Where am I?/Onde Estou? Automated Interpretation of Human Language Descriptions of Current Location

Kristin M. Stock[1] and Luciene S. Delazari[2]

[1]School of Engineering and Advanced Technology, Massey University, Albany, New Zealand
k.stock@massey.ac.nz

[2]Geodetic Science Program,Federal University of Parana, PO-BOX 19001, Curitiba, PR, Brazil
luciene@ufpr.br

Abstract. The automated interpretation of human language descriptions of current location has multiple potential applications, most importantly for emergency services. The paper proposes an approach for user interaction through graphical and conversational interfaces in which a user interacts with the system to describe his or her current location and then resolves ambiguities and refines the description as required in response to system attempts to map it to a spatial location. The paper also proposes a method for automated interpretation of user descriptions in human language (English and Portuguese) via a geometric configuration ontology, using both geometric and empirical approaches. Evaluation and refinement of both arms of the research are ongoing.

Keywords: User interfaces, natural language processing, spatial location

1 Introduction

If you find yourself lost in an unknown part of town; witness an accident and need an ambulance; or get a flat tire late at night in the countryside, you may need to call for assistance and provide details of your current location. Street signs and addresses are not always visible, GPS is sometimes unavailable or the signal unreliable; and location by mobile phone network is often of insufficient accuracy [15]. Human language (e.g. English, Portuguese) descriptions of current location are most likely to refer to visible landmarks, feature types and localities. An example of such a description is: *'I am near McDonalds, on a main road, there is a river to my left and a service station behind me.'*

In this paper, we describe a method for the automated mapping of descriptions of current location in natural language to a spatial query, from which an address or geographic coordinates can be determined, together with two user interfaces (a mobile web-based graphical interface and a speech-based conversational interface) to support user interaction. User input is used to resolve any ambiguity.

We are developing and testing the method with two languages: English and Portuguese, to widen the scope beyond a single language and to identify differences that

adfa, p. 1, 2011.

arise between language and their implications for the method. In later stages, we will consider languages that are less closely related.

This endeavor poses a number of challenges: (1) for the user because suitable landmarks may be limited; (2) semantically because natural language descriptions are inherently vague, imprecise, contextual and regularly used non-literally (3) computationally because the description may map to a query that results in either no results or a large number of results and (4) for data resources, because user descriptions may not match available data sets. Our focus is particularly on the second of these.

2 Related Work

While there has been extensive work exploring the nature of categories of geographic features that might be referenced in natural language location descriptions [e.g. 17, 18] and some work has investigated the kinds of descriptions that people use to explain their locations [e.g. 21, 7], the hierarchical structure of those descriptions [13] and the identification of place names [e.g. 10, 2, 14]; there has been little work on the fully automated interpretation of natural language location expressions. The work that has been done has focused on the interpretation of individual spatial relation words (e.g. spatial prepositions) or phrases (e.g. the road crosses the park) [e.g. 16, 12]. The Geometric Configuration Ontology (GCO) [19] that is the basis for mapping onto spatial queries in this paper used qualitative spatial reasoning (QSR) work from a number of different researchers [e.g 16, 4], and methods to identify locations by combining geometric and qualitative (QSR) relations with approximate boundaries have also been developed [3, 5]. The specific goal of the current work focusses on how we can map from human spatial language to QSR relations in ways that incorporate contextual factors. Previous research on context has included consideration of the factors that may result in different interpretations of prepositions [e.g. 1] and the types of linguistic devices that are used to specify a location [20].

Work on the provision of interfaces particularly focused around helping users to find their current locations is scarce and is usually considered a task appropriate for general web-mapping interfaces (e.g. Google Maps), which do not support the kinds of interactions described in this paper.

