Communicating Vague Spatial Concepts in Human-GIS Interactions: A Collaborative Dialogue Approach

Guoray Cai[1,3], Hongmei Wang[1,3], and Alan M. MacEachren[2,3]

[1] School of Information Sciences and Technology, Penn State University, University Park, PA 16802
{*cai, hwang*}@ist.psu.edu
[2] Department of Geography, Penn State University, University Park, PA 16802
maceachren@psu.edu
[3] GeoVISTA Center, Penn State University, University Park, PA 16802

Abstract: Natural language requests involving vague spatial concepts are not easily communicated to a GIS because the meaning of spatial concepts depends largely on the contexts (such as *task*, *spatial contexts*, and *user's personal background*) that may or may not be available or specified in the system. To address such problems, we developed a collaborative dialogue approach that enables the system and the user to construct shared knowledge about relevant contexts. The system is able to anticipate what contextual knowledge must be shared, and to form a plan to exchange contextual information based on the system's belief on who knows what. To account those user contexts that are not easily communicated by language, direct feedback approach is used to refine the system's belief so that the intended meaning is properly grounded. The approach is implemented as a dialogue agent, *GeoDialogue*, and is illustrated through an example dialogue involving the communication of the vague spatial concept *near*.

1 Introduction

Current geographical information systems (GIS) do not support human work effectively because users must interact with geographical data through formally defined textual or graphical query languages using keyboards and mouse. Recent research has paid more attention to human communication modalities as potential alternative modes of human-computer interactions, because it is relatively effortless for people to express their information needs in this way. One component in this line of research has been the use of natural language for submitting request to GIS[1-5]. These works have emphasized the scale-dependent and imprecise nature of natural language in describing spatial relations. However, actual implementations of natural language interfaces for GIS have been very limited (for examples see [4, 6] [7] [8] [9]. To overcome the difficulties of natural language for expressing metric details of spatial relations[10], many authors have explored the possibility of using gestures as more intuitive styles for interacting with spatial data. For example, pen-based gestures were used in *spatial query-by-sketch* [11] and *QuickSet[12]*. Free hand gestures were featured in *GIS Wallboard[13]*, *iMap[14-16]*, and *DAVE_G[17]*.

The main challenge for using human modalities (speech and gesture in particular) to interact with computers is that human communication are highly contextualized by the shared knowledge of the participants. Such contextual knowledge are often not available or properly captured in computer systems, which is the reason that spoken language requests to computer systems often appear to be incomplete, ambiguous, or vague. In this paper, we address the challenge of communicating spatial information requests that involve vague spatial concepts during human-GIS interaction. In particular, we focus on the concept of “near”, which is as a classic example of such concepts[3, 18].

Communicating spatial concepts to a computer system in natural language is especially challenging for a number of reasons. First, geographic concepts do not have unique and precise mapping to quantitative representation of space[10, 19-21]. Second, human uses of spatial language are qualitative and ambiguous, allowing meaning to vary with the contexts of use. In contrast, computer representations of spatial concepts are quantitative and precise[10], leaving no room to adapt to user’s conception of spatial concepts. Currently, formal computational models exist that allow GIS to form a quantitative representation of spatial concepts communicated through natural language. Examples of such computational models include those for distance relations[18, 22], directional relations[23, 24], and topological relations[10, 25-28]. These models, which mostly follow Rosch’s prototypical categorization theory of human cognition[29, 30], have serious limitations when used in human-GIS interaction. Human-GIS interactions have to deal with both the inherent uncertainties and context-varying nature of spatial concepts. Formal models of spatial concepts are useful in capturing the inherent uncertainty of spatial objects and spatial relations, but do not have the properties of human-human communication with contextualized interpretation and negotiation. Even formal models of common sense geography [31] in natural language do not have the necessary formalisms and parameters to incorporate contextual information that are crucial for communicating spatial concepts in human-GIS interaction.

Human-GIS interactions are usually situated in a complex set of contexts including the geographic site and situation, the task or problem being addressed, and the background (knowledge, experiences, and preferences) of the individual users. The spatial concepts communicated need to be interpreted within such relevant contexts[3, 19, 21, 32-37]. However, computer systems usually ignore the contexts when translating linguistic reference of spatial concepts into formal representations of meanings. This constitutes a reason for the system to misinterpret the user’s request and to cause breakdowns in human-GIS interaction[19].

The *collaborative dialogue approach*, as proposed in this paper, follows the principles of human-human communication in dealing with vague spatial concepts. Human-human conversation is grounded in a significant amount of knowledge (or context) shared among participants, which allows the hearer to disambiguate vague concepts. However, adequate sharing of relevant context may not always be established before a vague concept is communicated, causing difficulties for the hearer to construct an unambiguous mental model of that concept that matches the meaning intended by the speaker. In such situations, it is common that human will engage in dialogues that to acquire the meaning of the concept involved[22]. Dialogues as a general human communication strategies serve a diverse functions in communicating and sharing knowledge as well as coordinating actions. For the particular situations of communicating vague spatial concepts, dialogues make two kinds of collaboration possible: (1) soliciting and sharing additional contextual information, and (2) seeking feedback and confirmation. Applying such principles to human-GIS interactions, the *collaborative dialogue approach* addresses a number of computational issues, such as how to represent the belief states of context sharing between the user and the system, how to relate newly communicated knowledge with what has been shared so far, and how to build system intelligence for generating dialogues consistent with the goals of advancing context sharing. It is also about dialogue strategies for grounding linguistic message in a shared meaning reliably through confirmation and visual feedback. The user and the system is brought into collaboration[38], and follows the metal state view of human-human communication and collaborative planning[39, 40]. No other existing natural language based GIS [4, 17, 41, 42] have the abilities to initiate this kind of dialogue.

