contexts a system will be severely hampered unless it can reason about a wider variety of goal types. For example, Elzer (1995) argues that many utterances in a collaborative planning dialogue do not refer to specific domain goals but rather to how the agents should approach solving the problem. Unless a system can recognize problem-solving strategies and their relationship to the domain plan under construction, the system will be unable to achieve its full potential as a collaborative partner. Similarly, a natural language dialogue system must be able to recognize communicative goals such as expressing doubt at a proposition conveyed by another agent. For example, a Surface-negative question of the form *"Isn't PROP"*, where *PROP* is a proposition, can be simply a request for verification or can be an expression of doubt at some other proposition (Lambert and Carberry, 1992; Carberry and Lambert, 1999). If *PROP* is in fact true and the utterance is merely seeking verification, then it is sufficient to affirm the proposition's truth. However, if the utterance is intended to express doubt, then an appropriate response must address the implied relationship between the queried proposition and the proposition that is implicitly being doubted.

Wilensky (1981, 1983) was the first to address the importance of non-domain goals in plan inference. He investigated the recognition of metaplans (plans *about* plans), such as a metaplan for resolving conflict between two competing goals. Since often a character's actions in a story make sense only in terms of a metagoal and the character's plan for pursuing it, recognition of metaplans is essential for story understanding. In the area of language understanding, Litman and Allen (1987) introduced the notion of discourse metaplans to capture how a speaker might extend, continue, or modify the plan being pursued during a dialogue. As with Wilensky's metaplans, Litman and Allen's metaplans might be termed problem-solving plans since they reflected an agent's plan construction process. Litman and Allen developed a plan inference system that represented and reasoned about discourse metaplans and domain plans in a unified framework. This allowed them to recognize goals to clarify or correct an existing plan. Ramshaw (1989, 1994) expanded on Litman's work and developed a system for recognizing a rich set of problem-solving goals.

However, recognizing domain, problem-solving, and communicative goals during an extended interaction is not a simple task. As shown by Lambert and Carberry (1991), each kind of goal and plan must be recognized incrementally. For example, consider the following set of utterances:

The City of <xxx> is considering filing for bankruptcy.
One of your mutual funds owns <xxx> bonds.

Although neither utterance by itself constitutes a warning,a plan inference system must be able to recognize the *warning* from the two utterances together. Thus Lambert and Carberry developed a plan inference system that recognized each kind of plan (domain, problem-solving, and communicative) incrementally as the dialogue progressed. Focusing heuristics appropriate to each type of plan were used to capture expectations about potential shifts in attention. Rosé et al. present a system for handling more complex focus shifts (Rosé et al., 1995).

Although research has shown the necessity for recognizing more than just domain goals, relatively little effort has been devoted to incorporating such reasoning into plan recognition systems. Thus it remains a fruitful area for further research.

6. Robustness

Plan recognition systems are plagued by what might be termed *noise* in the data; this noise hampers the system's ability to accurately identify the agent's plan. Recall that plan inference systems take three inputs: (1) a plan library capturing the system's knowledge about potential goals and means of accomplishing them; (2) input from the agent whose plan is being recognized, and (3) whatever partial plan has already been inferred for the agent. Thus far we have made two assumptions: (1) the new input can be associated with one or more recipes in the plan library, and (2) the action/goal captured by this recipe and the partial plan already inferred for the agent can be meshed together in a manner that reflects the agent's actual intentions.

But disruptions or deficiencies in each of the three inputs can cause the system to infer an incorrect plan or fail to infer any plan at all. In this section, we examine how this can occur and research that has attempted to address issues of robustness.