3 Graphical and Speech Interfaces

The visual interface for interaction comprises an app for collecting the user position from the mobile device in use. If satisfactory GPS signals or mobile network location is unavailable, we allow the user to describe his or her current location. While the description of location in free text allows the user the most flexibility, there is some evidence that users prefer constraints in the kinds of format they can use to express natural language requests [9]. To this end, we are collecting a corpus of examples of natural language descriptions in English and Portuguese of current user location, from which we will distill common sub-expression formats to develop a suitable set of constraints to be used in the graphical interface, in which users will be given a series of templates to fill in conjunctively, with selection choices to fill in gaps in those templates (for example, ‘There is a” <geog feature> <spatial relation> me.” could be used to express: there is a pedestrian crossing in front of me). Preexisting sets of place names (Geonames[1]) and geographic features (GEMET[2]) will be used to populate the selection choices. If there is more than one possible result from the interpretation method described in Section 4, the user will be asked to select from a graphically displayed list of candidates.

The speech interface is designed to substitute for the graphical interface. In this case the user will describe his or her current location using an unconstrained natural language description, from which the method described below will attempt to identify a location that meets the description. The user will then be asked additional questions through a conversational interface to resolve any ambiguity in the location.

4 Automated Interpretation Method

4.1 Preprocessing of Input

The first stage of preprocessing converts speech into text in the case of input through the speech interface using third party speech recognition software. Templates (Section 3) are used to ensure that user input conforms to a structure that can be processed, as shown here (described in Extended Backus-Naur Form):

input expression = {[“I am” | “Estou”] , spatial relation , [place name | geographic feature type]] | [[“There is a” | Existe um/uma [place name | geographic feature] , spatial relation , [“me” | ?]]};

spatial relation = [“opposite” | “on” | “beside” | “behind” | “in front of” | “left of” | “right of” | “near” | “next to” | “do lado oposto” | “em” | “ao lado” | “atrás” | “em frente a” | “à esquerda de” | “ à direita de” | “perto” | “próximo a “];

place name = names listed in the GeoNames database;

geographic feature type = feature types listed in GEMET or similar ontology;

The second stage of preprocessing involves converting a text input expression that conforms to the syntax above into (x, spatial relation, y) 3-tuples. For the original example input expression, we have the following pairs, connected conjunctively since all are simultaneously true for the observer:

(“x”, “near”, “McDonalds”) ∧ (“x”, “on”, “main road”) ∧ (“river”, “left of”, “x”) ∧ (“service station”, “behind”, “x”)

In this work, we are most interested in fine scale descriptions of location of the kind shown here. As discussed in [13], location descriptions are likely to also contain higher level location information (e.g. in the city of Curitiba, Brazil), which in a production system we would establish either from the user or through a low-accuracy mobile phone network location.

---

[1] http://www.geonames.org/

[2] https://www.eionet.europa.eu/gemet/

4.2 Mapping onto Spatial Queries

We are testing two different approaches to the mapping of each clause in the conjunctive expression into a spatial query. The first approach adopts the geometric definitions of spatial queries for each natural language spatial relation type described in [19]. The second approach uses human interpretations of spatial expressions collected through empirical studies. By finding the most similar natural language expression for which we have a human interpretation, we can derive an interpretation for a new expression. This approach has the benefit of considering context and incorporating non-literal spatial relation descriptions, as manual human interpretations that include a specific context are included within the method.

Formulation of Queries using Geometric Definitions. The GCO queries conform to ISO 13249-3 [8], which is implemented by all major spatial database engines and geographic information systems. We search for members of the potential locations data set that have the required spatial relation with the other defined features. Thus the set of candidate locations for our example are defined as follows:

cl $\in$ CL $\Leftrightarrow$ cl $\in$ PL $\wedge$ ST_DWithin(cl,"McDonalds Restaurants",2σ) $\wedge$ (ST_Overlaps(cl, "Main Roads") = 1) $\wedge$ ((ST_Azimuth(cl,"Rivers") $<$ Θ) $\wedge$ (ST_Azimuth(cl, "Rivers") $>$ $\Theta \pm \pi$)) $\wedge$ ST_Angle(ST_Azimuth(cl, "Service Stations"), Θ) $> \pi/2$

Where σ is a constant [see 3] and Θ is the direction that the observer is facing. Since the observer is human and can move at will, Θ is arbitrary, but must be the same for all clauses in order for the spatial relations to all of the different features to be consistent. Figure 1 illustrates the example scenario.