In order to test the feasibility of our collaborative dialogue approach, we implemented a prototype system, *GeoDialogue*, which is a conversational agent that is able to engage in collaborative dialogue with the user. When a vague spatial concept is detected in a dialogue, GeoDialogue is able to invoke proper schemas (stored in its knowledge-base) about what contextual information needs to be shared. Such schemas are central organizing structures that link contextual knowledge to meanings of spatial concepts and guide the planning of communicative actions in exchanging contextual knowledge. As the dialogue proceeds, the user and the system will mutually adapt to each other and reconstruct meaning representation that reflects the shared contexts about tasks, spatial and geographical environment, and personal backgrounds. A common understanding of a spatial concept is collaboratively constructed and visually shared, resulting in a meaning that is neither what the user initially perceives nor what the system initially estimates. Currently, the collaborative dialogue approach and the *GeoDialogue* agent are already integrated as part of the core reasoning engine of the Dialogue-Assisted Visual Environment for Geoinformation (DAVE_G) [43]which is a generic environment in which the user can interact with the GIS through multimodal interactions.

The rest of the paper is organized as follow. In section 2, we discuss major factors affecting human-computer communication of spatial concepts. In section 3 we will describe our methodology, which incorporates a collaborative dialogue approach that captures, represents and uses the contextual information for understanding to a vague spatial concept. The section 4 will introduce implementation of this approach in the prototype system, the GeoDialogue agent. In the section 5, a sample dialogue is presented to illustrate the collaborative dialogue approach in GeoDialogue. Finally, we conclude with discussion on the implications of this development to the future natural language interfaces to GIS.

2 Factors Affecting the Communication of Spatial Concepts

Natural language terms may carry spatial concepts that refer to spatial objects or spatial relationships. When human communicate a spatial concept to a GIS through linguistic terms, there are two types of factors that present problems for a successful communication. First, the spatial concept itself may be inherently vague[44, 45]. Second, the meaning of a spatial concept may be dependent on a large number of contexts within which the concept is used. For human-GIS interactions, major parts of such contexts are spatial contexts (including geographical space [46] and environmental space[32, 36]), task contexts[10], and personal contexts of the individual users.

2.1 Inherent Uncertainties of Spatial Objects and Spatial Relations

There are two types of inherent uncertainties in the processing of spatial concepts[44, 45]. First, a spatial concept may refer to the spatial objects that have indeterminate boundaries. For example, mountainous areas are not easily determined because the boundary between a "mountainous area" and the surrounding areas is inherently uncertain. The second type of uncertainties is the fuzzy nature of spatial relations. For example, the concept '*near*' has long been considered to be an ill-defined[47]. It exhibits classic characteristics of a radial category, and it includes a range from exemplars to borderline cases[21, 36]. As an example, consider a request for a map showing 'hotels near *Miami*'. A user may consider hotels within 20 miles to *Miami* to be *near*, and those beyond 60 miles from *Miami* as *far*. However, a hotel that is 40 miles away from *Miami* may be considered neither near nor far from *Miami*.

2.2. Contextual Factors for Interpreting Spatial Concepts

Contextual factors about a particular use of a spatial concept refer to the knowledge that human uses to constrain the meaning of communication. To reach a common understanding of a vague concept, e. g. *near*, the system and the user need to share knowledge about the relevant contexts that affect the understanding of the vague concept. Among many potential contextual factors that may affect how people understand spatial concepts, we focus here on three of them: task, spatial contexts, and background of the user.

When a user requests a map to be displayed by a GIS, it is often because that the user is trying to perform a domain task that has some information needs. The task becomes an important part of the use context for spatial concepts (similar view was expressed in[10]). For example, the same request "show me a map *near* Miami" may be made by a person-*A* who is in a task situation of selecting a grocery store, and by a person-*B* who is planning a vacation for the weekend. However, person-*B* is likely to expect a map showing a larger geographical area comparing with person-*A*.

There are evidences that the meaning of spatial concepts, such as "*near*", is also dependent on the spatial context[18, 32, 36]. As detailed by Montello[48], human cognition supports multiple and qualitatively different conceptualizations of space. Two of these, which he labels environmental space and geographical space, are relevant here. Environmental space is the behavioral-scale space of physical entities, such as cities and buildings, in the real world. Worboy's [36, 37] and Fisher [32] provided evidences that environmental space affects people's interpretation of *near*. Geographical space is usually learned via symbolic representations, such as a map. It can also affect the understanding of spatial concepts, such as the distance relationship of *near[18, 22, 46]*. Since a map conveys information about geographical space in human-GIS interactions, properties of the map, such as the map scale[18, 22], the contents, and the distribution of objects in the map, affect people's perception on the relationship of *near*.