In addition to extraneous actions that have nothing to do with the agent's intentions in the domain (Albrecht et al., 1998), noise in the agent's input to the system can take the form of actions that are intended to advance the agent's goals but which are inappropriate (Pollack, 1987). Most plan inference systems assume that the user's knowledge is a subset of the system's and fail to account for the presence of misconceptions about how to achieve domain goals. Pollack (1986a) was the first to explicitly ascribe beliefs to an agent during plan inference. Although her system could ascribe the system's beliefs about recipes for achieving goals, it could also ascribe principled variations of these recipes. This allowed Pollack to infer ill-formed plans and thereby account for queries about actions that were inappropriate to a correct means of achieving a goal. Note that Pollack's mechanism differed from the inclusion of *buggy plans* (Brown and Burton, 1978), since the incorrect plans were not encoded in the system's knowledge base but were derived by applying principled mechanisms for hypothesizing variations in the system's recipes. Thus it theoretically is a much more powerful approach to plan recognition in the face of potential misconceptions. However, Pollack only experimented with very simple variations, such as the omission of a constraint; extending her work to more complex variations requires extensive research. Calistri-Yeh (1991) provided a classification of plan-based misconceptions that encompassed most of the categories identified by other researchers (Pollack, 1986a; Quilici et al., 1988; van Beek, 1987). He presented a plan inference system that used estimated probabilities to guide the search for an agent's intentions. Formulas attached to each class of misconception were used to estimate the probability of a particular kind of misconception based on the plan under consideration. Application of his system to several corpora supported this approach.

Besides misconceptions reflected in the user's input, noise can appear as an erroneous hypothesis about an agent's partially constructed plan. Since plan inference systems attempt to mesh together the action derived from the agent's input with the system's current hypothesis about the agent's plan, errors in the existing hypothesis will impede plan recognition. Eller and Carberry (1992) were the first to consider how the system's inference mechanisms might lead to incorrect beliefs about an agent's plan and thereby affect subsequent plan recognition. They proposed an approach to dealing with ill-formedness in which meta-rules relaxed the plan inference process and enabled the consideration of less well-formed hypotheses. The metarules could not only propose hypotheses reflecting possible misconceptions on the part of the agent but could also propose principled revisions of the system's existing beliefs about the agent's partial plan. However, as with Pollack's work, their implementation only suggested very simple variations.

In addition, plan recognition is hampered if the system's plan library does not capture all means of achieving a goal. Unfortunately, it is unrealistic to expect that the plan library will always be complete. Thus systems must also be able to reason in a principled manner about possible novel correct plans (correct plans that are not captured in the plan library) that an agent is pursuing. This is a very difficult problem and has received little attention in plan inference research. Cohen et al. (Cohen et al., 1991; Spencer et al., 1994) present a mechanism for updating the system's plan library with novel recipes, but they do not address the inference of novel plans.

Although most plan inference systems acknowledge the need to address the above problems, few actually do so. This is largely due to the problem of controlling inferencing once misconceptions, erroneous system beliefs, and novel plans must be considered. The problem of revising the system's beliefs may be alleviated with the advent of probabilistic approaches to plan recognition, since they construct multiple different hypotheses with associated probabilities. However,such systems still need to incorporate mechanisms for dealing with user misconceptions and novel plans and for revising system beliefs to account for misconceptions or novel plans that originally went undetected.

In addition to noise in the input, robustness is affected if the agent is deliberately attempting to thwart the plan inference process. Although Azarewicz et al. (1986; 1989) investigated plan recognition in an adversarial domain, they expanded their system's knowledge base to encode plans that they expected an adversary might pursue and did not propose principled mechanisms for hypothesizing how an agent might attempt to conceal his actual plan with misleading actions.

7. Acquisition and Representation of Knowledge

As discussed earlier, plan recognition systems require a knowledge base that encodes the system's beliefs about how goals can be achieved. The earliest plan libraries encoded recipes as collections of preconditions, constraints, subgoals, and effects. Kautz and Allen (1986) differentiated between *specialization* and *decomposition* recipes, where specialization captured alternative ways of performing a more general action and decomposition specified the subgoals that comprised an action. For example,the actions $Fly(X)$ and $Drive(X)$ might be represented in the knowledge base as specializations of the action $Travel\text{-}to(X)$.The plan hierarchy would capture subgoals that are part of all specializations of an action in the decomposition of that action,while the decomposition of a specialization of an action would capture subgoals that are specific to that specialization. More recently, researchers have developed representational systems based on description logic (Weida and Litman, 1992a, 1992b; Di Eugenio and Webber, 1992; Weida, 1995; Di Eugenio, 1995). For example, Weida and Litman (Weida and Litman, 1992a; Weida, 1995) developed an automatic classification system that incorporated temporal constraints. Their system could organize an initial knowledge base based on subsumption, could appropriately update the taxonomy with new or modified plans, and could efficiently retrieve plans from the taxonomy based on specified patterns. Such representation and retrieval systems are essential for efficient plan inference.

