What does a System Need to Know to Understand a User's Plans?

Bhavani Raskutti & Ingrid Zukerman
Computer Science Department
Monash University
Clayton, VICTORIA 3168, AUSTRALIA
email: {banu,ingrid}@bruce.cs.monash.edu.au

Abstract

During natural language interactions, it is often the case that a set of statements issued by a speaker/writer can be interpreted in a number of ways by a listener/reader. Sometimes the intended interpretation can not be determined by considering only conversational coherence and relevance of the presented information, and specialized domain knowledge may be necessary in choosing the intended interpretation. In this paper, we identify the points during the inference process where such specialized knowledge can be successfully applied to aid in assessing the likelihood of an interpretation, and present the results of an inference process that uses domain knowledge, in addition to other factors, such as coherence and relevance, to choose the interpretation intended by the speaker. Our mechanism has been developed for use in task-oriented consultation systems. The particular domain that we have chosen for exploration is that of a travel agency.

Introduction

During natural language interactions, it is often the case that a set of statements issued by a speaker/writer can be interpreted in a number of ways by a listener/reader. For instance, there are two possible interpretations of the statements "John was depressed. He bought a rope." One of them is that John has a plan to hang himself. Another is that John has a plan to do some rope skipping. The first one is the interpretation that most people would prefer. This is explained by the fact that in coherent conversations, people expect statements to be linked to a central theme. In addition, the interpretation that the rope is to be used for skipping has no explanation for John wanting to do skipping. Hence, all the information presented cannot be used relevantly. Thus, in this case, the preferred interpretation is arrived at by taking into account conversational coherence and the relevance of the information that was presented.

However, there are situations where the intended interpretation can not be determined by considering these two factors alone, and specialized domain knowledge may be necessary in choosing the correct interpretation. For instance, consider the following (real life) request of a lady to her husband, an hour before they were due to report at the airport for their return flight from overseas, "Can you fill petrol and have the films developed?" The husband inferred that the jobs could be performed in any order and hence filled petrol first. He had neglected to use the domain knowledge that the camera shop needs one hour to process the film, and found out that they could not process the film.

In this paper, we identify the points during the inference process where specialized knowledge can be successfully applied to aid in assessing the likelihood of an interpretation, and present the results from an integrated mechanism that considers multiple interpretations, and uses domain knowledge, in addition to other factors, such as coherence and relevance, to choose the interpretation intended by the speaker. Our mechanism has been developed for use in task-oriented consultation systems. The particular domain that we have chosen for exploration is that of a travel agency.

An interpretation of a user's statements is a set of plans that the user proposes to carry out, and a plan consists of an action with a number of parameters defining the action. For instance, in the travel domain, the proposal to fly from Melbourne to Sydney on December 1st, 1991, is a plan, where *flying* is the action, and the parameters *departure location*, *arrival location* and *departure date* are instantiated.

A number of researchers have considered the various factors that influence the choice of an interpretation. Litman and Allen (1987) take discourse coherence into account to prefer an interpretation. However, since they consider only one interpretation, they cannot cope with situations where a preferred interpretation must be given up in light of new information. For instance, if the above sample statements regarding John are followed by the revelation "Skipping always cheered him up," then the intended interpretation can be arrived at only if the previously discarded interpretation is reinstated. The problem of multiple interpretations has been considered by Carberry (1990) and by Goldman and Charniak (1991). However, the influence of different factors is hidden in the priors chosen for the different hypotheses or contexts. Thus, extensions to their systems to accommodate domain knowledge considerations will not be modular.

In this paper, we present a method for incorporating domain knowledge into a domain independent inference mechanism. In our discussion, we refer to the inference mechanism presented in [Raskutti & Zukerman 1991]. This mechanism addresses the problem of choosing the intended interpretation among multiple possibilities by generating all possible interpretations of a speaker's statements and tagging them with a likelihood measure. The likelihood tag is used to determine the interpretations that have to be maintained during the processing as well as to select the preferred interpretation(s). The likelihood measure used by the mechanism is calculated using Bayesian theory of probability and it is based on two factors: (1) discourse coherence — which favours discourses which are closer to normal patterns of conversation; and (2) relevance of the presented information — which is determined by the *information content* of an interpretation. In this paper, we add another factor, namely domain knowledge, which favours the interpretations that are more likely within a particular domain.