Individual users may differ in cultural and education background, personal experiences and preferences that affect how they form mental models geographical phenomena and assign meanings to vague concepts, such that they may perceive the same concept differently[36, 37]. In his experiments on the individual differences on the meaning of spatial relationships, Robinson [18] demonstrated that there are significant semantic variations among individuals even for relatively simple geographical datasets. Even in the same spatial context and task context, different individuals are often found to be different in determining whether specific locations are near or not.

3. Methodology

In this section, a collaborative dialogue approach is proposed as a more effective way to facilitate communication of vague spatial concepts in human-GIS interactions. In a collaborative dialogue system, the system and the user can construct and negotiate (perhaps through multiple exchanges) a shared understanding of the meaning of a vague concept by exchanging contextual information. The vague spatial concept *near* is used here to illustrate our approach.

In our approach, the process of communicating the meaning of vague spatial concepts between the system and the user is modeled as a collaborative discourse, following the theory of the collaborative discourse developed by Grosz and Sidner[49]. The process of communicating vague spatial concepts has all the central elements of a typical collaborative discourse: (1) it involves a shared goal among participants (which is to achieve common understanding of a vague spatial concept); (2) it requires a shared pool of knowledge that is initially distributed among participants; and (3) it uses communicative actions as the way to achieve collaboration. According to Grosz and Sidner[49], a collaborative discourse has three interrelated structures: a linguistic structure, an intentional structure and an attention state. The theory believes that there is always an intention behind any communicated messages (utterances), and they functionally contribute to the achievement of larger intentions of the dialogues within which messages are organized. The key to understand collaborative dialogue is to have a proper model for the intentional structure of the collaboration. A computational model, the recipe graph or *Rgraph*, was developed by Lochbaum [40, 50] to represent the intentional structure of a collaborative discourse. It is based on the *SharedPlan* theory[39, 51], which was developed to model mental attitudes that a group of agents must hold during a collaborative planning process.

The collaborative discourse theory provides the foundation for modeling any extended dialogues in which the participants of the dialogue have a common goal to achieve but each of the participants takes on different roles in their collaboration. In the collaboration process of a dialogue involving a vague spatial concept, participants exchange information about relevant contexts so that part of individual belief and knowledge become shared and integrated into a consistent set of knowledge held by all the participants. As a dialogue proceeds, the system uses an Rgraph [40, 50] to keep track of the dynamically changing set of knowledge and beliefs communicated and shared during a dialogue process. The *Rgraph* consists of actions and the belief of the agents being modeled. Each node in the *Rgraph* is a complex data structure which records an intended action together with a set of beliefs that the system and the user have towards that action. An action can be either basic (an action that can be executed directly) or complex (an action that requires further communication and coordination with the user before it can be executed). For each complex action, the system chooses a recipe from the recipe library maintained in its knowledge-base. A recipe specifies a way of achieving an action, and it usually includes knowledge preconditions (represented as *parameters*), subactions, and other constraints necessary for executing the recipe. The belief status of an agent on an action (or a parameter) represents the agent's belief on the status of collaboration on the action (or the parameter). An example of an Rgraph is shown in Figure 1. The belief status on actions and parameters and their corresponding meanings are described in Table 1.

,

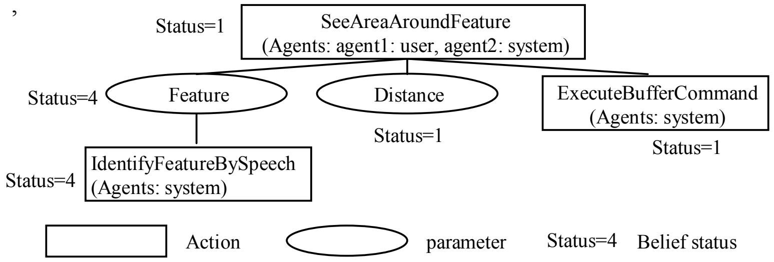

Fig. 1 Structure of Rgraph

With the representation of the dialogue status as an Rgraph, the effect of dialogue interactions will be the evolution of the *Rgraph* from the one that was only partially instantiated towards the one that is fully elaborated. The evolution begins with adding an action for achieving the common goal to the *Rgraph.* This action is the root action of the Rgraph. Then the system will understand and respond to the subsequent messaged based on the knowledge encoded in the Rgraph. Four steps of reasoning are conducted for each new input from the user: recognizing the intention of the utterance, explaining the intention by associating it with the Rgraph, elaborating the modified Rgraph by volunteering additional knowledge or performing executable parts of the Rgraph, and inferring the goal of generating questions or responses. If necessary, the system will call a GIS component to execute basic actions related to GIS commands and generate map responses, such as actions for doing buffering, displaying layers, and doing selections. By maintaining an Rgraph representation of a collaborative discourse, the system decides the process flow of its collaborative behavior (interpretation, elaboration, execution, response, and request,) by reasoning based on the Rgraph.

