Interpreting Destination Descriptions in a Cognitive Way

Yunhui Wu* and Stephan Winter

Department of Infrastructure Engineering, University of Melbourne
Parkville, VIC 3010, Australia
y.wu21@pgrad.unimelb.edu.au,winter@unimelb.edu.au

Abstract. This paper proposes a cognitively motivated approach to interpreting destination descriptions without computing spatial relations. In contrast to other computational approaches, this approach is based on a few assumptions drawn from human communication behavior. Although this cognitively motivated approach is relatively simple, the performance of the approach is almost as good as other computational approaches.

Keywords: Destination description, spatial reasoning, spatial relation

1 Introduction

People provide destination descriptions when they specify where to go to. Destination descriptions are referring expressions [1] of the form "x related to y", where x is the destination, and y is a reference feature. A destination description is a reflection of the speaker's conceptual map of the environment in their mind. In geographic environments, people perceive salient features (landmarks), anchoring their mental representations of the environment [2]. They update their knowledge by linking new experiences of other features to the existing ones. Therefore it is natural for people to describe the location of features by addressing their spatial relation to other, more salient features. Using landmarks in destination descriptions is also a way to set up the common ground between parties in the communication: the speaker expects that the listener knows the landmarks due to their salience in the urban environment, and then, through the spatial relation with the landmarks, figure out where the destination is. This paper focuses on the spatial reasoning of using spatial relations in human destination descriptions, and proposes an approach to interpret these descriptions automatically to smarten the user interaction of navigation services.

Although *humans* have the capability of understanding the spatial relations in destination descriptions, making sense of spatial relations is not an easy task for *computational systems*. The major challenge is interpreting the qualitative

* This research was supported under Australian Research Council's Discovery Projects funding scheme (project number 0878119).

relations frequently used in destination descriptions. Characterizing or interpreting *topological* relations, such as *in*, requires access to the spatial extents of the involved spatial features. *Orientation* directions, such as *in front of*, require size and shape information. *Distance* relations, such as *near*, require taking contextual factors into account such as the size of the features, the purpose of the located feature, the distance from the observer, the functional relationship and interaction between the two features, and the asymmetry from the order in which locations are retrieved in memory [3–5]. However, so far there is no comprehensive computational model able to handle all these factors for qualitative spatial relations. Given these difficulties, this paper studies the spatial reasoning behind human communication behavior, and suggests a cognitively motivated approach to interpreting destination descriptions. This approach does not require any of the additional information, but instead is built on point locations (as given in standard gazetteers), and a salience model of geographic features. We will in particular demonstrate that the cognitively motivated approach can identify destinations without computing any other spatial relations than neighborhoods based on salience.

2 Related Work

Common ground is the basis of joint actions by speakers and listeners [6]. People do things based on individual beliefs or assumptions about what is common ground between each other. Clark identifies two kinds of common ground: communal common ground and personal common ground. Communal common ground is based on factors, such as communication parties' nationality, residence, education, occupation, and religion. Personal common ground is based on joint personal experiences. This paper assumes that the speaker refers to communal common ground, as in talking to strangers, such that spatial databases can be used to enable the interpreting process of destination descriptions.

Research has been made on formalization and computational modeling of spatial relations. Models for characterizing topological relations exist (e.g., [7]). A cognitive and computational model for *nearness* has been developed before [8]. Schlieder et al. propose to encode neighborhood relations in gazetteers for retrieving qualitative information [9]. Freksa develops an approach for representing qualitative spatial reasoning using orientation information [10], which is later developed into reasoning toolboxes [11]. But yet a comprehensive computational model for qualitative relations is not found.

3 Cognitive Motivated Approach to Interpreting Destination Descriptions

If destinations are hard to recognize, ambiguous or lacking in the common ground, people usually refer to the most salient landmark nearby according to their knowledge, which is chosen from potentially large numbers of spatial features available. From the speaker's perspective, the more salient the landmark is in the environment, the more likely it is to be known to the listener. However people perceive the urban environment variously. Petrol stations are more meaningful, thus salient, to car drivers than to walkers. Therefore it is more likely that car drivers refer to petrol stations in destination descriptions than walkers. It appears that the choice of a landmark relies more on its salience than on the spatial relation between the destination and the addressed landmark. Furthermore, spatial relations are mental connections, or characterizations of the configuration of spatial features at particular locations [12]. Therefore when different speakers refer to the same landmark, they may use different terms to depict the spatial relation or even different relations. Here the first assumption is:

– It is always the most salient landmark chosen among others, no matter what the type of the spatial relation between the landmark and the destination is.