The incorporation of domain knowledge to determine the intended interpretation is based on the following factors: (1) the extent of the usage of a plan for a particular purpose, (2) the feasibility of a single plan and of the overall combination of the plans in an interpretation, and (3) the practicality of a single plan and of the overall combination of the plans in an interpretation. The subsequent sections discuss these factors with reference to the mechanism described below.

The Main Mechanism

In this section, we describe briefly our algorithm for generating and selecting interpretations from a user's utterances. The algorithm may be roughly divided into two inference processes: (1) Direct inference, and (2) Indirect inference. Direct inferences are those that are drawn on the basis of the user's statements, the definition of domain actions and discourse coherence considerations. Indirect inferences are those that are based on domain and world knowledge.

The direct inference stage consists of the inference of interpretations that are based on the definition of the basic actions of the domain and the user's statements. This stage is composed of three parts: (1) the inference of interpretations for each of the user's statements, (2) the inference of discourse relations between the interpretations generated from one statement and the interpretations of the earlier discourse, and (3) the generation and selection of new interpretations based on the first two parts.

The indirect inference stage is used to complete the definition of the plans in the interpretations generated by the direct inference process. During this stage, parameters of a plan are instantiated using sources other than the user's statements, such as plan relationships, and domain and world knowledge. The plan relations we consider are causal relations, such as ENABLE and CONSTRAIN, and temporal relations, such as BEFORE and AFTER. The temporal relations can be either inferred or explicitly stated. Other bases for inference, such as domain and world knowledge, are organized as indirect inference rules.

During both inference stages, a number of factors, such as discourse coherence, relevance of the presented information and domain knowledge, are taken into account to derive the likelihood of each of the interpretations. We have chosen Bayesian probability theory over other numerical methods for reasoning under uncertainty during plan inference (Raskutti & Zukerman 1991). Bayes theory has often been criticized for its computational complexity and the need for independence assumptions during conditional probability calculations. But other numerical methods, such as Dempster-Shafer(D-S) calculus, suffer from the same drawbacks. In addition, they are often explained in terms of probability thus admitting that probability theory is on a much firmer footing than other numerical methods (Goldman & Charniak 1991). For instance, D-S calculations are at least as expensive as probability updating since a limit case of the D-S calculus is the same as point-valued probability theory. Further, D-S calculus requires the availability of a complete set of disjoint hypotheses necessitating extensive independence assumptions.

After each inference stage, the likelihood calculated using Bayes theory is used to choose between competing interpretations by dropping those interpretations with likelihoods lower than a rejection threshold. This rejection threshold is a function of the maximum likelihood of the possible interpretations, so that only those interpretations with a low likelihood relative to the most likely interpretation are dropped. For instance, if there are two interpretations with likelihoods 0.54 and 0.46 (see Section *A Sample Run*), then surely there is no clear winner. Hence, both interpretations are retained and the information in the interpretations is used to query the user intelligently.

Domain Knowledge Considerations

In this section, we discuss the three types of domain knowledge applied by our system, namely extent of usage, feasibility and practicality.

Extent of Usage

The inference of interpretations from one statement issued by a user consists of inferring a set of possible plan schemas which match this statement, and computing the likelihood of each plan-schema. The inference of plan schemas is done by using a STRIPS-like operator library (Fikes & Nilsson 1971) and plan inference rules (Allen & Perrault 1980). Each operator in the operator library represents a domain action, and it is assigned a prior probability indicating the extent of its usage. In the travel domain, the extent of usage of an operator cannot be determined unless some of parameters in the operator are instantiated. Hence, all the operators, such as, FLY and TAKE_TRAIN are assigned an equal prior. However, in other domains, the prior can be used to favour a more commonly used operator. For instance, if a person wants to send a parcel overseas, it is more natural to assume that s/he would be going to the post office to do it, and the priors can be set to reflect this preference.

Extent of usage can also be used during the indirect inference process when indirect inference rules are used to instantiate the parameters of a plan. For instance, in our system, which operates in the travel domain, the transport schedule is organized so that the listing of available transports is in the order of the extent of usage of a particular mode of transport. During the inference process, when the *mode of transport* or the *departure/arrival time* is to be inferred, the first entry in the transport schedule that is in agreement with the earlier inferences is chosen to instantiate the required parameter(s). Thus, the extent of the usage essentially allocates a likelihood of 1 to the most commonly used option. If this option is unacceptable, it can always be altered at the user's request.