Table 1 Belief status numbers and their meanings

Status	Meaning on an action	Meaning on a parameter
1	The agent being modeled believes that the action has not been done.	The agent being modeled believes that a value for the parameter can be identified from the current discourse context
2	The agent being modeled believes that the result of the action is needed to be negotiated with the user.	The agent being modeled believes that the value of the parameter needs to be negotiated with the user.
3	The agent being modeled believes that the action fails.	The agent being modeled believes that instantiation of the parameter fails.
4	The agent being modeled believes that the action was successfully executed.	The agent being modeled believes that instantiation of the parameter succeeds.

Now we describe how the collaborative dialogue approach works for human-GIS communication on the meaning of spatial concepts. At the start of the process for communicating a vague spatial concept, the user and the system share the common goal that the system needs to show a map result matching the user's request, that is, the system needs to have a common understanding to the vague spatial concept with the user. However, there may be different initial sets of contextual knowledge and individual beliefs about the spatial concept that were available to the system and the user. The user is usually more knowledgeable about specifying the task and has a set of personal background towards the perception of spatial concepts that may not be available to the system. On the other hand, the system generally has more detailed knowledge on the environmental and geographical context due to its storage of spatial data, which may not be available in the memory of the human user. As the dialogue proceeds, both sides seek to maximize the shared knowledge by keeping track of what knowledge are already shared and what are to be further communicated. As each side shares more contextual knowledge with the other, they modify their beliefs on the correct understanding to the spatial concept, until finally they come to a common understanding to it.

As discussed in section 2, a successful communication of a spatial concept between a system and a user requires the sharing of three categories of contextual information. In our collaborative dialogue approach, each property of the contextual information is represented as an optional parameter of a recipe that represents a context-sensitive model of the spatial concept. For example, when the system receives a request (from a user) for a map *near* a place, it immediately realizes that an action *SeeAreaNearFeature* is to be planned. Suppose that the system adopts a "buffer zone" model of *near,* then a recipe as in figure 4 will be activated. Since the system understands that identifying the parameter *Distance* is a pre-condition for executing the action *SeeAreaNearFeature*, the system will search for a recipe to identify this parameter. Suppose that the system knows two strategies to acquire this parameter: (1) by making an arbitrary guess and ask the user to correct it though visual interaction (this is labeled as the '*direct approach*'), or (2) by pursuing sharing of contextual knowledge (this is labeled as the '*indirect approach*'). The system may choose to pursue either approach at any given time of the interaction, depending on the system's belief on which one is more effective. Suppose that the second (indirect approach) is applied initially. A recipe *IndentifyDistanceFromContext* (see the lower part of Figure 5) is instantiated, which includes parameters representing contextual knowledge such as map scale, the goal of the task, the vehicle (if any) that the user will use in the task, and the user's personal preferences on the *near* relationship. These parameters are optional to execute *IndentifyDistanceFromNear*. When a contextual parameter, such as task, is missing, the system realizes that there are opportunities to improve the system's understanding of near if such contextual information can be provided by the user. Hence, the system will initiate a subdialogue to request it from the user. If the user responded with the requested information, the system will re-estimate a buffer distance and generate a map showing a revised understanding of *near*. Since not all the contextual factors have equal influence on the model of *near*, our algorithm uses task knowledge as the primary determinant, and others factors (such as spatial contexts and personal preferences) as subsequent modifiers.

4 Implementation of the Approach

We implemented the collaborative dialogue approach as part of a research prototype, *GeoDialogue*, a natural language human-GIS dialogue manager. The architecture of GeoDialogue is shown in Fig. 2. It includes modules for Syntactic Parsing, Semantic Interpretation, Dialogue Control, Query formation and Map generation, and Response control. The Dialogue control module is the central intelligence module that maintains dynamic knowledge and dialogue context, we well as performing automated reasoning on collaborative plans. All modules have access to a Static Knowledge Base (Static KB) which provides knowledge about language structure (grammar and terms), task advancement, and information content of GIS databases.

The dialogue usually begins with a request from the user to the system. The user's natural language request is first analyzed by the Syntactic Parsing module so that words are grouped into meaningful phrases. The parsing result is sent to the Semantic Interpretation module to extract intended actions and associated parameters and constraints. If sufficient information has been accumulated for the system to issue a well-formed GIS request, then the job is continued by the Query Formation & Map Generation module, which translate active request into a GIS query and coordinate the execution of this query through a standard GIS query interface. In our current implementation, an ArcIMS service is used as the GIS component.

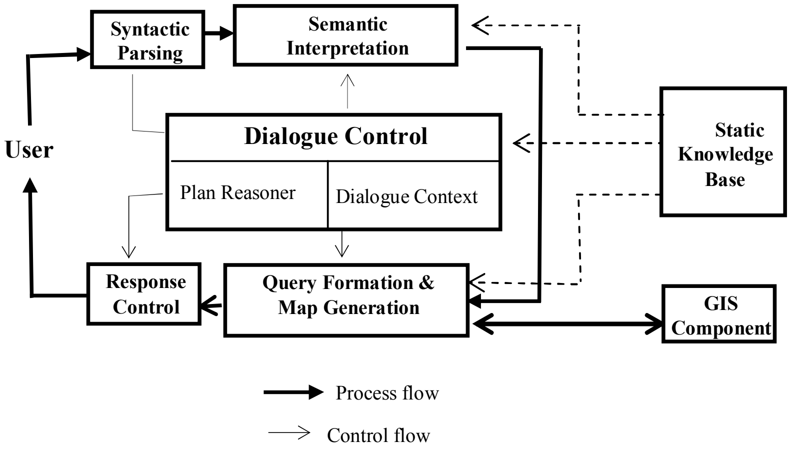

Fig. 2 **Architecture of GeoDialogue**

The Response Control module assembles response messages and coordinates the presentation of response to the user. A response message may include several components, such as a map generated by a GIS, summary status, reports of problems, and requests for further information, or clarification questions. To make the response as natural and effective as human communications, dialogue controller implements dialogue planning and response strategies as part of its *plan reasoner* functions.