Unfortunately, constructing a plan library requires much human effort. Recently, attention has been given to how machine learning might alleviate the problem of building the library. As noted in Section 3, Lesh proposed that the plan library be adapted to the individual user. Lesh and Etzioni (1996) presented a method for automatically constructing the plan library using goal and plan biases; nonetheless, the basic knowledge must still be encoded beforehand. Mooney (1990) added plan schemata to a plan library; however, the plan must first be recognized by plan inference techniques, and the insertion into the plan library only provided for more efficient recognition if the plan was encountered again.

Bauer (1998) presented a strategy for learning plan decompositions from a training set of goals and associated action sequences. Since often there are alternative ways of achieving a goal, he used a similarity metric to identify which action sequences represented the same basic plan. Given such a similar set of action sequences, he then constructed a plan decomposition that eliminated unnecessary actions, temporal orderings, and structural relationships, and that replaced each step that appeared in every action sequence with a more general action that subsumed the individual steps. Although Bauer's approach can construct decompositions even without domain knowledge such as an object taxonomy, the plans become very general and unrestricted. And in many cases, developing the appropriate object taxonomy is closely related to constructing the plan library. Although a very nice step forward, this work seems most appropriate for domains where data can be easily collected and where goals are achieved by a small set of alternatives.

In the case of Bayesian belief nets for plan inference, the structure of the network has been determined a priori by the system designers, and in many cases involves simplifying assumptions to reduce complexity. Research on learning the structure of a belief net is promising, but it is unclear how applicable this might be for large-scale plan recognition with complex influences among the actions and how extensive a training set will be required. The article in this special issue by Zukerman and Albrecht (2001) discusses recent research on predictive statistical techniques, and the article by Webb et al. (2001) discusses the role of machine learning in user modeling.

8. Other Challenges for the Future

Although formal plan inference techniques have been used in many applications, they have received relatively little attention in intelligent tutoring systems. As noted by Greer and Koehn (1995), the robustness problems discussed in Section 6 are exacerbated when inferring student problem-solving plans in tutoring systems, since the student is by definition incompetent in the domain and such users tend to have very novel ways of erring. Thus this domain is an excellent one for investigating strategies for increasing the robustness of plan inference. The article in this special issue by Kay (2001) discusses user modeling in intelligent tutoring systems.

Recent research has begun to address the problem of inferring plans involving several agents (Castelfranchi and Falcone, 1995; Lochbaum, 1998; Devaney and Ram, 1998). Although many of the basic issues remain the same, multi-agent plan recognition must identify which agents are contributing to a particular plan and how their activities are interwoven. Moreover, in the case of collaborative planning dialogues, more complex focusing heuristics must be devised to capture the focus of attention of the individual agents and allow their utterances to be correctly interpreted.

Plan inference research has concentrated on demonstrating how new techniques work on a small set of problems in a limited domain. One of the major problems is scaling plan recognition techniques to larger domains with thousands of goals and means of achieving them. Lesh and Etzioni (1996) proposed that version spaces be used to represent the space of possible hypotheses for efficient processing in large-scale plan recognition, and they presented a plan recognition algorithm based on version spaces that worked well for a certain class of goals. However, their system was only applied to simulated test examples. Further work is needed on developing scalable recognition algorithms that can address real-world plan recognition problems.

9. Summary

Research suggests that humans perform plan inference and that plan inference contributes to much of the intelligent processing done by humans. The contributions of plan inference to intelligent performance by computers has been demonstrated in a wide variety of applications. Unfortunately, a number of serious problems still impede the use of plan recognition in large-scale real-world situations. The most serious are (1) system robustness in the face of noise in the input; (2) effective discrimination among competing hypotheses, and (3) recognition algorithms that scale up to large domains. With the advent of large corpora of data, it should soon be possible to evaluate plan recognition techniques in real-world scenarios with real-world data.

Acknowledgements

I would like to thank Judy Kay, Diane Litman, and Ingrid Zukerman for their helpful comments and suggestions on an earlier draft of this paper.

References

Albrecht, D., Zukerman, I. and Nicholson, A.: 1998, Bayesian models for Keyhole Plan Recognition in an adventure game. *User Modeling and User-Adapted Interaction* **8**(1–2), 5–47.