Feasibility

Feasibility of a plan indicates whether a plan is achievable. Plans that are not feasible are not executable by any means. For instance, a plan where the departure date is after the arrival date is not feasible. Pollack (1986) also considers plans that are not executable. However, the focus of her research is on understanding the mental processes applied by the speaker in developing these plans, while our focus is on generating an interpretation of a user's request under the assumption that the preferred interpretation should be feasible.

An interpretation is feasible, if all the plans in it are feasible, and it is possible to achieve the plans in the inferred temporal order. For instance, while both the plans to go to Canberra today and Sydney tomorrow are feasible, the interpretation consisting of the two plans is feasible only if the trip to Canberra is before the trip to Sydney. Interpretations that are not feasible are assigned a likelihood of 0, thus eliminating them from the set of likely interpretations. Notice however, that if the user has explicitly requested a plan or a sequence of plans that is not feasible, then the interpretation containing this plan or sequence is retained, and the system must generate clarification queries. The mechanism for generating such queries is the subject of future research.

The feasibility of an interpretation is determined in two stages: (1) by determining the feasibility of all the plans in it, and (2) by determining whether the plans can be performed in the inferred temporal order. The system attempts to determine feasibility at the earliest possible opportunity. To this effect, whenever an inference gives rise to a new instantiation of a parameter, the system checks whether it has enough information to determine feasibility. For instance, if a new inference instantiates the parameter *mode of transport* to be train, and the original plan was to travel from Adelaide to Los Angeles, then clearly this plan, and hence the interpretation containing this plan, are not feasible.

The usage of domain knowledge to determine feasibility is done at the following points: (1) in the direct inference stage, during the inference of discourse relations, e.g., when a user requests for a flight to Hawaii and then asks for a means to get to Sydney, clearly the second request cannot be elaborating on the plan consisting of the flight to Hawaii; (2) in the indirect inference stage, during the inference of a temporal sequence of plans, e.g., the real life situation where the temporal sequence of filling petrol and then going to the camera shop makes the second plan unexecutable; (3) again, in the indirect inference stage, during the inference of parameters using indirect inference rules, e.g., if the *departure time* of a trip is inferred from the user's preference, and there is no transport scheduled at that time, then clearly the plan with the inferred departure time is not feasible.

Practicality

The practicality of a plan is a measure of the ease with which the effects of the plan can be achieved. A plan is practical when it is easily executable. For instance, a plan to service your car at a faraway garage is impractical, though feasible. In the travel domain, a plan to fly overseas from an international airport is practical. The practicality of a plan can be determined at all those points where the feasibility of an interpretation can be assessed. For instance, when both the *origin* and *destination* of a trip are known, then the practicality of the trip can be determined by checking for a direct route between the two places. If a direct means cannot be found, then the likelihood of the interpretation containing the plan is reduced to reflect this fact. However, the reduction is not to the extent that the interpretation would be dropped off, unless it is already an unlikely interpretation. Thus, the system can handle situations where the user intended to achieve an impractical plan by means of a series of sub plans.

The practicality of an interpretation is a measure of the ease with which the sequence of plans in the interpretation can be performed. For instance, the interpretation consisting of three plans, namely, dropping off the car for service, going to work from the garage and picking up the car on the way home, is practical only when the plans are performed in this sequence. A measure of the practicality of an interpretation can be determined only when the plans in the interpretation have been completely detailed. Hence, it is determined

Number of completely defined interpretations is 0

Interpretation 1 with probability 0.6666 **(0.6)** consists of 2 legs:
Fly to *Sydney*
Fly from *Adelaide* to *Hawaii*
departing on 16*th March* 1991 at 11:00 *am*

Interpretation 2 with probability 0.3333 **(0.4)** consists of 2 legs:
Fly from *Adelaide* to *Sydney*
Go to *Hawaii*
departing on 16*th March* 1991 at 11:00 *am*

Figure 1: **Interpretations after Direct Inferences**

Number of completely defined interpretations is 3

Interpretation 1.1 with probability 0.4444 **(0.3478)** consists of 2 legs:
Fly from *Melbourne* to *Sydney*
departing on date < 16*th March* 1991 at time < 10:00 *am*
arriving on date < 16*th March* 1991 at time < 11:00 *am*
Fly from *Adelaide* to *Hawaii*
departing on 16*th March* 1991 at 11:00 *am*

