Using the Situational Context to Resolve Frame of Reference Ambiguity in Route Descriptions

Robert J. Ross and John D. Kelleher

Applied Intelligence Research Center, School of Computing
Dublin Institute of Technology, Ireland
robert.ross@dit.ie, john.d.kelleher@dit.ie

1. Introduction

The application domain for this research is human-robot interaction, specifically the interpretation of route instructions by robots. Two features of spatial language that are of particular relevance to route instruction interpretation are the notions of perspective and spatial reference frame. This paper presents a computational model that addresses the problem of resolving frame of reference ambiguity in route descriptions. This model works by using non-linguistic information from the situational context.

(Levinson, 1996) describes three basic reference frames that can be used in English to describe the static spatial relationship between two objects: the *intrinsic*, the *relative*, and the *absolute*. There has been a number of studies and computational models that have addressed the effect of frame of reference ambiguity on the interpretation, e.g. (Carlson-Radvansky and Logan, 1997; Kelleher and Costello, 2005). By comparison there has been relatively little work on frame of reference ambiguity in route descriptions.

In a route description two other reference frames are commonly made use of (Klatzky, 1998). The *egocentric* perspective, tied to the intrinsic reference frame, is defined by a trajectory created by the direction of movement of the moving object. Allocentric perspectives on the other hand are related to absolute reference frames in that they are defined by virtue of global rather then mover properties. For the case of route directions given with respect to a map at least two allocentric frames of reference can be used: a cardinal reference frame (*north,* south, ...) and a reference frame defined relative to the intrinsic properties of the map, e.g., "drive to the left" where left is defined with respect to the map's left side.

The diversity of perspectives and reference systems choices for a given situation introduces significant complication in mapping between verbal route descriptions and space. Indeed, speakers frequently switch between allocentric and ego-centric (Tversky et al., 1999; Lawson et al., 2008). Furthermore, corpus studies have shown that verbalised routes are highly underspecified (Prévot, 2001; MacMahon, 2007). This combination of diversity of perspectives and underspecification can often give rise to potential ambiguity such that hearers can only produce the intended interpretation through the application of many layers of context – including both discourse and situational information.

1.1 Computational Models of Route Instruction Interpretation

A number of approaches to the computational interpretation of routes from a language understanding perspective have been proposed. These approaches can be split broadly into three categories based on the level of spatial representation afforded to the interpreting agent. *Perception-driven models* assume that the interpreting agent has no global representation of its environment and instead place emphasis on behavior models and the agent's perceived environment (e.g., (Bugmann et al., 2004; MacMahon, 2007)). *Qualitative models* assume that the agent does in fact have a representation of its global environment, but that this global representation is expressed in qualitative rather than quantitative terms (e.g., (Bateman et al., 2007)). Finally, *quantitative models* assume that the agent does have access to a quantitative representation of its environment (e.g., (Mandel et al., 2006; Levit and Roy, 2007; Tellex and Roy, 2006)).

While the various models just considered all include at their core the goal of interpreting linguistic descriptions of routes, the approaches taken vary widely in their cognitive plausibility, representational requirements, and computational costs. Moreover, with few exceptions, the route interpretation models described above ignore issues of linguistic underspecification and ambiguity, and often rely on pre-formalised semantic structures.

2. Application Scenario: Navspace

The development of this work was driven by the *Navspace* scenario. In this scenario, a user gives verbal instructions to direct an autonomous robot around a partially known office environment. Here, the user and system play the roles of Route Giver and Route Follower respectively. Reflecting the real-world capabilities of robotic systems, both the Route Follower and Route Giver are assumed to have shared knowledge of their environment. However, while the Route Follower is capable of moving in the environment and following the Route Giver's instructions, it is the Route Giver who has knowledge of their target destination. Moreover, during a typical interaction with the robot, a user views the shared simulated environment from a plan perspective which includes corridors, various rooms, and the robot's position in the environment at any given time. A user, with a particular known target destination, is then free to direct the robot towards that destination in whatever way that user sees fit.

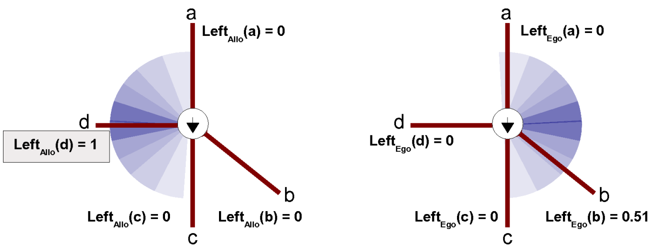

Figure 1: Modality only, i.e., "left". Evaluation of $Left_{Allo}$ on left, and $Left_{Ego}$ on right.

