A SPOKEN DIALOGUE SYSTEM UTILIZING SPATIAL INFORMATION

Annika Flycht-Eriksson and Arne Jönsson

Department of Computer and Information Science
Linköping University, S-581 83, LINKÖPING, SWEDEN
annfl@ida.liu.se arnjo@ida.liu.se

ABSTRACT

Spatial reasoning plays an important role in many spoken dialogue systems. One application area where it is especially important is timetable information for local bus traffic. Users of such systems often request information based on vague spatial descriptions and a usable system must be able to handle this. We have extended a dialogue system with abilities to transform vague spatial expressions into a form that can be used to access the information base. In our approach we use the power of a Geographical Information System (GIS) for the spatial reasoning.

1. Introduction

Spoken dialogue systems providing timetable information for local bus traffic requires both temporal and spatial reasoning capabilities. Temporal reasoning has been investigated in, for instance, Rail traffic information systems [7, 1], but as spatial reasoning is not that prominent in such applications less work has been done in this area. One important difference between Bus traffic information systems and Rail traffic information systems is that a user's natural way of expressing a departure or arrival location in a Bus traffic information system is not by means of the "official" name of a bus stop. Instead, other expressions are utilized, such as an area or town district, a set of reference points, or a street. Thus, a spoken dialogue system that provides timetable information for local bus traffic must be able to map such an imprecise description to a set of bus stops that corresponds to the description, i.e. a spatial reasoning mechanism is required.

2. Spatial information

The spatial reasoning methods presented in this paper are based on empirical investigations of dialogues between human traffic information providers and local bus traffic customers. These dialogues are not representative for users of a spoken dialogue system [4], but for exploring the possibilities of utilizing vague spatial description it forms a usable corpus. We have collected 40 dialogues with various users requesting information on buses in the county of Östergötland. From the corpus we have identified the linguistic means by which users present spatial information.

One common way of describing a departure or arrival location is by a town name, for example: *I wonder when the next bus from Skärblacka to Norrköping departs?* This implies that the spatial reasoning module must be able to map a town name to a small set of bus stops that can be used when searching the timetable database. Since small towns, in this case Skärblacka, often have few bus stops and are passed by few buses, they can be mapped to a set containing all of the bus stops in the town. Larger towns, Norrköping in the example, are more problematic to map to a set that contains few bus stops. One solution is to ask the user for information that can be used to limit the possibilities. Another approach, which is similar to the approach used by the human operators, is to use the bus stops that are most frequently passed by buses that travel between towns, such as the bus terminal or the Railroad station.

Another frequently used way to describe a location is to present an area such as a suburb. The problem is similar to the cases in which town names are used, sometimes further specification is needed in order to present a correct response, as in:

> U: Hi, I wonder when the next bus from Malmslätt to the city runs?
> S: Let's see, bus 213. Where in Malmslätt do you want to enter?

In this case the spatial reasoning module must provide information, to the dialogue manager of the interface, that a further specification is needed. The system can then use this information to formulate an appropriate clarification request.

Some users present a location by providing the exact address, e.g. *Hi, I am going to 8 Owl street.* For such examples the spatial reasoning module needs to map the address onto possible bus stops near the address. A similar example is when a user presents two places and expect the system to respond with the best route, e.g. *Hi, I would like to know how to go from the hospital down to IKEA in Linköping?*.

If a user specifies a bus stop, the system can perform a search directly in the timetable database. However, this may sometimes result in a travel route that is much worse than a travel route from another nearby bus stop. The spatial module should therefore allow fuzzy matching of a bus stop to include the nearby bus stops as alternative departure or arrival locations.

A problem not encountered in the corpus but likely to appear in a spoken dialogue system is that a name of an area, bus stop, place or an address can be ambiguous. The hospital can for example refer to many places in different towns or areas. The spatial reasoning module must therefore be able to reason about which location the user is referring to, or inform the user that further specification is necessary.