This assumption establishes a basis for interpreting spatial relations without computing them explicitly. By saying "the most salient landmark among others", there must be implied a spatial restriction from which landmarks are selected. This restriction can be derived from the principle of relevance [13], which we apply here by a second assumption. We expect that the landmark has to include the destination within their neighborhood – a concept that needs to be further formalized. If the destination is not in the neighborhood of the landmark, the relationship is too weak to use in the destination description, since the relationship to another landmark is stronger. So the second assumption is:

– The landmark is chosen only if the destination is within the landmark's neighborhood.

By referring to a chosen landmark, the speaker wants to ensure that the listener can figure out the destination effectively and unambiguously. If there are two pizza shops in the neighborhood of the landmark, the speaker has some choices to disambiguate. They can specify the name of the target one, such as "the *Pizza Hut* next to the petrol station" (i.e., not *Domino's*), or employ another landmark to avoid such confusion, like "the pizza shop *opposite the church*" (which is also next to the petrol station, but does not apply for *Domino's*). Or they can name a disambiguating spatial relation, like "the pizza shop *left of* the petrol station" (instead of the one right of the petrol station). Except for the third case, it can be inferred that in destination descriptions the destination is unique within the neighborhood of the chosen landmark. The third case can be discovered either from inflection (where the relation would be stressed), or from discovering the ambiguities in the interpretation. The third case requires special treatment, but for the other cases we can make our third assumption:

– The landmark is chosen because it is sufficient enough to disambiguate the destination.

These assumptions require a computational model of neighborhood. Moulin et al. advise that the influence area of a spatial feature defines the portion of neighborhood in which every other features are spatially related to the located feature in a qualitative way [14]. The influence of a spatial feature can be used to define its neighborhood in the environment. Salience of landmarks represents their influence: the more salient a landmark, the larger its influence area. Winter et al. suggest a method of generating a hierarchical partition establishing the neighborhoods of landmarks at different levels of salience (or context) [15]: landmarks are grouped by their salience at different levels in a hierarchy, and then Voronoi cells representing the neighborhood of landmarks are created at each level (Figure 1).

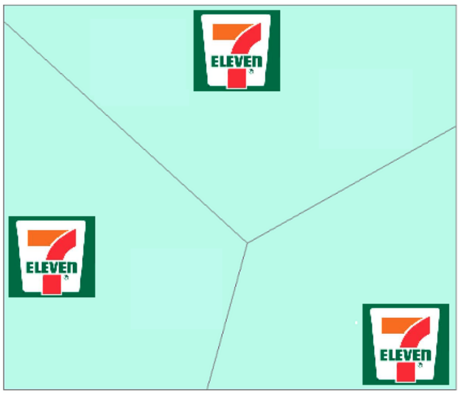

Fig. 1. Illustration of Voronoi diagram between landmarks of similar salience.

The interpretation of a destination description starts with a list of identified potential destination candidates. In the example above, it would start with a list of all known pizza shops. If there is no destination (no pizza shop) found in the database, the interpretation fails, similar to lacking common ground in human-to-human communication. If there is only one pizza shop found in the database, the interpretation completes successfully, and the relation to the landmark can only be used in an affirmative way. But if there are multiple destination candidates found, then the assumptions above will allow the disambiguation of destination candidates. This disambiguation process will use only the locations of landmarks, the salience of landmarks, and the location of destination candidates. The interpretation process computes a second list, namely a list of landmark candidates, e.g., all petrol stations. If this list is empty, no common ground could be established. If exactly one landmark candidate is found, the nearest destination is considered as a solution. If multiple landmarks are found, their neighborhoods are computed [15], and the one that has a unique destination candidate in its neighborhood identifies the destination.

This algorithm discovers automatically the third case – where the spatial relation is used for disambiguation – when no landmark has a unique destination candidate in their neighborhood. In this case the algorithm has to fall back to computing the spatial relations (which is possible, but not addressed in this paper).

This interpretation process offers an approach that avoids in many cases computing of spatial relations. The next section explains by example how this approach works.

4 Example and Discussion

Angela wants to meet her friend for lunch, and says "let's meet at the pizza shop next to the 7-Eleven". The pizza shop is the destination (x), and the 7-Eleven is chosen as the landmark (y). This section demonstrates how the cognitively motivated approach interprets the spatial relation in this destination description, and finds "the pizza shop". At first it is supposed that the spatial restriction of this communication is known from context (the area shown in Figure 2). In this area three 7-Eleven and three pizza shops are found (Figure 2, left). The 7-Eleven are of similar salience, therefore no hierarchy is created. The neighborhoods of three 7-Eleven are defined by Voronoi cells. In Figure 2, left, the 7-Eleven at the