The knowledge about understanding natural language words/phrases, actions for achieving goals and relevant datasets needed in the collaboration are stored in the Static Knowledge Base module. In GeoDialogue, the Static KB is implemented as a relational database. The Dynamic Knowledge module stores all information involving in the dynamic discourse context and a dynamic map context. The discourse context is updated dynamically as the dialogue proceeds and the map context is updated as a new map response is generated by the system. As described in the third section, the collaborative discourse context is modeled as an Rgraph.

5 Example User-System Dialogue Involving *Near*

In this section, a sample dialogue is given to illustrate how the collaborative dialogue approach is used in human-GIS communications involving a vague concept *near*. Within this interaction scenario (Fig. 3), the system knows extensive information about the geographical space, which can be shared with the user through an interactive map. The user has detailed information about the task, which can be communicated to the system through natural language input. Personal background and preferences are hard to be articulated in language, and these factors are taken into account using the direct approach (discussed in section 3). The combination of these dialogue strategies ensures that a shared understanding of the vague spatial concept ‘near’ is grounded in the shared contextual knowledge, and the process of such communication is more efficient than using only one of these strategies. For the purpose of convenience for description, we assume the user to be female.

The example dialogue begins with the user’s request U1, which involves a vague concept ‘near’. The system interprets the user’s intention of U1 as “to see a map covering an area near some features”, and forms a plan of ‘show map’ action using *SeeAreaAroundFeature* as the recipe. So, an *Rgraph* is initiated with the action *SeeAreaAroundFeaure* as its root action after the interpretation of U1 and the current focus in the Rgraph is the root action. The system elaborates this action with its recipe, *SeeAreaAroundFeature*, and updates the *Rgraph* (Fig. 4)

Note: The dialogue begins with an empty map shown on the screen.
(U- the user; G-the system, GeoDialogue)

(**U1**): I would like to see area near *Miami*.
(**G1**): (show a map) What are you going to do with this map?
(**U2**): Oh, I am just planning a trip to *Miami* and will have fun there by car.
(**G2**): (show a new map) Here it is. Is this what you want?
(**U3**): Yes, but I want to see a wider area near *Miami*.
(**G3**): (shows a new map) Is this what you want?
(**U4**): wider.
(**G4**): (shows a new map) Is this what you want?
(**U5**): The previous one is fine.
(**G5**): (Show the previous map) Here it is.

Fig. 3 An example dialogue involving near between the system and the user

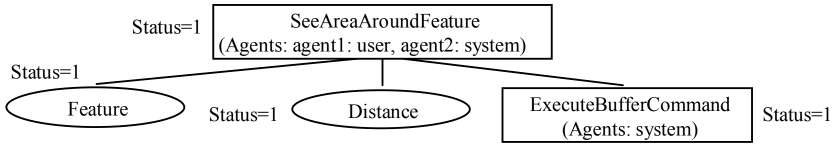

Fig. 4 Initial Rgraph

Based on the understanding of input U1, the system starts to reason about what actions should be taken next. By inspecting the Rgraph, the system found that there are two knowledge pre-conditions, ‘Feature’ and ‘Distance’, that need to be satisfied before the ‘buffer zone’ model’ of the near concept can be calculated. Since *Miami* is a feature mentioned in the user’s input, it is used to instantiate the parameter *Feature* (through a basic action *IdentifyFeatureBySpeech*). However, there is no clue from the user on how to determine the second parameter ‘Distance’. Assuming that the system has a number of stored model of ‘near’ that could potentially be used to determine the buffer distance, but it need help from the user to select an appropriate model based on the use contexts. For this reason, the system chooses to use the ‘indirect approach’ by instantiating a recipe, *IdentifyDistanceFromContext*, pursue the goal of context sharing. The recipe *IdentifyDistanceFromContext* is added to the *Rgraph* by attaching it to the parameter *Distance* node. The action *IdentifyDistanceFromContext* contains four (4) optional parameters, *Map_Scale*, *Task_Goal*, *Task_Vehicle* and *User_Preference* (For simplicity in describing the approach, we implement only four parameters for the three categories of contextual information here). While reasoning about a way to get knowledge on the 'task_goal' parameter, the system believes that the user has the knowledge about the intended task, and subsequently generates a question (G1 in Figure 3). In the same time, the system will generate a distance estimate (by executing the basic action *ExecuteEstimationOnNear*) using generic model (such as the Fuzzy Logic model). Although such an estimate has little chance to match with the user's intended meaning, it at least allows the system to generate a visual response. The Rgraph at this moment is shown in Fig. 5.