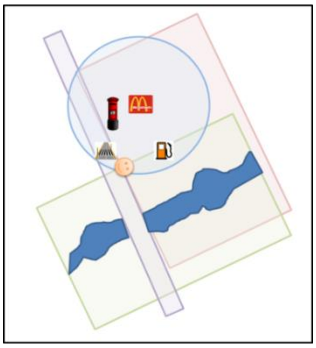

Figure 1: Candidate Locations

Left of and *behind* are difficult to convert into queries given that they are egocentric relations, referring to the current direction of the observer. For the purpose of finding his or her location, the observer's current direction is not important, but in this case, the relationship between the observer and the river, and the observer and the service station, tell us something about the relationship between the river and the service station that may be important.

Formulation of Queries using Empirical Data. *On* is an example of a spatial relation whose interpretation is very context dependent. *The house is on the island*, *the shop is on the main road* and *the boat is on the river* all describe quite different spatial relations, and this cannot be handled by simple geometric models for a given spatial relation as described above. For this reason, we develop a method in which the spatial relation expression including feature types, is matched to a similar expression in a knowledge base of human interpretations of spatial relation expressions. We are testing three different similarity matching expressions: direct term; ontology-based and Wordnet-based. If a sufficiently similar match is found in the knowledge base, we use the GCO concept for the most semantically similar expression in the knowledgebase. For example, *on main road* most closely matches the very near (a,b) GCO concept, rather than the substantially collocated (a,b) GCO concept used for the *on* spatial relation in the geometric method, thus:

cl ∈ CL ⇔ cl ∈ PL ∧ ST_DWithin(cl,"McDonalds Restaurants",2σ) ∧ ST_DWithin(cl,"Main Roads",σ) ∧ ((ST_Azimuth(cl,"Rivers") $< \Theta$) ∧ (ST_Azimuth(cl, "Rivers") $> \Theta \pm 2\pi$)) ∧ ST_Angle(ST_Azimuth(cl, "Service Stations"), Θ) $> \pi/2$

4.3 Resolution of Ambiguities

The final stage involves user confirmation of the selected location. The speech interface confirms or clarifies the location through landmarks and routes as important elements of human wayfinding [11]. We use a hierarchy of landmark types to identify the most prominent landmark that is within a specified distance of the candidate location. If there are no landmarks available, the geometric layout of features is used (intersections, bends in roads, building corners, etc.). The graphical option presents the user with a map of each candidate area showing the landmarks in the region as well as the features mentioned in the original location description and asks the user to confirm or select an alternative.

5 Conclusions and Further Research

Open questions that are being explored as part of this research include the best methods for user interaction both visually, through written input (e.g. the degree of constraint) and verbally, and further extension and refinement of the interpretation method is ongoing, to include additional components (e.g. adjective descriptions, inter-object locations).

Acknowledgements

We gratefully acknowledge the anonymous reviewers for their feedback, and the support of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Governo Federal Brazil under the Science without Borders Program, processes 401264/2014-9 and 301980/2014-4 – research grant.

