Classification of Localization Utterances using a Spatial Ontology

Mohammad Fazleh Elahi[1], Hui Shi[1], John A. Bateman[1], Kathleen M. Eberhard[2], Matthias Scheutz[3]

[1] SFB/TR8 Spatial Cognition, Universität Bremen, Germany
{fazleh, shi, bateman}@informatik.uni-bremen.de
[2] Department of Psychology, University of Notre Dame, USA
eberhard.1@nd.edu
[3] Human-Robot Interaction Lab, Tufts University, USA
mscheutz@cs.tufts.edu

Abstract. Dialogue systems for spatially situated tasks need to provide referential descriptions of spatially located objects and understand such descriptions from users. To construct such dialogue systems, it is useful to investigate how humans describe object locations in their immediate environment and how they ask about object locations in remote environments. In this paper, we address the semantic classification of the localization utterances found in the CReST corpus, which is a dialogue corpus of humans performing a cooperative, remote, search task. The aim is to explore the relation between specific semantic configurations and the dialogically and situationally embedded linguistic forms employed. Specifically, we first extracted different types of localization utterances from the corpus and then paired these with semantic categories provided by the linguistically motivated spatial ontology GUM. The paper concludes with a discussion of the characteristics of different types of localization expressions on the basis of spatial concepts and descriptions employed.

1 Introduction

The localization of objects in an indoor environment poses a considerable challenge for dialogue-based intelligent systems, which are employed in a variety of tasks, such as mapping and localization Kruijff *et al.* [6], and indoor wayfinding Cuayahuitl *et al.* [3]. However, the support for object localization in intelligent systems is currently rather limited. For effective interaction, such dialogue systems should be able to deal with place-related knowledge acquired from verbal expressions, such as "now you're in the long hallway", in which the place "hallway" is described as "long". This requires them to support place-related information extraction, representation and reasoning. Our aim here is to explore the semantic classifications of localization utterances, a fundamental effort toward the construction of sophisticated place-aware spoken dialogue systems.

Vasardani M., Winter S., Richter K-F., Janowicz K. and Mackaness W. (eds.); P-KAR 2012, pp. 13-18

Spatial localization has two main aspects: location description and location query. The recently collected human-human dialogue corpus CReST (cooperative, remote, search task) of Eberhard *et al.* [4] covers both communicative aspects and, therefore, provides useful data. In addition, research investigating the relation between linguistic expressions involving space and semantic representations is of relevance (e.g., broad-coverage characterizations of semantics for naturally occuring texts and dialogues [8]). Hence, we will focus on the relationship of natural linguistic expressions with a broader range of more formal spatial characterizations capable of covering more of the semantics of naturally occuring, task-based expressions, combining a data-driven and a semantic classification approach.

The Generalized Upper Model (GUM) is a general task and domain independent "linguistically motivated ontology" that provides linguistic semantics for spatial expressions. It serves as an intermediate "interface ontology" mediating between linguistic forms and contextualized interpretations. In previous work, Bateman *et al.* [2] employed GUM to classify spatial relations in three spatial language corpora: the *Trains* 93 Dialogues (Heeman & Allen [5]), the *HCRC Map Task* (Anderson *et al.* [1]), and the *IBL Corpus* (Lauria *et al.* [7]), and showed that the GUM spatial ontology provided a characterization of the semantics that was of immediate use for several natural language processing tasks. Here we follow this line of research and present the semantics of a set of localization expressions from the CReST corpus using GUM, with focus on static spatial configurations and their elements.

2 Corpus Analysis

The CReST corpus [4] of natural language dialogues was obtained from an experiment involving humans performing a cooperative, remote, search task. The experiment required individuals in a dyad to coordinate their actions via remote audio communication in order to accomplish several tasks with target objects ("colored boxes") that were scattered throughout an indoor search environment (see Figure 1). Neither individual was familiar with the environment before the experiment. One individual was designated as the director (D), and the other as the searcher (S).