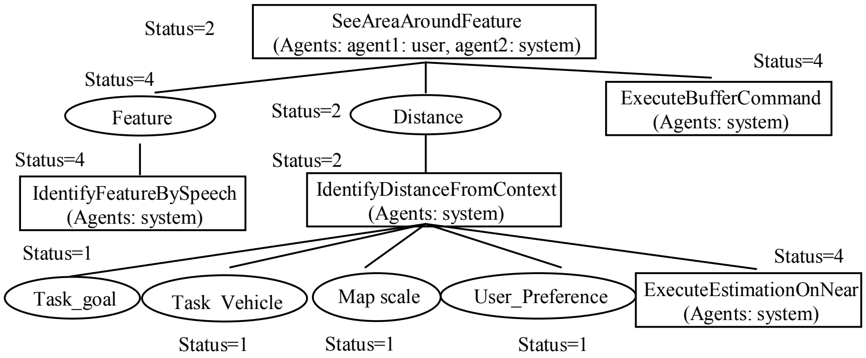

Figure 5 A snapshot of the Rgraph for the example dialogue (at G1).

The map response in the (G1) of Figure 3 serves two purposes. First, the map provides a shared geographical context (e.g, the distribution and sizes of features on the map) for the user to reflect on his understanding of 'near'. Second, it serves as a representation of how the system currently understands 'near'. In the same time, the system also realizes (by consulting the Rgraph in Figure 5) that the first parameter *Task_Goal* information is missing, which means that task information (that may be known to the user) has not been shared with the system. So it generates a question to the user in G1 in order to elicit the task information from the user.

After the user receives the (G1) response, she replies U2, which provides information not only about the goal of the task but also the information about the vehicle that will be used in the task. U2 provides information to instantiate two parameters: *Task_Goal* and *Task_Vehicle*. Additionally, the system can get the map scale information from the current map context. After acquiring more contextual information such as the task ("having fun by car") and the current map context, the system modifies its belief on the buffer distance (such as '2 hours of driving distance', or '100 miles') as its new understanding of *near*.

Although important contextual information about the tasks and spatial contexts are already communicated and shared at this time, the system and the user may still differ in their understanding due to the user's specific characteristics. The user provides her own preferences in U3 to answer C2: "Yes, but I want to see a wider area near

Miami." From the system's Static KB, the system knows that "wider" is a vague concept and it, in the current discourse context, represents longer distance. The system modifies the *Distance* value to a larger value based on the value of *Distance* estimated in the last time through a formal method and generates a new map response and a question response G3 to the user.

After the dialogue exchange at G3, the user still wants a wider area. Her reply in U4 led the system to generate a question in G4 and a new map with a larger value of the parameter *distance* than the previous one (with G3 together). With sharing the map context shown by the system, the user finally changes his previous understanding (in U4) of *near* and thinks the map sent with G3 is fine. Based on his reply in G5, the system generates a map by using the same distance used in the previous map (sent with G3). Finally, the system and the user come to the common understanding of *near* based on their shared knowledge of the relevant context through a collaborative dialogue process.

6. Discussions and Conclusions

In this paper, we have developed a collaborative dialogue approach to facilitate communication of vague spatial concepts in Human-GIS interaction. This approach allows contextual information to be captured, represented and shared through mixed-initiative dialogues. The dialogue between the system and the user is modeled according to the collaborative discourse theory developed by Grosz and Sidner [49]. For the vague spatial concept *near*, the context factors that we consider in this paper include the task information, the geographical information, and the user's individual background and preferences. The system contributes more knowledge on generating a shared map context, and the user contributes more on providing the task information and the individual information.

The distinctive feature of our approach is that it is based on a well-received theory, the *SharedPlan* theory of multi-agent collaboration [51]. The theory models the mental attitudes for an agent to participate in a collaborative discourse. By actively modeling the intentional structure of the dialogue as it progresses, the system is able to adapt to users with different level of contextual knowledge. The system participates in the dialogue by interleaving the actions of understanding the user's request, finding a plan to achieve it (which may include collaboration with the user), executing the plan, and generating responses to the user.

Supporting communications of vague spatial concepts between a user and the system is the key to enable natural language based GIS. Towards this goal, the collaborative dialogue approach, as proposed in this paper, is not an alternative to the existing formal models of spatial concepts. Instead, the two approaches are complementary to each other. In particular, our approach emphasize the role of shared contextual knowledge in communicating a vague concept, while formal models of spatial concepts captures the inherent uncertainties of spatial objects and relationships themselves. Ideally, the system should be able to adapt its behavior in a situated interaction that may range from no contextual information at all to the sharing of full knowledge about the contexts.

Our experience with developing collaborative dialogue interfaces [17] for GIS indicates that there are unexplored potentials to incorporate the principles of human-human communication into interfaces for geographical information systems. As exemplified in this paper, dialogues and collaborations are fundamental strategies for human to communicate vague (spatial) concepts. When people's communication involves a vague spatial concept, people usually exchange relevant knowledge about the context to come to a common understanding of the concept. Further research is needed to understand how these multiple contextual factors interact in affecting the meaning of vague spatial concepts, and to develop computational models that are more cognitively plausible.

