Interactive Problem Solving and Dialogue in the ATIS Domain

Stephanie Seneff, Lynette Hirschman, and Victor W. Zue

Spoken Language Systems Group
Laboratory for Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT

This paper describes the present status of the discourse and dialogue models within the MIT ATIS system, extended to support the notion of booking a flight. The discourse model includes not only the resolution of explicit anaphoric references, but also indirect and direct references to information mentioned earlier in the conversation, such as a direct reference to an entry in a previously displayed table or an indirect reference to a date, as in “the following Thursday.” The system keeps a history table containing objects such as flights and dates, represented as semantic frames, as well as the active ticket, previously booked tickets, and previously displayed tables. During flight reservations scenarios, the system monitors the state of the ticket (which is displayed to the user), making sure that all information is complete (by querying the user) before allowing a booking. It may even initiate calls to the database to provide additional unsolicited information as appropriate. We have collected several dialogues of subjects using the system to make reservations, and from these, we are learning how to design better dialogue models.

INTRODUCTION

A key goal of spoken language systems is to provide support for interactive problem solving. To be effective in this role, a system should permit the user to build up a solution incrementally, and therefore to make implicit and explicit reference to information from earlier parts of the system/user dialogue. In addition, in certain scenarios, e.g., making flight reservations, it is natural to have the system play an active role in the dialogue. It is in fact our belief that spoken language systems will never reach a practical level of performance unless special attention is paid to issues of discourse and dialogue. To this end, we spent considerable effort over the last several months augmenting our system’s discourse capabilities, and developing a much more active role for the computer during the dialogue with a user.

We can view interactive problem solving systems on a continuum of system/user control. At one extreme would be a system that answers user’s questions without at any point asking questions or making suggestions. At the other extreme would be a menu system that forces the user into a very limited range of choices. This paper will describe an attempt at MIT to explore a middle ground in this continuum, in the context of allowing users to make flight reservations.

There are a number of advantages to an approach which explores a mixed initiative system for flight reservations. First, it provides the user with a more focused way to interact: making a reservation provides a natural notion of task completion. Second, a system that can take the initiative on occasion not only can provide a more flexible and natural interaction (e.g., asking for clarification, eliciting missing information) but also can reduce the search space for spoken input, thus potentially improving performance.

There are also a number of potential disadvantages to such an approach. One problem with making reservations is the difficulty in eliciting clean data that does not contain extraneous information (names of travelers, VISA card numbers, etc.) A problem that arises when the system takes the initiative is that it must have substantial problem solving expertise; it crosses the line into an expert system, and therefore may extend beyond the scope of the Spoken Language Systems effort as currently defined. Third, such systems are much more difficult to evaluate automatically. Finally, by making the system more natural, we run the risk of raising user expectations too high.

This paper describes the present status of the discourse and dialogue models within the MIT ATIS system. After describing the models, we will illustrate some of the system’s capabilities by way of an example. We then describe our preliminary attempts at collecting data in a booking mode, for which we have included a complete dialogue elicited from one of our subjects. Finally, potential implications for improvements in speech recognition are discussed.

MODELLING METHODOLOGY

The back-end component of the MIT ATIS system has been completely redesigned since last June [5]. The main system is described in detail in [4], and will only be briefly

1This research was supported by DARPA under Contract N00014-89-J-1332, monitored through the Office of Naval Research.
A flight that is incompletely specified in a new sentence inherits modifiers that are consistent with its current state. Defining consistency is tricky and requires knowledge of how information is structured in the domain. Within the ATIS domain, the explicit mention of a flight number is taken to mark a change of focus, and therefore blocks almost all inheritance modifiers except source and destination. Similarly, the modifier “cheapest” would block an inheritance of a specific flight number, since it implies taking a subset of a previously mentioned set. Whenever both a new source and a new destination are present, all inheritance is blocked, unless the new sentence was a clarifier, such as “How about from Boston to Denver?” Of course, a modifier always blocks inheritance of an entry under the same key in the history table. Exactly which modifiers should block which others was determined empirically from subject data through the data collection episodes.