The CReST corpus is particularly relevant for the localization problem because the experiment directly involves several distinct localization subtasks. In one subtask, the searcher was to report the locations of eight green boxes in the environment and the director was to mark their locations on the map. Hence, the director needed to learn about the boxes' locations through dialogue. In a second subtask, the director was to direct the searcher to the cardboard box at the furthest point in the search environment, while the searcher was to collect blocks from blue boxes and put them into the cardboard box. Here, the director needed to track the searcher in the environment (from dialogue only, with no visual feedback) and then give further instructions to direct the searcher to the blue boxes.

In order to extract probable relatum and locatum from the dialogues, three wordlists were created. The first wordlist (BP) contains the words related to the places in the environment (room, cubicle, office, hallway, doorway, etc), the second (FO) words related to the fixed objects (door, wall, steps/stairs, stage, booth, etc), and the third (MO) words related to the movable objects (computer, chair, table, shelf, cabinet, etc). The localization-related utterances of the directors were first extracted on a word-by-word basis through comparing each word with those in the wordlists. These utterances were then processed using a set of tools (e.g., tokenizer, part-of-speech tagger, and regular expression based NP chunker) developed in NLTK toolkit[1], in order to find the meaningful noun-phrase(s) (NP) of individual utterances. For instance, NP [the/DT little/JJ tiny/JJ room/NN] is the meaningful phrase of the utterance "is it in the little tiny room?".

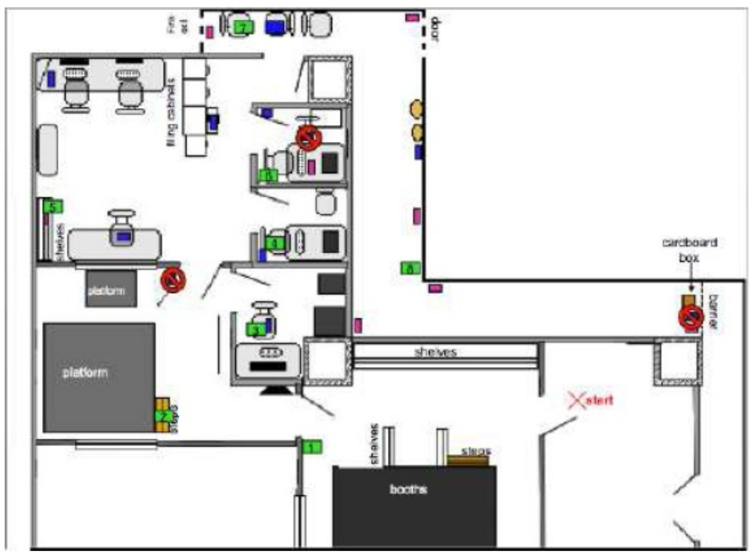

Fig. 1. The figure shows the actual locations of the eight green boxes in the environment which the director was to indicate on the map (CReST; Eberhard *et al.*: [4]). The map of the search environment displayed to the director on a computer screen is as same as this map but without green boxes. As can be seen from the figure, the environment consisted of three big rooms, three small rooms and a surrounding hallway.

Analyzing the number of utterances of each group, we find that rooms of the wordlist BP and doors of the wordlist FO are most frequently used. Moreover, we found that "small", "big","tiny", "first", etc. (see Table 1) are often used to distinguish rooms.