The main purpose of the spatial module is, thus, to match the locations provided by the user to bus stops that can be used to search the timetable database. In many cases this can be carried out in a straightforward fashion by simply listing the matching stops, but if this set of possible matches becomes to large there must be means to restrict the number of items.

3. Spatial representations and Geographical Information Systems

Reasoning about the relations between bus stops and other geographical objects such as buildings, addresses and regions requires a representation of spatial and geographical information. Such information can be represented and reasoned with in different ways.

In the traditional quantitative approach a coordinate system is used as a basis for the representation of objects and regions. Information about the objects properties and relations among objects is extracted by means of arithmetic and trigonometrical computations. The representation and manipulation of spatial data is done numerically [5].

Another approach is to use a qualitative representation and reasoning mechanism. Qualitative representations are symbolic and based on discrete values, the distance between two objects can for example be expressed as one of the values very close, close, far or very far [2].

Maps are also a medium for representing geographic information. In a map both geometric aspects and symbolic representations of objects can be integrated [6].

In the field of Geographical Information Systems, quantitative representations are combined with maps. A geographic information system (GIS) can be defined as "A computer-based information system that enables capture, modelling, manipulation, retrieval, analysis, and presentation of geographically referenced data." [10, p. 1].

A GIS consists of three components; a database, an analytic engine and an interface. The design of the database affects how the GIS stores and models the reality. The analytic engine is responsible for the manipulations and transformations of the data. The interface differs depending on the domain and application but one common feature is a map that can be used for visualisation of the spatial data in the database [9].

Using a GIS in a spoken dialogue system has a number of advantages. For instance, a GIS comes with a variety of predefined functions for spatial reasoning. Furthermore, much spatial data is available in a form ready to use in a GIS. Since GISs support maps there is also the possibility of constructing a multi-modal web-based interface, without modifying the underlying spatial representation.

4. The spatial reasoning module

The spatial reasoning module utilizes a GIS to represent the spatial information about bus stops, streets, places, suburbs and towns. Since traditional GISs are of a quantitative nature while spatial relations expressed in natural language are more qualitative, the spatial reasoning module must be able to transform qualitative concepts to quantitative. From our corpus, we have identified two major concepts, *in* and *near*, that need to be transformed. The spatial reasoning module maps the qualitative term *near* onto a precise distance in the quantitative representation, i.e. into an area near the location. The concept *in* describes the topological relation between bus stops, places or streets and a town or a suburb.

When the spatial reasoning module maps a departure or arrival location specified by the user in terms of a bus stop, a street, a place, a suburb and/or a town it utilizes the *in* and *near* relations. All bus stops within the distance which specify the *near* concept are gathered when mapping a place, street or bus stop to the nearby bus stops. For areas such as suburbs all the bus stops that lies in the region are collected. The same applies for small towns. Large towns are mapped to the set of bus stops which lies in the town and are key bus stops, i.e. bus stops that most of the bus routes pass. If a user has specified a location by means of different kinds of spatial information, for example a place and a street, the spatial reasoning module maps each kind of information separately to sets containing the nearby bus stops and then takes the intersection of the sets.

The *in* and *near* concepts are also used by the spatial reasoning module to resolve ambiguous spatial information which can refer to a number of different locations. When the name of a bus stop, a street, a place or an area is ambiguous the spatial reasoning module identifies the alternative locations and systematically examine how the alternatives are related to other spatial information expressed by the user.

To illustrate this, consider the ambiguous name of a place, such as *the hospital.* The reasoning steps needed to disambiguate this are:

Identify the locations with the name "the hospital", these form the set of possible alternative locations, L
If the user has specified a bus stop, B
  Collect the places, P, *near* B
  Select the subset of places which belong to both L and P, this is the new set, L, of alternative locations
  If only one alternative remains in L return L
If the user has specified a street, S
  Collect the places, P, *near* S
  Select the subset of places which belong to both L and P, this is the new set, L, of alternative locations
  If only one alternative remains in L return L
If the user has specified an area, A
  Collect the places, P, *in* A
  Select the subset of places which belong to both L and P, this is the new set, L, of alternative locations
  If only one alternative remains in L return L
If the user has specified a town, T
  Collect the places, P, *in* T
  Select the subset of places which belong to both L and P, this is the new set, L, of alternative locations
  If only one alternative remains in L return L
If no unique location has been singled out, i.e. L contains more than one alternative
  ClarificationRequest(town or suburb)