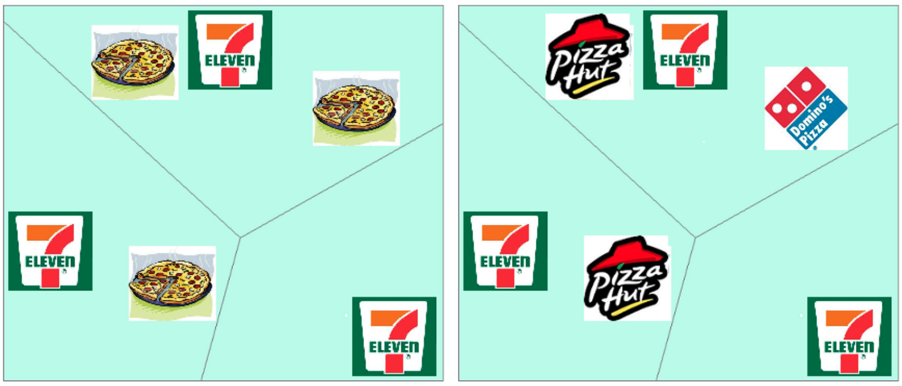

Fig. 2. Left: Three 7-Elevens and three pizza shops found in neighborhoods of the landmarks; Right: Two pizza shops are *Pizza Hut*, and one is *Domino's*.

right-bottom corner has no pizza shop in its neighborhood, thus does not define any destination; the 7-Eleven on the top has two pizza shops in its neighborhood, thus destination candidates are found ambiguous; and the 7-Eleven on the left has a unique pizza shop in its neighborhood. Assuming the rules of relevance theory, this pizza shop would be the target destination. Therefore, the hypothesis is proven. This is the general process of the cognitively motivated approach, and the computation complexity is $O(n)$. In comparison, other computational approaches need to compute the spatial relation between each 7-Eleven and pizza shop, and then check whether the relationship is a "next to" relation. If this process identifies any nearest pizza shop to a 7-Eleven, it could omit less relevant pizza shops. Thus the computation complexity is $O(n^2)$.

In some cases, the general process cannot obtain a unique result. For example, Angela specifies the name of the pizza shop by saying "let's meet at the Pizza Hut next to the 7-Eleven". Figure 2, right, shows a unique Pizza Hut in two 7-Eleven's neighborhood separately. In this case, the two 7-Eleven (on the top and on the left) are results through this cognitively motivated approach, and further refinement to resolve the remaining ambiguity is needed.

Computational approaches require separate computation algorithms for various types of spatial relations, i.e., topology, orientation, distance relations. As natural language is flexible, categorizing spatial relation in destination descriptions may introduce error. Mismatching between the identified types from descriptions and the preset types in algorithms will also cause failure. This cognitively motivated approach only checks the uniqueness of destination candidates within neighborhood of each landmark candidates, therefore avoids these risks. However the precision of this approach depends on the appropriate definition of landmark neighborhood. Imprecision may also be produced when spatial relations are used to disambiguate destinations.

5 Experimental Evaluation

To evaluate the performance of the cognitively motivated approach, a gazetteer was built, including 36,134 instances. The gazetteer data is based on the point of interest data from Whereis[1] and VicNames data[2] over the entire area of Victoria, Australia. Each instance consists of three essential attributes: the place name, the category of place, and a geographic location provided by the data sources [16].

The only other attribute needed is a salience value. Salience is derived here by a method suggested by Duckham et al. [17] utilizing the categories of gazetteer instances. Each category in a gazetteer is assessed by an expert on nine criteria (physical size, proximity to road, visibility, difference from surroundings, ubiquity, nighttime vs. daytime salience, permanence, length of description and spatial extents) in two ways: the average salience of individual instances in a category (suitability) and their standard deviation (typicality). The final salience of each category is then normalized in the range [0,1]: 1 represents the highest suitability, and 0 represents the lowest suitability.

From salience, the influence areas of all instances are computed at all levels of a salience hierarchy, according to Winter et al. [15]. This concludes the preprocessing of generating a suited gazetteer.

After preparing the gazetteer data, we collected 57 destination descriptions given by participants in an interview experiment. Examples of these collected destination descriptions are "Yarra Bend Park near Alphington", and "Lorne on the Western Coast Road, between Geelong and Apollo Bay, about half way

[1] www.whereis.com.au

[2] http://services.land.vic.gov.au/vicnames/

between each". In total, 80 individual places were mentioned in the collection of destination descriptions, including street names, suburb names, names of stations, restaurants, shopping centers, hospitals, clubs, universities, and parks. Besides individual places, there are 15 paraphrased places found in the collection, such as "the library". In the collected data, there are 38 descriptions (67%) including spatial relations and reference place names.