References

1. Aurnague, M., Vieu, L.: Towards a Formal Representation of Space in Language: A Commonsense Reasoning Approach. IJCAI-93 Workshop on Spatial and Temporal Reasoning, 123-158 (1993)
2. Derungs, C., Purves, R. S.: From text to landscape: locating, identifying and mapping the use of landscape features in a Swiss Alpine corpus. International Journal of Geographical Information Science, 28(6), 1272-1293 (2014)

3. De Felice, G., Fogliaroni, P. and Wallgrün, J.O., A Hybrid Geometric-Qualitative Spatial Reasoning System and Its Application in GIS, in: Spatial Information Theory, pages 188-209, Springer Berlin Heidelberg, 2011
4. Du, H., Alechina, N., Stock, K., Jackson, M.: The logic of NEAR and FAR. Conference on Spatial Information Theory, Scarborough, UK, 2-6 September 2013. Lecture Notes in Computer Science 8116 (LNCS) (2014)
5. Fogliaroni,P. and Hobel, H., Implementing Naïve Geography via Qualitative Spatial Relation Queries, 18th AGILE Conference on Geographic Information Science, Lisbon, 9-12 June (2015)
6. Goyal, R., Egenhofer, M.: The Direction-Relation Matrix: A Representation for Directions Relations between Extended Spatial Objects, Proc. Univ. Consortium for Geographic Information Systems Science (UCGIS) Ann. Assembly and Summer Retreat, June, (1997)
7. Hirtle, S., Richter, K-F., Srinivas, S., Firth, R: This is the tricky part: When directions become difficult. Journal of Spatial Information Science 1, 53-73 (2010)
8. International Standards Organisation. Information technology–Database languages–SQL Multimedia and application packages–Part 3: Spatial. ISO 13249-3. 2011-04-15.
9. Kaufmann, E., Bernstein, A., Evalauting the Usability of Natural Language Query Languages and Interfaaces to Semantic Web Knowledge Bases. Journal of Web Semantics: Science, Services and Agents on the World Wide Web 8, 377-393 (2010)
10. Leidner, J.L., Lieberman, M.D.: Detecting geographical references in the form of place names and associated spatial natural language. *SIGSPATIAL Special* 3, 2,5-11 (2011)
11. Lynch, K. The Image of the City. Cambridge, Massachusetts, MIT Press (1960)
12. Mark, D. M., Egenhofer, M. J.: Topology of Prototypical Spatial Relations Between Lines and Regions in English and Spanish. In: Proceedings, Auto Carto 12, pp. 245-254, Charlotte, North Carolina, (1995)
13. Richter, D., Winter, S., Richter, K. F., Stirling, L.: Granularity of locations referred to by place descriptions. Computers, Environment and Urban Systems, 41, 88-99 (2013)
14. Santos, J., Anastácio, I., Martins, B.: Using machine learning methods for disambiguating place references in textual documents. GeoJournal, 1-18 (2014)
15. Schmidt-Dannert, A., & SNET, S. M. M. (2010). Positioning technologies and mechanisms for mobile devices. In *Seminar Master Module SNET2 (TU-Berlin, ed.)*.
16. Schwering, A.: Evaluation of a semantic similarity measure for natural language spatial relations. In Proceedings of the International Conference on Spatial Information Theory (COSIT 2007), Melbourne, Australia (2007)
17. Smith, B., Mark, D.: Ontology with human subjects testing. American Journal of Economics and Sociology 58(2), 245–312 (1998)
18. Smith, B., and David M. M.. "Geographical categories: an ontological investigation." *International journal of geographical information science* 15.7 (2001): 591-612.
19. Stock, K.: A Geometric Configuration Ontology to Support Spatial Querying. 17th AGILE Conference on Geographic Information Science, Castellon, 3-6 June, (2014).
20. Tomai, E., Kavouras, M.: Where the city sits? Revealing Geospatial Semantics in Text Descriptions. In: 7th AGILE Conference on Geographic Information Science, Association of Geographic Information Laboratories for Europe, Heraklion, Greece, 189-194 (2004)
21. Vasardani, M., Winter, S., Richter, K-F.: Locating place names from place descriptions. International Journal of Geographical Information Science (2013)
22. Vieu, L., Aurnague, M.: Part-of relations, functionality, dependence. In M. Aurnague, M. Hickmann and L. Vieu, editors, The categorization of spatial entities in language and cognition, John Benjamins (Human Cognitive Processing 20), 307-336 (2007).