[1] http://nltk.org/

Table 1. Some example utterances from different groups

Group	Examples	Descriptions from the corpus
BP	are you still in the smaller room? okay now you're in the long hallway?	little, small, big, long, large, tiny, initial, first, third, etc
FO	there's a open door in front of you right? is there a door with a single door?	open, close, single, double, small, two, etc
MO	are there filing cabinets you see? are there two chairs at the desk?	filing cabinets, two chairs, etc

3 Semantic Classification of Localization Expressions

The GUM ontology [2] is a *linguistically motivated* ontology based on grammatical evidence from a broad range of linguistic spatial expressions which provides a unified account of spatial concepts and their relations. We use the GUM concept SpatialLocating (SL) to define and classify the semantics of the localization utterances in CReST for two reasons. First, GUM's SpatialModality (SM) covers distance, direction and relationship between the locatum (L) and relatum (R) which are crucial for localization expression. Second, GUM provides spatial semantics to localization expressions regardless of contextual interpretations. As can be seen from table 2, in GUM, "you were in the little cubicle, right?" and "are there like a desk in front of the computer chair?" receive spatial semantics SpatialLocating, where SpatialLocating is the concept that specifies the place (as a placement relation) where an entity (as a locatum) is being positioned. A placement relation can be a GeneralizedLocation (e.g., "on the table") or a GeneralizedRoute (e.g., "through the hallway"). The GeneralizedLocation binds together a relatum and a spatial relationship hasSpatialModality within a single structured entity that may stand in a placement relation within a spatial configuration. The utterance "There is an open door in front of you", for example, is bound with the following GUM semantics:

Configuration:	SpatialLocating	
Locatum:	an open door	
Placement:	GeneralizedLocation	
	hasSpatialModality:	FrontProjectionExternal
	relatum:	you(person)

Exploring the different forms of localization expressions, we found several utterances (e.g, 19.5% utterances of SpatialLocating), in which the locatum (L) is unknown and the relatum (R) is either an object or a person (e.g, the searcher). As can be seen from table 2, the locatum of the sentence "what do you see immediately to your left?" (Configuration: SpatialLocating, Locatum: unknown, Relatum: person, SpatialModality: LeftProjectionExternal) is unknown. From the analysis of spatial relations between a locatum and a relatum, we have found that there are many utterances (e.g, 33.8% utterances of SpatialLocating) related to the GUM concept Containment of a person (e.g., "Are you in a hallway or small room?") or an object (e.g., "Is there like a room that has a bunch of desks with tables and chairs?") in a spatial place (i.e., BP). This is because locating the person first and then the entities that can be seen from that place is essential for solving localization problems in a search task. The GUM concepts ProjectionRelation (e.g., "on the left"), MultipleDirectional (e.g., "it's in the bottom right or bottom left?"), and GeneralDirectional (e.g., "on the opposite side of the filing cabinets") are frequently used for directional relations in the corpus, together with clauses containing words such as "face" or "look" for orientation; for example: "Okay three steps and now when you look to your right there's another open door." and "if you look outside the cubicle there should be a door to your right is that correct?".

Many utterances (e.g, 35.1% utterances of SpatialLocating) in CReST contain what in GUM terms are called ComplexConfigurations, in which the director describes two or more spatial locations in a single utterance. The specifications of such utterances usually combine several SpatialLocating configurations with conjunction ("SLCjSL") or disjunction ("SLDjSL") as shown in table 2. For example, by saying "there's a cubicle on your right hand side and then straight in front of you there's also a door?", the director required the searcher to confirm a spatial setting which only matches with the room in the upper left of the setup shown in Figure 1. Table 2 shows that the complex sentence can be divided into two sentences: "there's a cubicle on your right hand side" and "there is also a door straight in front of you" respectively.

Table 2. Some examples of different kinds of utterances of *SpatialLocating (SL)*

	Examples	L	R	SM	%
SL	what do you see immediately to your left?	-	per	LeftProEx	19.5
	to the left of the filing cabinets what do you see?	-	MO	LeftProEx	
	you were in the little cubicle, right?	per	BP	Contain	33.8
	is there a door anywhere near you?	FO	per	Proximal	16.9
	there's a filing cabinet in front of you, right?	MO	per	FrontProEx	
	is there like a desk in front of the computer chair?	MO	MO	FrontProEx	6.5
SL-CjSL	there's a cubicle on your right hand side and	BP	per	RightProEx	35.1
	then straight in front of you there's also a door?	FO	per	FrontProEx	
SL-DjSL	are you in a hallway or	per	BP	Contain	
	(are you in a) small room?	per	BP	Contain	