The history table contains not only the frames associated with previously mentioned noun phrases and their modifiers, but also the previously displayed table, the previous state of the ticket under development, and previously booked tickets or first legs of a round trip ticket. The system frequently consults the tickets, as well as other elements from the history, to decide what directed questions to ask the user.

Occasionally, the history elements have to be reinterpreted after being inherited. This is particularly true for “return” flights, which can be mentioned in a number of different ways: using an adjective or a verb phrase modifier, with or without explicit mention of a source, destination, or date, and with or without a mention of a forward leg in the same sentence. It turned out to be quite difficult to make all conditions work out, inheriting source and destination when appropriate, and reversing them only if they came from a history flight that was not also marked as a return flight. Return flights also inherit fare class, fare restrictions, and airline. In addition, a restricted weekday fare is generalized to include a compatible weekend fare, and some restrictions require a minimum/maximum stay restricting the return date. Finally, users often mention the return date early in the dialogue, in which case the system stores it and recalls it later, when the topic turns to the return leg.

Figure 2 gives a block diagram of the control flow for managing discourse and dialogue. As shown in the figure, both the user and the computer may issue questions to the back-end component. These questions are processed the same discourse references [1]. However it has been sufficient for the ATIS domain to date.
way, updating both the discourse and dialogue components accordingly. For instance, when the user has booked a particular flight but has said nothing about fares, the system can simply issue the request “Show fares” to the back-end. The discourse history will incorporate automatically the relevant flight information. If a user query is ambiguous, the system defers calling the database until it has queried the user for resolution of the ambiguity. After the computer has answered the user’s question, it assesses the dialogue state, which is maintained as a stack. When the state stack is popped, the system may update the information contained in the ticket. The computer may decide at this point to take the initiative, anticipating the user’s needs.

Consider the example in which the user says, “Book the cheapest flight.” The system does not immediately know whether it should find the cheapest one-way fare or the cheapest round-trip fare. It also must remember, however, that a booking has been requested. The system pushes <booking> onto the dialogue state, followed by <resolve flight cycle>. This is similar to the stack-based approach described in [2]. The database query function examines the top of the dialogue state and finds that more information is needed before a table can be displayed, so it does nothing. Control now passes to the computer, which asks the directed question, “One way or round trip?” and pops the top of the dialogue stack. After the question has been answered by the user, the user’s answer is incorporated into the flight-event object, and a table is displayed. Now it consults the dialogue state once again and finds <booking>, so it calls up the booking routine to fill in all the ticket information and make the next decision about what to ask.

Before deciding to query the user about the one-way/round-trip ambiguity, the computer tries hard to infer the answer from history information. Of course, if this had been specified in a previous sentence, then it would be available as a flight-modifier frame in the history. It is also possible that the table previously displayed contained only round-trip (or only one-way) fares, in which case it could decide based on the table. Finally, if the user had previously specified a return date (as in the example below), the computer would assume a round-trip fare was wanted. Only when all these conditions fail does it resort to asking the user.

**AN EXAMPLE**

The easiest way to explain some of the discourse/dialogue aspects of the system is through an example of a simulated dialogue, as shown in Table 14. The dialogue involves making a round trip reservation from Boston to San Francisco. The subject’s first sentence indicates only a destination. Rather than displaying all flights to San Francisco from anywhere, the system asks instead a directed question for clarification of the source location. The subject responds with only the word “Boston,” but the system infers “from Boston” based on the dialogue state. The system then immediately asks for a date, since this will dictate whether the flight is available on that day and also which fares apply. The next sentence from the subject, “I’ll be leaving Boston next Sunday and returning the following Tuesday,” provides two dates, the first relative to today’s date and the second relative to the first one. The system updates the reference date “on the fly” so that the appropriate date is available for reference when the phrase “the following Tuesday” is processed. The system now answers by focusing on the forward leg, holding the return date for later reference. It also provides verbal feedback verifying its understanding of the flight conditions: “These are the flights from Boston to San Francisco on Sunday, January 6,” where it has interpreted “next Sunday” as “Sunday of next week.” Finally, it fills in the source, destination, and date on the ticket form being displayed continuously to the subject.