Albrecht, D., Zukerman, I., Nicholson, A. and Bud, A.: 1997, Towards a Bayesian model for Keyhole Plan recognition in large domains. In: *Proceedings of the Sixth International Conference on User Modeling*. Chia Laguna, Sardinia, Italy, pp. 365–376.

Alexandersson, J.: 1995, Plan recognition in VERBMOBIL. In: *Proceedings of the IJCAI-95 Workshop on the Next Generation of Plan Recognition Systems*. Montreal, Canada, pp. 2–7.

Alexandersson, J., Reithinger, N. and Maier, E.: 1997, Insights into the Dialogue Processing of VERBMOBIL. In: *Proceedings of the 5th Applied Natural Language Processing Conference*. pp. 33–40.

Allen, J., Miller, B., Ringger, T. and Sikorski, T.: 1996, A Robust system for natural spoken dialogue. In: *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics*. Santa Cruz, California, pp. 62–70.

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

Appelt, D. E. and Pollack, M. E.: 1991, Weighted abduction for plan ascription. *User Modeling and User-Adapted Interaction* **2**(1–2), 1–25.

Ardisonno, L. and Sestero, D.: 1996, Using dynamic user models in the recognition of the plans of the user. *User Modeling and User-Adapted Interaction* **5**(2), 157–190.

Azarewicz, J., Fala, G., Fink, R. and Heithecker, C.: 1986, Plan recognition for airborne tactical decision making. In: *Proceedings of the Fifth National Conference on Artificial Intelligence*. Philadelphia, Pennsylvania, pp. 805–811.

Azarewicz, J., Fala, G. and Heithecker, C.: 1989, Template-based multi-agent plan recognition for tactical situation assessment. In: *Proceedings of the Fifth IEEE Conference on Artificial Intelligence Applications*. Miami, Florida.

Bauer, M.: 1994, Quantitative modeling of user preferences for plan recognition. In: *Proceedings of the Fourth International Conference on User Modeling*. Hyannis, Massachusetts, pp. 73–78.

Bauer, M.: 1996, A Demster–Shafer approach to modeling agent preferences for plan recognition. *User Modeling and User-Adapted Interaction* **5**(3–4), pp. 317–348.

Bauer, M.: 1996, Acquisition of user preferences for plan recognition. In: *Proceedings of the Fifth International Conference on User Modeling*. Kailua-Kona, Hawaii, pp. 105–112.

Bauer, M.: 1998, Acquisition of abstract plan descriptions for plan recognition. In: *Proceedings of the Fifteenth National Conference on Artificial Intelligence*. Madison, Wisconsin, pp. 936–941.

Brown, J. and Burton, R.: 1978, Diagnostic models for procedural bugs in basic mathematical skills'. *Cognitive Science* **2**(2), 155–192.

Calistri-Yeh, R. J.: 1991, Utilizing user models to handle ambiguity and misconceptions in robust plan recognition. *User Modeling and User-Adapted Interaction* **1**(4), 289–322.

Carberry, S.: 1983, Tracking user goals in an information-seeking environment. In: *Proceedings of The National Conference on Artificial Intelligence*. Washington, D.C., pp. 59–63.

Carberry, S.: 1988, Modeling the user's plans and goals. *Computational Linguistics* **14**(3), 23–37.

Carberry, S.: 1990a, Incorporating default inferences into plan recognition. In: *Proceedings of the Eighth National Conference on Artificial Intelligence*. Boston, Massachusetts, pp. 471–478.

Carberry, S.: 1990b, *Plan Recognition in Natural Language Dialogue*, ACL-MIT Press Series on Natural Language Processing. Cambridge, Massachusetts: MIT Press.

Carberry, S., Chu-Carroll, J. and Elzer, S.: 1999, Constructing and utilizing a model of user preferences in collaborative consultation dialogues. *Computational Intelligence* **15**(3), 185–217

Carberry, S. and Lambert, L.: 1999, A process model for recognizing communicative acts and modeling negotiation subdialogues. *Computational Linguistics* **25**(1), 1–53.

Castelfranchi, C. and Falcone, R.: 1995, From single-agent to multi-agent: challenges for plan recognition systems. In: *Proceedings of the IJCAI-95 Workshop on The Next Generation of Plan Recognition Systems*. Montreal, Canada, pp. 24–32.