In cases where a location is mapped to a too large set of bus stops the spatial reasoning module must find a way of narrowing down the alternatives or ask the user for a specification. If the user has specified a bus route, this information can be used to select the subset of bus stops that are passed by that bus route. In cases where the fuzzy mapping of a bus stop given by the user has resulted in a too large set, the module returns only the specific bus stop named by the user. If a departure or arrival location is described in terms of a place that are mapped onto many nearby bus stops, the spatial reasoning module asks the user to select one of the stops from the set. The mapping of streets to bus stops that results in too many alternative bus stops are treated in two different ways. If the street is long, i.e. more than 300 meters, the spatial reasoning module asks for a clarification by means of a bus stop, place or suburb. Otherwise the spatial reasoning module presents the alternatives and asks the user to select one. Suburbs and areas are treated in a similar way. In case of a large area, a bus stop, place or street is requested for specification of the location. If the area is small the user has to choose one bus stop from the set of possible bus stops.

The spatial reasoning module must also be able to discover inconsistent information that the user may have provided or may be the result of a misinterpretation of an utterance. An example is a location specified in terms of a place and a town where there doesn't exist a place with that name in the town. Using the *in* and *near* relations the spatial reasoning module can find the inconsistency and explain it to the user, using the following algorithm:

If the user has specified a place, P
  If the user has specified a bus stop, B
    If P is not *near* B
      Error(P is not near B)
  If the user has specified a street, S
    If P is not *near* S
      Error(P is not near S)
  If the user has specified an area, A
    If P is not *in* A
      Error(P is not in A)
  If the user has specified a town, T
    If P is not *in* T
      Error(P is not in T)

5. The Dialogue System

The spatial inference module is integrated with the LIN-LIN dialogue system [3]. The interaction between the interface and the spatial reasoning module is mediated through the dialogue manager. The dialogue manager controls the interaction with the user. This involves mechanisms for posing questions to the spatial inference module and the timetable database based on the kind of information requested by the user. It also involves mechanisms for requesting further information from the user when a clarification is needed. Information from the spatial reasoning helps the dialogue manager to specify clarification requests based on information from the dialogue, and the GIS. The new information is then integrated with the previous description and the result is used to access the time table database.

The dialogue manager also records the entities being discussed so far. The constituents of the dialogue is modelled in dialogue objects. The dialogue objects are responsible for maintaining the dialogue and also contains focus information, such as town, street, place, bus-stop, and arrival and departure time of a bus line.

During the course of interaction a dialogue tree is built up from instances of dialogue objects. The dialogue tree serves two purposes. First, it serves as a vehicle for monitoring the dialogue, to guide decisions on how to proceed in the dialogue and where a user move fits into the dialogue, if it is to be regarded as a new initiative, a clarification request, or a response to a system initiative. The dialogue tree also records the focus parameters to be used by the referent resolving algorithms of the interpretation and generation modules.

6. An example

Consider the utterance *I would like to know how I can travel from the hospital down to IKEA in Linköping.*

The phrase "the hospital" is recognised by the parser [8] as a departure location and "IKEA in Linköping" as an arrival location. The information is passed to the spatial reasoning module as a place named "the hospital" and a pair with a place named "IKEA" and a town named "Linköping". It is the spatial reasoning module's task to map the locations to sets of bus stops. When the spatial reasoner discovers that "the hospital" is an ambiguous reference to a location, it tries to disambiguate the information. Since no more spatial information about the departure location is given, a clarification is needed. The spatial reasoning module sends a request to the dialogue manager which poses a question to the user.

S: There are many places named "the hospital".
Which town or suburb are you in?
U: Linköping.