Interpretation 2.1 with probability 0.2222 **(0.2398)** consists of 2 legs:
Fly from *Adelaide* to *Sydney*
departing on date < 16*th March* 1991 at time < 9:30 *am*
arriving on date < 16*th March* 1991 at time < 11:00 *am*
Go from *Sydney* to *Hawaii*
departing on 16*th March* 1991 at 11:00 *am*

Interpretation 2.2 with probability 0.1111 **(0.1449)** consists of 2 legs:
Go from *Melbourne* to *Hawaii*
departing on 16*th March* 1991 at 11:00 *am*
arriving on 16*th March* 1991 at 11:00 *pm*
Fly from *Adelaide* to *Sydney*
departing on date > 16*th March* 1991 at time > 11:00 *pm*
arriving on date > 16*th March* 1991 at time > 12:30 *pm*

Interpretation 1.2 with probability 0.2222 **(0.2173)** consists of 2 legs:
Fly from *Adelaide* to *Hawaii*
departing on 16*th March* 1991 at 11:00 *am*
Fly from *Hawaii* to *Sydney*

Figure 2: **Interpretations during Indirect Inferences**

after all the inferences have been performed in order to enable pruning of the set of likely interpretations.

Our system, operating in the travel domain, determines a measure of practicality of an interpretation as follows: (1) Between an interpretation with a zig-zag route and another one with a straightforward route, the interpretation with the straightforward route is preferred, (2) If there are trips overseas, an interpretation that has unnecessary country crossings is not preferred. We consider country crossings to be unnecessary, when the distances between the countries are relatively large, e.g., between U.S.A. and Australia. On the other hand, in Europe, where the distances between countries are much smaller, an extra trip between countries may be justified.

In addition to the above, other criteria may also be used to determine the practicality of an interpretation. For instance, the duration of travel to a place and the duration of the stay at a place may be taken into account. However, we have not found the necessity to do so in the current implementation.

A Sample Run

Consider the following scenario:
Traveler: "Get me a flight ticket to Sydney.
I am going to fly to Hawaii at
11:00 am the day after tomorrow.
I'll be leaving from Adelaide."

These statements are input to the mechanism as the following three predicates:
(1) FLY (destination = *Sydney*)
(2) FLY (dep_time = *11:00 am*,
dep_date = *day after tomorrow*,

Number of completely defined interpretations is 1

Interpretation 2.1.1 with probability 0.3781 (**0.5400**) consists of 4 legs:
Fly from *Melbourne* to *Adelaide*
departing on date < *16th March* 1991 at time < 8:00 *am*
arriving on date < *16th March* 1991 at time < 9:30 *am*
Fly from *Adelaide* to *Sydney*
departing on date < *16th March* 1991 at time < 9:30 *am*
arriving on date < *16th March* 1991 at time < 11:00 *am*
Fly from *Sydney* to *Hawaii*
departing on *16th March* 1991 at 11:00 *am*
arriving on *16th March* 1991 at 11:00 *pm*
Fly from *Hawaii* to *Melbourne*
departing on date > *16th March* 1991 at time > 11:00 *pm*
arriving on date > *16th March* 1991 at time > 11:00 *pm*

Interpretation 1.1.1 with probability 0.6218 (**0.4599**) consists of 4 legs:
Fly from *Melbourne* to *Sydney*
departing on date < *16th March* 1991 at time < 8:30 *am*
arriving on date < *16th March* 1991 at time < 9:30 *am*
Fly from *Sydney* to *Adelaide*
departing on date < *16th March* 1991 at time < 9:30 *am*
arriving on date < *16th March* 1991 at time < 11:00 *am*
Fly from *Adelaide* to *Hawaii*
departing on *16th March* 1991 at 11:00 *am*
Fly from *Hawaii* to *Melbourne*

Figure 3: **Interpretations after Completion and Pruning**

destination = *Hawaii*)
(3) LEAVE (origin = *Adelaide*)

The mechanism was run first without domain knowledge, and then run with domain knowledge. The actual output of these runs, with today's date set to the 14th of March 1991, is presented in the Figures 1, 2 and 3. The likelihoods with domain knowledge considerations appear in brackets in boldface, next to the likelihoods without domain knowledge considerations. In this example, all the interpretations generated without using domain knowledge considerations are feasible even when domain knowledge is used. Therefore, the interpretations generated during each run are identical, but the likelihoods of the interpretations are different due to the effect of the practicality considerations during the run when domain knowledge is used. These differences are discussed below.