Charniak, E.: 1986, A neat theory of marker passing. *Proceedings of the Fifth National Conference on Artificial Intelligence*, Philadelphia, Pennsylvania, pp. 584–588.

Charniak, E. and Goldman, R.: 1991, A probabilistic model of plan recognition. In: *Proceedings of the Ninth National Conference on Artificial Intelligence*. Anaheim, California, pp. 160–165.

Charniak, E. and Goldman, R.: 1993, A Bayesian model of plan recognition. *Artificial Intelligence Journal* **64**, 53–79.

Cohen, P. R., Perrault, C. R., and Allen, J. F.: 1981, Beyond question answering. In: W. Lehnert and M. Ringle (eds.): *Strategies for Natural Language Processing*, pp. 245–274.

Cohen, R., Song, F., Spencer, B. and van Beek, P.: 1991, Exploiting temporal and novel information from the user in plan recognition. *User Modeling and User-Adapted Interaction* **1**(2), 125–148.

Dean, T. and Wellman, T.: 1991, *Planning and Control*. San Mateo, California: Morgan Kaufmann.

Devaney, M. and Ram, A.: 1998, Needles in a haystack: plan recognition in large spatial domains involving multiple agents. In: *Proceedings of the Fifteenth National Conference on Artificial Intelligence*. Madison, Wisconsin, pp. 942–947.

Di Eugenio, B. and Webber, B.: 1992, Plan recognition in understanding instructions. In: *Proceedings of the First International Conference on Artificial Intelligence Planning Systems*. pp. 52–61.

Di Eugenio, B.: 1995, Plan recognition and natural language understanding. In: *Proceedings of the IJCAI-95 Workshop on the Next Generation of Plan Recognition Systems*. Montreal, Canada, pp. 42–47.

Eller, R. M. and Carberry, S.: 1992, A meta-rule approach to flexible plan recognition in Dialogue'. *User Modeling and User Adapted Interaction* **2**(1–2), 27–54.

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*. Montreal, Canada, pp. 37–41.

Elzer, S., Chu-Carroll, J. and Carberry, S.: 1994, Recognizing and utilizing user preferences in collaborative consultation dialogues. In: *Proceedings of the Fourth International Conference on User Modeling*. Hyannis, Massachusetts, pp. 19–24.

Ferguson, G., Allen, J. and Miller, B.: 1996, TRAINS-95: towards a mixed-initiative planning assistant. In: *Proceedings of the Third Conference on Artificial Intelligence Planning Systems*. pp. 70–77.

Ferguson, G. and Allen, J. F.: 1998, TRIPS: an intelligent integrated problem-solving assistant. In: *Proceedings of the Fifteenth National Conference on Artificial Intelligence*. Madison, Wisconsin, pp. 567–573.

Finin, T.: 1983, Providing help and advice in task oriented systems. In: *Proceedings of the International Joint Conference on Artificial Intelligence*. Karlsruhe, West Germany, pp. 176–178.

Forbes, J., Huang, T., Kanazawa, K. and Russell, S.: 1995, The batmobile: towards a Bayesian automated taxi. In: *Proc. of the Fourteenth International Joint Conference on Artificial Intelligence*. Montreal, Canada, pp. 1878–1885.

Gertner, A.: 1997, Plan recognition and evaluation for on-line critiquing. *User Modeling and User-Adapted Interaction* **7**(2), 107–140.

Gertner, A. and Webber, B. L.: 1996, A bias towards relevance: recognizing plans where goal minimization fails. In: *Proceedings of the Thirteenth National Conference on Artificial Intelligence*. Portland, Oregon, pp. 1133–1138.

Goodman, B. A. and Litman, D. J.: 1992, On the interaction between plan recognition and intelligent interfaces. *User Modeling and User-Adapted Interaction* **2**(1–2), 83–115.

Greer, J. and Koehn, G.: 1995, The pecularities of plan recognition for intelligent tutoring systems. In: *Proceedings of the IJCAI-95 Workshop on the Next Generation of Plan Recognition Systems*. Montreal, Canada, pp. 54–59.

Haller, S. and Shapiro, S.: 1996, IDP – an interactive discourse planner. In: *Trends in Natural Language Generation – An Artificial Intelligence Perspective*. Springer-Verlag, pp. 144–167.