Acknowledgement

We would like to acknowledge the funding and support from the National Science Foundation under Grant No.BCS-0113030. This work has benefited from discussion with Sven Fuhrmann and Isaac Brewer and has received programming assistance from Ping Xia and Levent Bolelli.

References

1. Frank, A. U. and Mark, D. M.: Language issues for GIS. In: Macguire, D., Goodchild, M. F. and Rhind, D.(eds): *Geographical Information Systems: Principles and Applications*, Wiley, New York (1991) 147-163
2. Mark, D. M. and Frank, A. U.: User interfaces for Geographic Information Systems: report on the specialist meeting. National Center for Geographic Information and Analysis (1992)
3. Mark, D. M., Svorou, S. and Zubin, D.: Spatial terms and spatial concepts: Geographic, cognitive, and linguistic perspectives. *Proceedings of International Geographic Information Systems (IGIS) Symposium: The Research Adgenda:*, Arlington, VA. (1987) 101-112
4. Shapiro, S. C., Chalupski, H. and Chou, H. C.: Linking ARC/INFO with SNACTor. Santa Barbara, California.: National Center for geographic Information and Analysis (1991)
5. Neal, J. G. and Shapiro, S. C.: Intelligent Multi-Media Interface Technology. In: Sullivan, J. W. and Tyler, S. W.(eds): *Architectures for Intelligent Interfaces: Elements and Prototypes*, Addison-Wesley, Massachusetts (1991) 11-44
6. Zue, V., Glass, J., Goddeau, D., Goodine, D., Leung, H., McCandless, M., Phillips, M., Polfroni, J., Seneff, S. and Whitney, D.: Recent progress on the MIT voyager spoken language system. *Proceedings of the ICSLP*, (1990) 1317-1320

7. Lokuge, I. and Ishizaki, S.: Geospace: An interactive visualization system for exploring complex information spaces. *CHI '95 Conference Proceedings*, New York, 1995 (1995) 409-414
8. Wang, F.: Towards a natural language user interface: An approach of fuzzy query. *International Journal of Geographical Information Systems* 2 (1994) 143-162
9. Wang, F.: Handling Grammatical Errors, Ambiguity and Impreciseness in GIS Natural Language Queries. *Transactions in GIS* 1 (2003) 103-121
10. Egenhofer, M. J. and Shariff, A. R.: Metric details for natural-language spatial relations. *ACM Transactions on Information Systems* 4 (1998) 295 - 321
11. Egenhofer, M. J.: Query Processing in Spatial-Query-by-Sketch. *Journal of Visual Languages and Computing* 2 (1997) 403-424
12. Cohen, P. R., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L. and Clow, J.: QuickSet: multimodal interaction for distributed applications. *Proceedings of the fifth ACM international conference on Multimedia*, Seattle, Washington (1997) 31 - 40
13. Florence, J., Hornsby, K. and Egenhofer, M. J.: The GIS wallboard: interactions with spatial information on large-scale displays. In: Molenaar, M.(eds): *International Symposium on Spatial Data Handling*, 7. Taylor and Francis, Delft, The Netherlands (1996) 449-463
14. Sharma, R., Pavlovic, V. I. and Huang, T. S.: Toward a multimodal human computer interface. *Proceedings of the IEEE* 5 (1998) 853-69
15. Sharma, R., Poddar, I., Ozyildiz, E., Kettebekov, S., Kim, H. and T.S. Huang: Toward Interpretation of Natural Speech/Gesture: Spatial Planning on a Virtual Map. *Proceedings of ARL Advanced Displays Annual Symposium*, Adelphi, MD (1999) 35-39
16. Kettebekov, S., Krahnstöver, N., Leas, M., Polat, E., Raju, H., Schapira, E. and Sharma, R.: i2Map: Crisis Management using a Multimodal Interface. *ARL Federate Laboratory 4th Annual Symposium*, College Park, MD (2000)
17. Rauschert, I., Agrawal, P., Fuhrmann, S., Brewer, I., Wang, H., Sharma, R., Cai, G. and MacEachren, A.: Designing a User-Centered, Multimodal GIS Interface to Support Emergency Management. *ACM International Symposium on Advances in Geographical Information Systems*, (2002)
18. Robinson, V. B.: Individual and multipersonal fuzzy spatial relations acquired using human-machine interaction. *Fuzzy Sets and Systems* 1 (2000) 133 - 145
19. Duckham, M., Mason, K., Stell, J. and Worboys, M. F.: A formal approach to imperfection in geographic information. *Computers, Environments and Urban Systems* (2001) 89-103
20. Goodchild, M. F.: Sharing imperfect data. In: Onsrud, H. J. and Rushton, G.(eds): *Sharing geographic information*, Rutgers, New Jersey (1995) 413-425
21. Fisher, P.: Sorites paradox and vague geographies. *Fuzzy Sets and Systems* (2000) 7–18
22. Robinson, V. B.: Interactive Machine Acquisition of a Fuzzy Spatial Relation. *Computers and Geosciences* 6 (1990) 857-872