4 Conclusions

In this paper, we analyzed location descriptions and location queries in the CreST corpus which naturally includes different forms of localization utterances in a dialogue context. First, to analyze the probable locatum and relatum, we extracted localization utterances which are more frequent in the corpus by classifying the environment into places (BP), fixtures (FO) and objects (MO). Second, we extracted descriptions of each group in order to explore the levels of descriptions attached with locatum and relatum. Finally, we paired these utterances with semantic categories provided by the GUM ontology to explore the specific semantic configurations and linguistic forms employed in localization query. The findings include (1) the adjectives and clauses are used to describe the locatum and relatum, (2) the entities of the locatum and relatum are often found unknown or variable, (3) the spatial relations such as Containment and Projectional are found frequently used, and (4) the query which describes two or more spatial locations are used to uniquely identify a spatial setting in the environment. Since linguistic semantics of localization expressions formalized in GUM is domain-independent, the semantic classifications and findings of this research offer a general conceptualization for relating place-sensitive natural expressions to their spatial semantic interpretation and therefore, provide a sophisticated foundation for the contextualization and generation of place-aware natural expressions in situated dialogue systems, like DAISIE (see Ross & Bateman [9]).

References

1. A. H. Anderson, M. Bader, E. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. Thompson, and R. Weinert. The HCRC Map task Corpus. *Language and Speech*, 34(4):351–366, 1991.
2. J. A. Bateman, J. Hois, R. J. Ross, and T. Tenbrink. A linguistic ontology of space for natural language processing. *Artificial Intelligence*, 174(14):1027–1071, September 2010.
3. H. Cuayáhuitl, N. Dethlefs, K.-F. Richter, T. Tenbrink, and J. Bateman. A dialogue system for indoor wayfinding using text-based natural language. In *Proceedings of the 11th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2010): posters and short presentations*, Iasi, Romania, 2010. March 21-27.
4. K. Eberhard, H. Nicholson, S. Kbler, S. Gunderson, and M. Scheutz. The indiana "cooperative remote search task"(crest) corpus. In *Proceedings of the Seventh Language Resources and Evaluation Conference (LREC 10)*, 2010.
5. P. A. Heeman and J. Allen. The Trains 93 Dialogues. Trains Technical Note 94-2, Computer Science Dept., University of Rochester, mar 1995.
6. G.-J. M. Kruijff, H. Zender, P. Jensfelt, and H. I. Christensen. Clarification dialogues in human-augmented mapping. In *Proceedings of the 1st Annual Conference on Human-Robot Interaction (HRI'06)*, Salt Lake City, UT, March 2006.
7. S. Lauria, G. Bugmann, T. Kyriacou, J. Bos, and E. Klein. Training Personal Robots via Natural-Language Instructions. *IEEE Intelligent Systems*, pages 38–45, 2001.
8. I. Mani, C. Doran, D. Harris, J. Hitzeman, R. Quimby, J. Richer, B. Wellner, S. Mardis, and S. Clancy. SpatialML: annotation scheme, resources, and evaluation. *Language Resources and Evaluation*, 44(3):263–280, 2010. Special Issue LREC 2008: Selected papers, Edited by: N. Ide and N. Calzolari.
9. R. J. Ross and J. A. Bateman. Daisie: Information state dialogues for situated systems. In V. Matoušek and P. Mautner, editors, *Text, Speech and Dialogue*, volume 5729 of *Lecture Notes in Computer Science*, pages 379–386. Springer, Berlin / Heidelberg, 2009. 12th International Conference, TSD 2009, Pilsen, Czech Republic, September 13-17, 2009. Proceedings.