The subject now asks a question referring to a specific entry in the list, “the third one.” The system extracts airline, flight number, source and destination from the table entry, and constructs the appropriate semantic frame, inheriting the date from the history. Again it confirms understanding with a verbal response.

Question 5 from the subject is a complicated one, containing two superlative adjectives. The system knows to first extract all non-stop flights that serve dinner⁴, then to pick the earliest among them, and finally to find the cheapest fare for that flight. If these modifiers were processed in a different order, it would not get the right answer. Usually the system would ask for a one-way/round-trip disambiguation at this point, but since the subject has already specified a return date, the system assumes they want a round-trip ticket. The cheapest fare is an unrestricted one because the restrictions are failing on advance-purchase requirements, given the date.

In Question 6, the subject says, “Book it,” where “it”

---

Footnotes:

- Tables have been abbreviated somewhat so that they will fit within a single column on the page.
- It inherits source, destination, and date from the flight set history created by Question 3.
Subject: I WANT TO GO TO SAN FRANCISCO
ATIS: Where from?
Subject: Boston.
ATIS: What date will you be travelling on?
Subject: I'LL BE LEAVING BOSTON NEXT SUNDAY AND RETURNING THE FOLLOWING TUESDAY
ATIS: These are the flights from Boston to San Francisco on Sunday January 6.

| AIRLINE | NUMBER | FROM  | TO     | DEPARTURE | ARRIVAL | STOPS |
|---------|--------|-------|-------|-----------|---------|-------|
| UNITED  | 21     | BOS   | SFO   | 8:00 A.M. | 11:31 A.M. | 0     |
| UNITED  | 343    | BOS   | SFO   | 9:40 A.M. | 1:36 P.M.  | 1     |
| AA      | 813    | BOS   | SFO   | 11:49 A.M. | 3:58 P.M.  | 1     |
| CO   | 1765   | BOS   | SFO   | 5:40 P.M. | 11:10 P.M. | 1     |
| UNITED  | 93     | BOS   | SFO   | 5:45 P.M. | 9:27 P.M.  | 0     |
| TWA     | 61     | BOS   | SFO   | 6:10 P.M. | 9:37 P.M.  | 0     |

Subject: WHERE DOES THE THIRD ONE STOP
ATIS: American flight 813 from Boston to San Francisco on Sunday January 6 stops in the following places.

| AIRLINE | NUMBER | FROM  | TO     | DEPARTURE | ARRIVAL | STOP | AIRPORT |
|---------|--------|-------|-------|-----------|---------|------|---------|
| AA      | 813    | BOS   | SFO   | 11:49A.M  | 5:58P.M. | 1    | DFW     |

Subject: WHAT IS THE CHEAPEST FARE FOR THE EARLIEST NONSTOP FLIGHT THAT SERVES DINNER
ATIS: This is the cheapest round-trip fare for the earliest non-stop flights from Boston to San Francisco serving dinner on Sunday January 6.

| AIRLINE | NUMBER | RESTRICTION | ROUND TRIP | FARE CLASS |
|---------|--------|-------------|------------|------------|
| UNITED  | 93     | NONE        | $1210.00   | Y:COACH    |

Subject: BOOK IT
ATIS: I'll show it to you again.

Airline number restriction round-trip fare class
UNITED 93 NONE $1210.00 Y:COACH CLASS
I'll book United flight 93 from Boston to San Francisco for you. Can I help you with the return flight on Tuesday January 8?

Subject: YES PLEASE
ATIS: These are the United flights from San Francisco to Boston class Y on Tuesday January 8.

| AIRLINE | NUMBER | FROM  | TO     | DEPARTURE | ARRIVAL | FARE | CLASS |
|---------|--------|-------|-------|-----------|---------|------|-------|
| UNITED  | 92     | SFO   | BOS   | 8:00A.M.  | 4:27P.M. | $1210.00 | Y:COACH |
| UNITED  | 20     | SFO   | BOS   | 1:40P.M.  | 10:02P.M.| $1210.00 | Y:COACH |

Please select a booking.

