Toward a Tutorial Dialogue System for Urban Navigation

Christopher Fox and Kristy Elizabeth Boyer
Department of Computer Science
North Carolina State University
Raleigh, North Carolina, USA
{cdfox,keboyer}@ncsu.edu

GPS navigation systems are tremendously beneficial for drivers, affording their users the ability to navigate to any destination regardless of their prior knowledge of the roads. However, as previous studies have shown, these systems can inhibit the formation of a *cognitive map* of a driver's local area (Jackson, 1996; Burnett and Lee, 2005). This lack of development of navigational skills and knowledge poses difficulties in situations where navigation systems fail, such as when network connectivity is lost, a battery is drained, or there is not enough time to input a destination into the system. Burnett and Lee (2005) called for a new "learning-oriented" user interface design for navigation systems, and Oliver and Burnett (2008) later observed that adding landmarks and traces of previous trips to the visual map display encouraged cognitive map development. We hypothesize that a learning-oriented navigation system based on dialogue and long-range navigation instructions will be equally or more effective, while providing a more natural interaction modality that does not require visual attention to the navigation interface. In the system we envision, the next instruction offered would be based on a model of the user's navigational knowledge, which would be estimated from observations of past navigation sessions and from dialogue with the user. Such a system may be thought of as an intelligent tutoring system for urban navigation. Following the "scaffolding" paradigm from education research, the strategy of the tutor would be to gradually reduce the level of assistance along frequent routes until the user achieves mastery (Wood et al., 1976). This abstract describes the planned navigation tutor in more detail, as well as results from a preliminary experiment in urban navigation.

The goal of the tutor is to facilitate the user's development of a cognitive map of his or her local area, while also providing assistance for unfamiliar routes. Jackson (1996) found that, for study participants watching a video of a particular route taken from the driver's perspective, the introduction of a narrator reading turn-by-turn instructions caused one group of participants to remember the details of the route less well than another group that had no narrator. Burnett and Lee (2005) made similar observations of users who were asked to complete several routes in a driving simulator, where one group of users was given turn-by-turn guidance as they drove and the other was not as they drove. Based on these results, it appears that drivers are better able to form a cognitive map of an area if assistance is limited. At the same time, if a route is not known to the driver, then he or she will need detailed instructions on at least the first occasion. We propose, therefore, that the tutor will adapt the granularity of the instructions to the user's cognitive map. A fine-grained instruction specifies the next turn only, as in current navigation systems, while a coarse-grained instruction specifies an intersection or landmark that requires multiple turns to reach from the driver's current location.

In order to make appropriate decisions about the next instruction, the tutor will engage the user in dialogue to determine if portions of the user's existing cognitive map are relevant to the route. For example, after the user has input the destination, the system might give the prompt, "Do you know how to get part of the way there?" If the user responds by specifying a landmark he or she knows how to reach, the system might follow up by asking, "Could you tell me how to get there?" The problem then is understanding the user's spoken route instructions, which previous studies have addressed successfully in a limited domain (Johansson et al., 2011; Meena et al., 2012; Meena et al., 2013). Once the user's description of the route is understood and checked for correctness, the system can update its own model of the user's knowledge.

Another important question to consider is the choice of representation for navigational knowledge. We propose to use a modified version of the *conceptual route graph*, in which there is a node for every place where a turn could occur along the route, rather than only at places where turns do occur (Müller et al., 2000; Johansson et al., 2011; Meena et al., 2012; Meena et al., 2013). The rationale for this modification is that the user's knowledge of the segments (triples consisting of an edge, an end node, and a turning action to perform) of one route should transfer to overlapping segments of a different route. Figure 1 shows a part of the representation for two such overlapping routes.

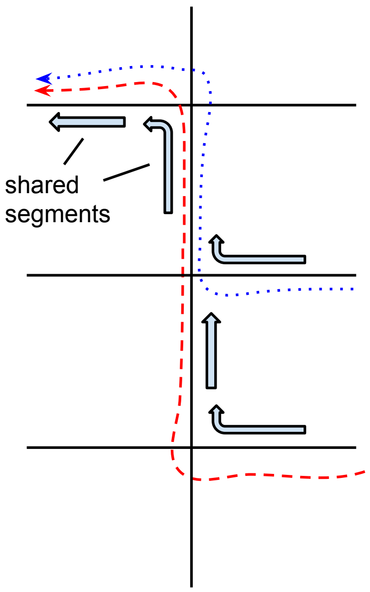

Figure 1: Modified route graph showing two overlapping routes. The routes enter the road drawn vertically from different points, but the overlapping parts of the routes along this road are captured in this representation.