For this experiment the cognitively motivated approach was implemented to interpret the given destination descriptions. Participants were asked to judge the interpretation results. For comparison, we also developed an approach computing topological, orientation and distance relations. According to their judgement, 27 destination descriptions were interpreted correctly by the cognitively motivated approach, and 28 by the approach computing relations explicitly (the cognitively motivated approach was not allowed to fall back on the explicit computation of relations). The results show that the performance of the cognitively motivated approach is almost as good as the approach with explicit computation of relations.

6 Conclusions

Destination descriptions can use various qualitative spatial relations, thus computing spatial relations can be computationally expensive. This may be one of the reasons why no commercial navigation system has implemented methods for interpreting spatial relations (Google Maps, for example, ignores any given relationship and imposes a 'near' relationship on any destination description, of which the semantics remains opaque, of course). Compared to other computational approaches, this cognitively motivated approach is relatively simple, because no computation of individual spatial relation needed.

This paper proposes a cognitively motivated approach to interpreting destination descriptions without computing spatial relations. This approach is based on disambiguating combinations of the multiple destination and landmark candidates found in gazetteers. Given the context, further discussion is needed to retrieve relevant destination candidates and salient landmark candidates from gazetteers. Furthermore, the cognitive adequacy of the construction of the hierarchy of neighborhoods also needs further study.

References

1. Reiter, E., Dale, R.: Building Natural Language Generation Systems. Cambridge University Press, Cambridge, UK (2000)
2. Couclelis, H., Golledge, R.G., Gale, N., Tobler, W.: Exploring the anchorpoint hypothesis of spatial cognition. Journal of Environmental Psychology **7** (1987) 99–122
3. Morrow, D.G., Clark, H.H.: Interpreting words in spatial descriptions. Language and Cognitive Processes **3**(4) (1988) 275–291
4. Ferenz, K.S.: The role of nongeometric information in spatial language. dissertation, Dartmouth College, Hanover, NH (February 2000)

5. Mcnamara, T.P., Diwadkar, V.A.: Symmetry and asymmetry of human spatial memory. Cognitive Psychology **34**(2) (11 1997) 160–190
6. Clark, H.H.: Using Language. Cambridge University Press, Cambridge, UK (1996)
7. Egenhofer, M., Franzosa, R.: Point-set topological spatial relations. International Journal of Geographical Information Systems **5**(2) (1991) 161–174
8. Duckham, M., Worboys, M.: Computational structure in three-valued nearness relations. In Montello, D., ed.: Spatial Information Theory. Volume 2205 of Lecture Notes in Computer Science. Springer, Berlin (2001) 76–91
9. Schlieder, C., Vögele, T., Visser, U.: Qualitative spatial representation for information retrieval by gazetteers. In Montello, D.R., ed.: Spatial Information Theory. Volume 2205 of Lecture Notes in Computer Science. Springer, Berlin (2001) 336–351
10. Freksa, C.: Using orientation information for qualitative spatial reasoning. In Frank, A., Campari, I., Formentini, U., eds.: Theories and Methods of Spatio-Temporal Reasoning in Geographic Space. Volume 639 of Lecture Notes in Computer Science. Springer, Berlin (1992) 162–178
11. Wallgrün, J.O., Frommberger, L., Wolter, D., Dylla, F., Freksa, C.: A toolbox for qualitative spatial representation and reasoning. In Barkowsky, T., Knauff, M., Ligozat, G., Montello, D.R., eds.: Spatial Cognition V: Reasoning, Action, Interaction. Volume 4387 of Lecture Notes in Artificial Intelligence. Springer, Berlin (2007) 39–58
12. Van der Zee, E., Nikanne, U., eds.: Cognitive interfaces: constraints on linking cognitive information. Oxford University Press, New York, USA (2000)
13. Tomko, M., Winter, S.: Pragmatic construction of destination descriptions for urban environments. Spatial Cognition and Computation **9**(1) (2009) 1–29
14. Moulin, B., Kettani, D., Gauthier, B., Chaker, W.: Using object influence areas to quantitatively deal with neighborhood and perception in route descriptions. In Hamilton, H., ed.: Advances in Artificial Intelligence. Volume 1822 of Lecture Notes in Computer Science. Springer, Berlin (2000) 69–81
15. Winter, S., Tomko, M., Elias, B., Sester, M.: Landmark hierarchies in context. Environment and Planning B **35**(3) (2008) 381–398
16. Hill, L.: Core elements of digital gazetteers: Placenames, categories, and footprints. In: Research and Advanced Technology for Digital Libraries. Volume 1923 of Lecture Notes in Computer Science. Springer, Berlin (2000) 280–290
17. Duckham, M., Winter, S., Robinson, M.: Including landmarks in routing instructions. Journal for Location-Based Services **4**(1) (2010) 28–52