Subject: I'D LIKE TO BOOK UNITED FLIGHT NINETY TWO PLEASE
ATIS: This is United flight 92 from San Francisco to Boston class Y on Tuesday January 8.

| AIRLINE | NUMBER | FROM  | TO     | DEPARTURE | ARRIVAL | FARE | CLASS |
|---------|--------|-------|-------|-----------|---------|------|-------|
| UNITED  | 92     | SFO   | BOS   | 8:00A.M.  | 4:27P.M. | $1210.00 | Y:COACH |

I'll book united flight 92 from San Francisco to Boston for you. Can I help you with something else?

Subject: NO THANK YOU
ATIS: Have a nice day.

Table 1: Simulated dialogue involving a round trip flight reservation.

Figure 3: The ticket facsimile displayed to the subject, as it appears after Question 6 in the simulated dialogue from Table 1.

Table 3: The ticket facsimile displayed to the subject, as it appears after Question 6 in the simulated dialogue from Table 1.

Table 4: The ticket facsimile displayed to the subject, as it appears after Question 6 in the simulated dialogue from Table 1.

Table 5: The ticket facsimile displayed to the subject, as it appears after Question 6 in the simulated dialogue from Table 1.

Table 6: The ticket facsimile displayed to the subject, as it appears after Question 6 in the simulated dialogue from Table 1.

Table 7: The ticket facsimile displayed to the subject, as it appears after Question 6 in the simulated dialogue from Table 1.

Table 8: The ticket facsimile displayed to the subject, as it appears after Question 6 in the simulated dialogue from Table 1.

Table 9: The ticket facsimile displayed to the subject, as it appears after Question 6 in the simulated dialogue from Table 1.

Table 10: The ticket facsimile displayed to the subject, as it appears after Question 6 in the simulated dialogue from Table 1.

Table 11: The ticket facsimile displayed to the subject, as it appears after Question 6 in the simulated dialogue from Table 1.
frame in the history table, reversing source and destination.

The system now shows the subject two United flights from San Francisco to Boston, and asks the subject to select one. The subject selects flight 92, and the system is now ready to form a complete booking consisting of two flights tied to a single fare, adding it to a list of previous bookings. The system finally asks the subject, “Can I help you with something else?” and the subject concludes the dialogue.

COLLECTING DIALOGUE DATA

We have collected several thousand sentences from subjects using our system [3], but only about ten of the subjects were allowed to use the system in booking mode. Even for these ten, we only asked them to do one booking scenario, in addition to several non-booking scenarios. In part, this was done because data collection at TI is done in non-booking mode, and we wanted our data to be better matched to the likely TI test data. In addition, we were not confident that our booking dialogue was sufficiently robust to be ready for data collection until the last month or so. We are encouraged, however, by preliminary results of the booking scenarios. While a subject almost never gets through a booking “without a hitch,” we do find that subjects are able to specify flights to be booked and successfully complete scenarios. With each data collection episode, we gain new insights on faulty assumptions in the system. A close coupling between data collection and system development should ultimately yield a robust dialogue model, which could possibly be viewed as an “artist’s conception” of a useful system.

In the appendix can be found an example of an actual dialogue between one of our subjects and the computer. Twice during the dialogue the computer made faulty assumptions, but the subject was able to recover from the error and ultimately achieve all the goals of the scenario. Obviously, we are improving the system so that, next time around, these errors will not recur.

HELPING THE RECOGNIZER

There are potentially many ways to use discourse and dialogue to make the speech recognition task more successful. A simple step that we are taking is to restrict flight numbers to be only those that have previously been displayed in a table. Hypotheses with unlicensed flight numbers would be pruned away. This can be effective, since numbers are relatively difficult to recognize correctly.