23. Peuquet, D. J. and Zhan, C.-X.: An algorithm to determine the directional relationship between arbitrarily-shaped polygons in the plane. *Pattern Recognition* (1987) 65-74
24. Matsakis, P., Keller, J. M., Wendling, L., Marjamaa, J. and Sjahputera, O.: Linguistic Description of Relative Positions in Images. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS* 4 (2001) 573-588
25. Egenhofer, M. J. and Franzosa, R.: Point-set topological spatial relations. *International Journal of Geographical Information Systems* 2 (1991) 161-174
26. Zhan, F. B. and Lin, H.: Overlay of Two Simple Polygons with Indeterminate Boundaries. *Transactions in GIS* 1 (2003) 67-81
27. Hazelton, N. W., Bennett, L. and Masel, J.: Topological structures for 4-dimensional geographic information systems. *Computers, Environment, and Urban Systems* 3 (1992) 227-237
28. Zhan, F. B.: Topological Relations Between Fuzzy Regions. *Proceedings of the 1997 ACM Symposium on Applied Computing*, San Jose, CA (1997) 192-6
29. Rosch, E.: Principles of Categorization. In: Rosch, E. and Lloyd, B.(eds): *Cognition and Categorization*, Erlbaum, Hillsdale, NJ (1978)
30. Rosch, E.: On the Internal Structure of Perceptual and Semantic Categories. In: Moore, T.(eds): *Cognitive Development and the Acquisition of Language*, Academic Press, New York, NY (1973) 111–144
31. Egenhofer, M. J.: User Interfaces. In: Nyerges, T., Mark, D. M., Laurini, R. and Egenhofer, M. J.(eds): *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems*, Kluwer Academic Publishers, Dordrecht, The Netherlands (1995) 143-145
32. Fisher, P. and Orf, T.: An investigation of the meaning of near and close on a university campus. *Computers, Environment and Urban Systems* (1991) 23–25
33. Mark, D. M.: Spatial Representation: A Cognitive View. In: Maguire, D. J., Goodchild, M. F., Rhind, D. W. and Longley, P.(eds): *Geographical Information Systems: Principles and Applications*, 1. John Wiley & Sons, New York (1999) 81-89
34. Mark, D. M. and Frank, A. U.: NCGIA Initiative 2, "Languages of Spatial Relations," Closing Report., Santa Barbara, CA: National Center for Geographic Information and Analysis (1992)
35. Landau, B. and Jackendoff, R.: "What" and "where" in spatial language and spatial congnition. *Behavioral and brain sciences* (1993) 217-265
36. Worboys, M. F.: Nearness relations in environmental space. *International Journal of Geographical Information Science* 7 (2001) 633–651
37. Worboys, M. F.: Communicating geographic information in context. *Meeting on Fundamental Questions in GIScience*, Manchester (2001)
38. Terveen, L. G.: Overview of human-computer collaboration. *Knowledge-Based Systems* 2-3 (1995) 67-81
39. Grosz, B. J. and Kraus, S.: Collaborative plans for complex group action. *Artificial Intelligence* 2 (1996) 269-357

40. Lochbaum, K. E.: A collaborative planning model of intentional structure. *Computational Linguistics* 4 (1998) 525-572
41. Oviatt, S. L.: Multimodal interfaces for dynamic interactive maps. *Proceedings of the Conference on Human Factors in Computing Systems (CHI'96)*, (1996) 95-102
42. Oviatt, S. L. and Cohen, P.: Multimodal interfaces that process what comes naturally. *Communications of the ACM* 3 (2000) 45-53
43. Rauschert, I., Sharma, R., Fuhrmann, S., Brewer, I. and MacEachren, A.: Approaching a New Multimodal GIS-Interface. *Proceeding of the 2nd International Conference on GIS (GIScience)*, CO, USA (2002)
44. Zhan, F. B.: A Fuzzy Set Model of Approximate Linguistic Terms in Descriptions of Binary Topological Relations between Simple Regions. In: Matsakis, P. and Sztandera, L. M.(eds): *Applying Soft Computing in Defining Spatial Relations*, Physica-Verlag, Heidelberg, Germany (2002) 179-202
45. Robinson, V. B. and Frank, A.: About different kinds of uncertainty in collections of spatial data. *Proceedings of 7th International Symposium on Computer-Assisted Cartography*, (1985)
46. Lloyd, R. and Heivly, C.: Systematic distortions in urban cognitive maps. *Annals of the Association of American Geographers* 2 (1987) 191-207
47. Denofsky, M.: How near is near? , Cambridge, MA: MIT AI Lab (1976)
48. Montello, D.: Scale and multiple psychologies of space. *Spatial Information T heory: A Theoretical Basis for GIS (COSIT'93)*, Marciana Marina, Elba Island, Italy (1993)
49. Grosz, B. J. and Sidner, C. L.: Attention, intentions, and the structure of discourse. *Computational Linguistics* 3 (1986) 175--204
50. Lochbaum, K. E.: Using Collaborative Plans to Model the Intentional Structure of Discourse. *Devision of Applied Sceince*, Cambridge: Harvard Univsersity (1994) 158
51. Grosz, B. J. and Kraus, S.: Collaborative plans for group activities. *Proceedings IJCAI 93*, Chambery (1993) 367-373