During the direct inference stage, while inferring discourse relations, the run without domain knowledge prefers the interpretation that the user is flying from Adelaide to Hawaii due to coherence considerations. The run with domain knowledge decreases the likelihood of the first interpretation, since there is no direct means of transport from Adelaide to Hawaii (see Figure 1). During the indirect inference stage, while using the inferred temporal sequence of plans, the likelihoods of interpretations 2.1 and 2.2 are higher with domain knowledge than without, since Hawaii is an overseas location and Melbourne and Sydney are international ports (see Figure 2). During the final stages of both runs, interpretation 2.2 is rejected owing to its low initial probability, while interpretation 1.2 is rejected due to its low information content (Raskutti & Zukerman 1991). The interpretation intended by the user, i.e., interpretation 2.1.1 in Figure 3, is determined to be more practical since the path of its itinerary is less meandering than the path of the other interpretation. Thus, with the use of domain knowledge the interpretation intended by the user is the one with the highest likelihood. However, since the likelihoods of the two interpretations are close there is no clear winner and the user has to be queried to select one of these interpretations. During the generation of a dicriminating query, the information inferred during the plan recognition process can be used to generate a sensible query. Without the use of domain knowledge, the incomplete interpretation that does not represent the user's intent is clearly preferred, and this interpretation needs to be completed by means of an information seeking query that would probably confuse the user. Hence, the application of domain knowledge to aid plan recognition and the subsequent response generation can make sense out of an incoherently phrased request.

Conclusion

In this paper, we have demonstrated the need to use domain knowledge to choose between interpretations. We have described the basis of the application of domain knowledge in our system, and have indicated the points during the inference process where domain knowledge can be successfully applied. We have also presented the results of the application of our ideas in processing a sample request at a travel agency.

Work in Progress

In order to generate cooperative responses, the interpretations generated by the plan inference mechanism have to be analyzed as follows: (1) If there is only one valid interpretation with sufficient relevant information, then it is sent to the planner for planning. If the planner can come up with an itinerary that satisfies the user's needs, then the system has to inform the user about the proposed itinerary; (2) if there are multiple possibilities, as in the case of the example presented above, then the possibilities must be analyzed to determine the points that discriminate between them, so that a clarification question can be generated; finally, (3) if there are no complete interpretations, the user must be queried to complete an interpretation. Presently, we are in the process of incorporating the response generation module into the inference mechanism.

Our mechanism for plan inference has been mainly used in a travel planner where the domain constraints, such as time and location dependencies, are fairly clear. We are currently considering other domains, such as Yellow Pages directory assistance system, to evaluate the portability of our mechanism (Zukerman 1991).

References

Allen, J.F. & Perrault, C.R. (1980), Analyzing Intention in Utterances. In *Artificial Intelligence* 15, pp. 143-178.

Carberry, S. (1990), Incorporating Default Inferences into Plan Recognition. In *Proceedings of the National Conference on Artificial Intelligence*, pp. 471-478.

Fikes, R.E. & Nilsson, N.J. (1971), STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving. In *Artificial Intelligence* 2, pp. 189-208.

Goldman, R. & Charniak, E. (1991), A Probabilistic Model for Plan Recognition. In *Proceedings of the Ninth National Conference on Artificial Intelligence*, pp. 160-165.

Litman, D. & Allen, J.F. (1987), A Plan Recognition Model for Subdialogues in Conversation. In *Cognitive Science* 11, pp. 163-200.

Pollack, M. (1986), A Model of Plan Inference that Distinguishes between the Beliefs of Actors and Observers. In *Proceedings of the 24th Annual Meeting of the Association of Computational Linguistics*, pp. 207-214.

Raskutti, B. & Zukerman, I. (1991), Generation and Selection of Likely Interpretations during Plan Recognition in Task-oriented Consultation Systems. In *User Modeling and User Adapted Interaction — an International Journal* 1(4).

Zukerman, I. (1991), Towards a System for the Generation of Cooperative Responses in Information Seeking Environments. Technical Report, 91/154, Department of Computer Science, Monash University.