Another feature that we have implemented has to do with ease of recovery from a recognition error. A system which remembers the history can become quite confused if it remembers false information from sentences that were incorrectly recognized. We have therefore implemented a “scratch that” or “erase that” command which allows the system to completely “forget” all information that was newly introduced in the most recent sentence. This includes erasing entries from the ticket display if the sentence requested a booking. This command is distinct from a “clear history” command, which has a more global effect of erasing all records in the history.

Finally, we hope to be able to use the dialogue state to dynamically modify probabilities on arcs in the grammar. Once we have a fully integrated recognizer, with parse probabilities incorporated into the scores of partial theories, perplexity can be reduced by rewarding paths that are supported by the dialogue state. For example, we could introduce a bonus on the <date> node whenever the system asks the question, “What date will you be travelling on?” We hope to be able to explore some of these ideas in the near future.

SUMMARY

This paper describes our current model for discourse and dialogue within the ATIS system. Heavy emphasis has been placed on the specific goal-directed task of making flight reservations. This is in part because we believe that this particular capability represents the most useful possible outcome of our efforts in developing a system based on the Official Airline Guide, potentially reaching a large population of users with no prior expert training. But in part this choice was also made because it offers a very well-defined goal with a number of well-defined subgoals, ideally suited for exploring issues in interactive human/computer dialogues. The task is nontrivial, but the potential pay-off is large, both in terms of reducing the complexity of the recognition task and improving user satisfaction. We believe that it would be particularly worthwhile to narrow down the ATIS system to be focused only on making flight reservations, discarding, for example, all references to ground transportation and questions probing the database for other types of information, such as capacity of aircraft, wingspan, etc. If we can demonstrate a system that actually works with speech input and speech/text/tables as output, then such a system should stand as a convincing demonstration of the utility of spoken language systems.

REFERENCES

[1] Dahl, D. A. and C. N. Ball, “Reference Resolution in FUn- DIT.” Chapter 8 in Logic and Logic Grammars for Language Processing, P. Saint-Dizier and S. Szpakowicz, Ed., Ellis Horwood, New York, 1990.
[2] Grosz, B. J. and C. L. Sidner, “Attention, Intentions, and the Structure of Discourse,” Computational Linguistics, Vol. 12, No. 3, pp. 175-204, July-September, 1986.
[3] Polifroni, J., S. Seneff, V. W. Zue, “Collection of Spontaneous Speech for the ATIS Domain and Comparative Analyses of Data Collected at MIT and TI,” These Proceedings.
[4] Zue, V., J. Glass, D. Goddeau, D. Goodine, L. Hirschman, H. Leung, M. Phillips, J. Polifroni, and S. Seneff, “Development and Preliminary Evaluation of the MIT ATIS System,” These Proceedings.
[5] Zue, V. J. Glass, D. Goodine, H. Leung, M. Phillips, J. Polifroni, and S. Seneff, “Preliminary ATIS Development at MIT,” Proc. Third DARPA Speech and Natural Language Workshop, Hidden Valley, PA, June 1990.
APPENDIX: A dialogue between a subject and the computer, from our data collection files.

SCENARIO: You live in Philadelphia. You need to make a business trip to San Francisco next week. You have an old friend in Dallas and you'd therefore like to spend the afternoon in Dallas on your way to San Francisco. You'd prefer to fly first class on American. Find out what kind of aircraft you'll be flying on.

SUBJECT: I'D LIKE TO FLY FROM PHILADELPHIA TO SAN FRANCISCO
ATIS: What date will you be travelling on?
SUBJECT: NEXT TUESDAY
ATIS: These are the flights from philadelphia to san francisco on Tuesday January 16 (four flights)
SUBJECT: ARE THERE FLIGHTS GO FROM PHILADELPHIA TO SAN FRANCISCO ON MONDAY
ATIS: These are the flights from philadelphia to san francisco on Monday January 15 (four flights)
SUBJECT: WHAT KIND OF AIRCRAFT IS THAT
ATIS: This is aircraft for first class one-way american flight 1074 from san francisco to philadelphia on Thursday January 18 (one entry)
SUBJECT: END SCENARIO EIGHT