The new information is integrated with the old by the dialogue manager and now the spatial reasoning module is given a departure location specified by the place "the hospital" and the town "Linköping". This time the spatial reasoning module succeeds when it tries to disambiguate the location of the place "the hospital". The place referred to by the user is mapped onto the bus stops *near* the place. "IKEA" is not ambiguous and is therefore mapped to the nearby bus stops. The two sets of bus stops are returned to the dialogue manager and are used to request information from the timetable database.

7. Discussion

The spatial reasoning module is developed for use in both a telephone based spoken system and a multi-modal system utilizing speech as one important modality. We can, however, use the same knowledge bases and reasoning modules for both interfaces. In the case of the multi-modal interface we can also use the map representation provided by the GIS.

The information represented in the GIS and the reasoning mechanisms in the spatial reasoning module are based on the spatial information and the problems identified in the corpus. Since humans do not interact with computer systems in the same way as they interact with other humans [4] it is possible that the spatial reasoning module can not handle all situations that occur when the spoken dialogue system is used. An evaluation of the representation and the reasoning mechanisms should therefore take place in a real setting where users interact with the system, or at least think that they interact with a system, i.e. a wizard of Oz study.

8. Conclusions

For many applications of spoken dialogue system utilizing a GIS and spatial reasoning mechanisms is beneficial. First of all it has the advantage of providing a more natural interaction, as users can express locations in a way natural to them. In our system the user does not have to know the name of a bus stop to be able to ask for timetable information which is a common restriction in this kind of bus traffic information systems.

Furthermore, the interaction becomes more robust since the spatial inference module can discover and handle inconsistent information provided by the user or due to misinterpretations. Error messages can be more informative and the spatial reasoning module can offer the user ways of correcting a mistake and continue the dialogue.

The use of spatial information also helps the dialogue manager in other kinds of clarification situations. When more information is needed to resolve an ambiguity or narrow down a large set of possible bus stops the spatial inference module can point out what kind of information the user should provide.

9. ACKNOWLEDGMENTS

This work is supported by The Swedish Transport & Communications Research Board (KFB). We are indebted to Oskar Lindwall who collected and transcribed the dialogues.

10. REFERENCES

1. S. Bennacef, H. Bonneau-Maynard, J. L. Gauvin, L. Lamel, and W. Minker. A spoken language system for information retrieval. In *Proceedings of ICLSP'94*, 1994.
2. A.G. Cohn. Qualitative spatial representation and reasoning techniques. In G Brewka, Ch. Habel, and B Nebel, editors, *KI-97 : Advances in artificial intelligence : proceedings of 21st Annual German Conference on Artificial Intelligence*, volume 1303 of *Lecture Notes in Artificial Intelligence*, pages 1–30. Springer-Verlag, 19.
3. Arne Jönsson. A model for habitable and efficient dialogue management for natural language interaction. *Natural Language Engineering*, 3(2/3):103–122, 1997.
4. Arne Jönsson and Nils Dahlbäck. Talking to a computer is not like talking to your best friend. In *Proceedings of the First Scandinavian Conference on Artificial Interlligence, Tromsø*, 1988.
5. Robert Laurini and Derek Thompson. *Fundamentals of Spatial Information Systems.* Academic press, 1992.
6. M. Monmonier. *How to Lie with Maps.* Univ. of Chicago Press, second edition, 1996.
7. Albert Russel, Ingmar Herberts, Els den Os, and Louis Boves. Dialog management issues in the localization of a train time table information system. In *Proceedings of ISSD'96, Philadelphia*, pages 81–84, 1996.
8. Lena Strömbäck and Arne Jönsson. Robust interpretation for spoken dialogue systems. In *Proceedings of ICSLP'98, Sydney, Australia*, 1998.
9. W.P.A. van Deursen. *Geographical Information Systems and Dynamic Models.* PhD thesis, University Utrecht, 1995.
10. Michale F. Worboys. *GIS: A Computing Perspective.* Taylor & Francis, 1997.