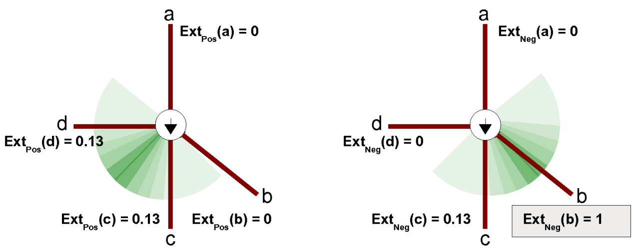

Figure 2: Extent only, i.e., "45 degrees". Evaluation of positive extent on left and negative extent on right.

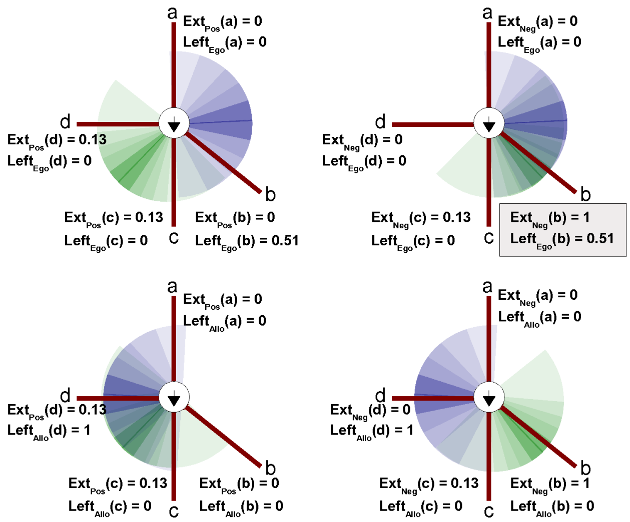

Figure 3: Modality and extent, i.e., "45 degrees to the left". The four figures depict combinations of the evaluation results from Figures 1 and 2.

3. The Computational Model

The computational model is designed to address frame-of-reference ambiguity in route directions given to a robot as it travels through an indoor environment. The model defines parameterised spatial template models for each frame-of-reference and overlays these spatial templates on the situational context. These spatial template models can be understood as defining a distribution of acceptability of a route description, under the assumed frame of reference, on the spatial context. Frame of reference ambiguity is resolved by taking the path in the environment that has the highest acceptability rating as defined by the overlaid spatial templates. For example, a robot that is approaching a junction may be told to "turn left", "turn 45 degrees", or "turn 45 degrees to the left". These different examples illustrate that an instruction may include directional (or modality) and metric information. Figures 1, 2, 3 illustrates how the model overlays on the environment the spatial templates for each frame of reference and computes the acceptability associated with each of the potential paths.

References

John Bateman, Stefano Borgo, Klaus Luttich, Claudio Masolo, and Till Mossakowski. 2007. Ontological modularity and spatial diversity. *Spatial Cognition & Computation*, 7:1.

Guido Bugmann, Ewen Klein, Stanislao Lauria, and Theocharis Kyriacou. 2004. Corpus-Based Robotics: A Route Instruction Example. In *Proceedings of IAS-8*.

L.A. Carlson-Radvansky and G.D. Logan. 1997. The influence of reference frame selection on spatial template construction. *Journal of Memory and Language*, 37:411–437.

John Kelleher and Fintan Costello. 2005. Cognitive representations of projective prepositions. In *Proceedings of the 2nd workshop on the linguistic dimensions of prepositions*, pages 119–127, University of Essex, Colchester, United Kingdom. Association for Computational Linguistics.

Roberta L. Klatzky. 1998. Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. *Lecture Notes in Computer Science*, 1404.

Shaun Lawson, Emile van der Zee, and Laura Daley. 2008. Spatial language in computer mediated communication. In Phil Turner and Susan Turner, editors, *Exploration of Space,Technology, and Spatiality: Interdisciplinary Perspectives*. Hershey, Pennsylvania: Information Science Reference.

S.C. Levinson. 1996. Language and space. *Annual Review of Anthropology*, 25:353–382.

M. Levit and D. Roy. 2007. Interpretation of spatial language in a map navigation task. *IEEE Trans. Systems, Man and Cybernetics*, 37(3):667–679, June.

Matthew Tierney MacMahon. 2007. *Following Natural Language Route Instructions*. Ph.D. thesis, The University of Texas at Austin.

C. Mandel, U. Frese, and T. Röfer. 2006. Robot navigation based on the mapping of coarse qualitative route descriptions to route graphs. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006)*.

Laurent Prévot. 2001. Topic structure in route explanation dialogues. In *Proceedings of the workshop "Information structure, Discourse structure and discourse semantics" of the 13th European Summer School in Logic, Language and Information*.

Stefanie Tellex and Deb Roy. 2006. Spatial routines for a simulated speech-controlled vehicle. In Michael A. Goodrich, Alan C. Schultz, and David J. Bruemmer, editors, *HRI*, pages 156–163. ACM.

Barbara Tversky, Paul Lee, and Scott Mainwaring. 1999. Why do speakers mix perspectives? *Spatial cognition and computation*, 1:399–412.