Kautz, H. and Allen, J.: 1986, Generalized plan recognition. In: *Proceedings of the Fifth National Conference on Artificial Intelligence*. Philadelphia, Pennsylvania, pp. 32–37.

Kay, J.: 2001, Learner control. *User Modeling and User-Adapted Interaction* **11**(1–2), 111–127 (this issue).

Konolige, K. and Pollack, M.: 1989, Ascribing plans to agents: preliminary report. *Proceedings of the International Joint Conference on Artificial Intelligence.* Detroit, Michigan, pp. 924–930.

Lambert, L. and Carberry, S.: 1991, A tripartite plan-based model of dialogue. In: *Proceedings of the 29th Annual Meeting of the ACL.* Berkeley, CA, pp. 47–54.

Lambert, L. and Carberry, S.: 1992, Modeling negotiation subdialogues. In: *Proceedings of the 30th Annual Meeting of the ACL.* Newark, Delaware, pp. 193–200.

Lesh, N.: 1997, Adaptive goal recognition. In: *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence*. Madison, Wisconsin, pp. 1208–1204.

Lesh, N. and Etzioni, O.: 1996, Scaling up goal recognition. In: *Proceedings of the 5th International Conference on Knowledge Representation and Reasoning*. pp. 178–189.

Lesh, N., Rich, C. and Sidner, C.: 1999, Using plan recognition in human-computer collaboration. In: *Proceedings of the Seventh International Conference on User Modeling*. Banff, Canada, pp. 23–32.

Litman, D.: 1986, Linguistic coherence: a plan-based alternative. In: *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*. New York, New York, pp. 215–223.

Litman, D. and Allen, J.: 1987, A plan recognition model for subdialogues in conversation'. *Cognitive Science* **11**, 163–200.

Lochbaum, K.: 1998, A collaborative planning model of intentional structure. *Computational Linguistics* **24**(4), 525–572.

Mayfield, J.: 1992, Controlling inference in plan recognition. *User Modeling and User-Adapted Interaction* **2**(1–2), 83–115.

Mooney, R. J.: 1990, Learning plan schemata from observation: explanation-based learning for plan recognition. *Cognitive Science* **14**, 483–509.

Pearl, J.: 1988, *Probabilistic Reasoning in Intelligent Systems*. San Mateo, California: Morgan Kaufman.

Perrault, R. and Allen, J.: 1980, A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics* **6**(3–4), 167–182.

Pollack, M.: 1986a, A model of plan inference that distinguishes between the beliefs of actors and observers. In: *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*. New York, New York, pp. 207–214.

Pollack, M.: 1986b, Inferring domain plans in question-answering. Ph.D. thesis, University of Pennsylvania, Philadelphia, Pennsylvania.

Pollack, M.: 1987, Some requirements for a model of the plan-inference process in conversation. In: R. Reilly (ed.): *Communication Failure in Dialogue*. Amsterdam, The Netherlands: Elsevier, pp. 245–256.

Pynadath, D. V. and Wellman, M. P.: 1995, Accounting for context in plan recognition, with application to traffic monitoring. In: *Proceedings of the 11th International Conference on Uncertainty in Artificial Intelligence*. pp. 472–481.

Quilici, A., Dyer, M. and Flowers, M.: 1988, Recognizing and responding to plan-oriented misconceptions. *Computational Linguistics* **14**(3), 38–51.

Ramshaw, L.: 1994, Correcting real-word spelling errors using a model of the problem-solving context. *Computational Intelligence*, pp. 185–211.

Ramshaw, L. A.: 1989, A metaplan model for problem-solving discourse. In: *Proceedings of the Fourth Conference of the European Chapter of the Association for Computational Linguistics*. Manchester, England, pp. 35–42.

Raskutti, B. and Zukerman, I.: 1991, Generation and selection of likely interpretations during plan recognition in task-oriented consultation systems. *User Modeling and User-Adapted Interaction* **1**(4), 323–353.

Retz-Schmidt, G.: 1991, Recognizing intentions, interactions, and causes of plan failures. *User Modeling and User-Adapted Interaction* **1**(2), 173–202.