Once the user's knowledge has been assessed, the system must decide on which instruction to offer to the user. Our thought is that the tutor should offer an instruction that matches the user's needs, that is, a turn-by-turn-style instruction when the next part of the route is unfamiliar to the user and a multi-turn instruction when the next part is known. Once an instruction is offered, a brief dialogue between the user and the tutor begins. If the user expresses uncertainty about the instruction offered, the tutor will choose a more fine-grained instruction, which might be comprised of a long-range instruction involving fewer turns or a turn-level instruction. The tutor will need to be mindful of how much time is left before the user approaches the next turn. Similar issues of managing the time available for dialogue during navigation have been considered by Janarthanam et al. (2013). We will consider a reinforcement learning approach to determining an effective strategy for stepping back to more fine-grained instructions in light of the user's response to the initial instruction offered. Reinforcement learning has been been applied to tutorial dialogue systems in other domains to, for example, decide whether to provide feedback after a student's response, or to decide between telling the student a target concept and prompting the student to describe it in his or her own words (Chi et al., 2010; Tetreault and Litman, 2006).

Ahead of more focused studies aimed at addressing the questions raised here, we have carried out a preliminary experiment in a pedestrian navigation scenario. The purpose of the experiment was to make observations about how people give directions remotely. For each of two pairs of participants, one participant acted as the tutor and the other as the user. The tutor directed the user by cell phone to walk to a series of destinations. One tutor could see the user's location in real-time, but the other tutor could not. As a result, the tutor that had access to user location tended to spend more time issuing commands, with limited feedback from the user. By comparison, the tutor that could not see the user's location frequently asked the user to report his location, and the user soon began to volunteer this information as he approached intersections. The user in this latter pair noted that having to report his location helped him to be more aware of where he was. In light of this, we plan to explore prompting the user to self-report progress along a route as another potentially useful tutorial strategy.

We have proposed an approach to designing an intelligent tutoring system for urban navigation. There are several key challenges to be addressed, including how to represent navigational skills and knowledge, how to estimate this information from dialogue with the user and observations of past navigation sessions, and how to choose the next instruction so as to maximize navigational learning over time. If successful, this system would fulfill a need that is not satisfied by current GPS navigation systems: for a navigation system that increases the user's ability to navigate autonomously over time.

References

Gary E Burnett and Kate Lee. 2005. The Effect of Vehicle Navigation Systems on the Formation of Cognitive Maps. In *International Conference of Traffic and Transport Psychology*, pages 407–418.

Min Chi, Kurt VanLehn, and Diane Litman. 2010. Do micro-level tutorial decisions matter: Applying reinforcement learning to induce pedagogical tutorial tactics. In *Intelligent Tutoring Systems*, pages 224–234. Springer.

P. G. Jackson. 1996. How Will Route Guidance Information Affect Cognitive Maps? *The Journal of Navigation*, 49(02):178–186.

Srinivasan Janarthanam, Oliver Lemon, Xingkun Liu, Phil Bartie, William Mackaness, and Tiphaine Dalmas. 2013. A Multithreaded Conversational Interface for Pedestrian Navigation and Question Answering. In *SIGDIAL*, pages 151–153, August.

Martin Johansson, Gabriel Skantze, and Joakim Gustafson. 2011. Understanding Route Directions in Human-Robot Dialogue. In *SemDial*, number September, pages 19–27.

R. Meena, G. Skantze, and J. Gustafson. 2012. A Data-driven Approach to Understanding Spoken Route Directions in Human-Robot Dialogue. In *INTERSPEECH*.

Raveesh Meena, Gabriel Skantze, and Joakim Gustafson. 2013. Human Evaluation of Conceptual Route Graphs for Interpreting Spoken Route Descriptions. In *CoSLI*, pages 30–35.

R. Müller, T. Röfer, A. Lankenau, A. Musto, K. Stein, and A. Eisenkolb. 2000. Coarse qualitative descriptions in robot navigation. In *Spatial Cognition II*, pages 265–276.

Keith J. Oliver and Gary E. Burnett. 2008. Learning-oriented vehicle navigation systems: a preliminary investigation in a driving simulator. In *MobileHCI*, pages 119–126.

Joel Tetreault and Diane Litman. 2006. Using Reinforcement Learning to Build a Better Model of Dialogue State. In *EACL*.

David Wood, Jerome S Bruner, and Gail Ross. 1976. The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, 17(2):89–100.