Rose, C. P., Di Eugenio, B., Levin, L. and Van Ess-Dykema, C.: 1995, Discourse processing of dialogues with multiple threads. In: *Proceedings of the 33rd Meeting of the Association for Computational Linguistics*. Cambridge, Massachusetts, pp. 31–38.

Schmidt, C. F., Sridharan, N. S. and Goodson, J. L.: 1978, The plan recognition problem: an intersection of psychology and artificial intelligence. *Artificial Intelligence* **11**, 45–82.

Shafer, G.: 1976, *A Mathematical Theory of Evidence*. Princeton, New Jersey: Princeton University Press.

Shrager, J. and Finin, T.: 1982, An expert system that volunteers advice. In: *Proceedings of the Second National Conference on Artificial Intelligence*, pp. 339–340.

Sider, J. S. and Burger, J. D.: 1992, Intention structure and extended responses in a portable natural language interface. *User Modeling and User-Adapted Interaction* **2**(1–2), 155–179.

Sidner, C. L.: 1983, What the speaker means: the recognition of speakers' plans in discourse. *Computers and Mathematics With Applications* **9**(1), 71–82.

Sidner, C. L. and Israel, D.: 1981, Recognizing intended meaning and speakers' plans. In: *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*. Vancouver, Canada, pp. 203–208.

Smith, R., Hipp, R. and Biermann, A.: 1995, An architecture for voice dialogue systems based on prolog-style theorem proving. *Computational Linguistics*, pp. 281–320.

Spencer, B., Cohen, R. and Hoyt, P.: 1994, 'Designing a tool to allow updates during plan recognition – challenges and applications. In: *Proceedings of the IEEE Conference on Tools for Artificial Intelligence*. New Orleans, LA, pp. 63–70.

van Beek, P.: 1987, A model for generating better explanations. In: *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*. Stanford, California, pp. 215–220.

Van Beek, P. and Cohen, R.: 1991, Resolving plan ambiguity for cooperative response generation. In: *Proceedings of the 12th International Joint Conference on Artificial Intellligence*. Sydney, Australia, pp. 938–944.

Waern, A.: 1994, Plan inference for a purpose. In: *Proceedings of the Fourth International Conference on User Modeling*. Hyannis, Massachusetts, pp. 93–98.

Webb, G., Pazzani, M. and Billus, D.: 2001, Machine learning for user modeling. *User Modeling and User-Adapted Interaction* **11**(1–2), 19–29 (this issue).

Weida, R.: 1995, Knowledge representation for plan recognition. In: *Proceedings of the IJCAI-95 Workshop on the Next Generation of Plan Recognition Systems*. Montreal, Canada, pp. 119–123.

Weida, R. and Litman, D.: 1992a, Terminological plan reasoning and recognition. In: *Proceedings of the Third International Workshop on User Modeling*. Dagstuhl, Germany, pp. 177–191.

Weida, R. and Litman, D.: 1992b, Terminological reasoning with constraint networks and an application to plan recognition. In: *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning*. Cambridge, Massachusetts, pp. 282–293.

Wilensky, R.: 1978, Why John married Mary: understanding stories involving recurring goals. *Cognitive Science* **2**, 235–266.

Wilensky, R.: 1980, Meta-planning. In: *Proceedings of the First National Conference on Artificial Intelligence*. Stanford, California, pp. 334–336.

Wilensky, R.: 1981, Meta-planning: representing and using knowledge about planning in problem solving and natural language understanding. *Cognitive Science* **5**, 197–233.

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 Modeling and User-Adapted Interaction* **1**(2), 149–172.

Zukerman, I. and Albrecht, D.: 2001, Predictive statistical models for user modeling. *User Modeling and User-Adapted Interaction* **11**(1–2), 5–18 (this issue)

Zukerman, I. and Litman, D.: 2001, Natural language processing and user modeling: synergies and limitations. *User Modeling and User-Adapted Interaction* **11**(1–2), 129–158 (this issue).

Author's Vita

Sandra Carberry is professor and chair of the computer science department at the University of Delaware. She received a B.A. in mathematics from Cornell University, an M.S. in computer science from Rice University, and a PhD in computer science from the University of Delaware. Her current research interests include user modeling, computational linguistics with an emphasis on dialogue systems, and human computer interfaces. Her book 'Plan Recognition in Natural Language Dialogue' was published by MIT Press in 1990.