Natural Language Question Understanding Requires Concept-Level Reasoning for Resolving Reference Ambiguities

Anne Schleicher and Diedrich Wolter
University of Bamberg

Abstract

Mastering GI science work involves learning to map intuitive questions to concrete computations. Since language is tightly connected to human thinking we expect to be facing fundamental challenges underlying design of GI science teaching also in the area of automated spatial question understanding. In this paper we argue that interpretation of a question as well as answering the question requires concept-level reasoning involving manifold knowledge sources. We conclude that ontological and qualitative spatial reasoning are important elements that should thus be addressed.

keywords: natural language processing, question answering, reasoning

1 Introduction

Our research is driven by prospects of navigation assistants that are able to interpret complex phrases, allowing the user to ask questions about routes, landmarks, destinations, etc. We expect to be facing the same challenges underlying spatial question understanding as in other application domains involving natural language interfaces to a Geographic Information System (GIS). These interfaces have to translate intuitively phrased questions to precise technical queries that a GIS can answer. GI science students have to learn performing this translation manually. Instructing GI science students and developing an autonomous system thus both requires a thorough understanding of underlying principles. In the following we argue that many intuitively phrased questions are ambiguous and it requires concept level reasoning, i. e. reasoning based on mental representations, even beyond Kuhn's core computations [1] to identify their intended interpretation.

In our research, we focus on understanding knowledge questions. These include those that access stored data in the geo-database ("What is x called?", "What is that?") and those that require a calculation in the sense of core computing ("How far is it to x?"). In our opinion, the core computations of a navigation system consist mainly of topological operations, e. g. "Where is a parking space within a region x?", as well as route/distance calculations, including their optimization (e. g. "near x").

2 Resolving Spatial References

For answering natural language questions, it is crucial to understand what is asked for. Spatial language is however very flexible and even a spatial relation term can have different, potentially not even spatial, interpretations. Additional context information may be required, too. For example, a question like “What is the name of the mountain ahead?” could interpret *ahead* with respect to location of the user and his or her direction of view or with respect to a planned route. In the following we outline main components of our approach aimed to resolve such relations.
First, a mostly shallow syntactic analysis of the given question is performed using Part-of-Speech Tagging, Named Entity Recognition, and Dependency Parsing. Second, semantic analysis is performed using a variety of Knowledge Bases (KBs) to explicate potential interpretations. Third, reasoning is applied to select the most plausible interpretation. Finally, the question can be answered. The last two steps closely relate to core computations with geographic knowledge

2.1 Syntactic Parsing

The first step in processing sentences that are already in text form is to reveal the linguistic structure of the question. A large number of different language processing tools and training corpora are available, but additional work is often required, particularly if not working with English language (cp. [2]). Parsers also tend to present a single interpretation where multiple interpretations are possible, potentially missing the intended structure. For example, “Where is Bamberg in Upper Franconia located on the river Regnitz?” may either be interpreted to ask for the whereabouts of Bamberg on river Regnitz, or within Upper Franconia which should be on the river Regnitz. Rather than embarking on parser improvement we propose to relax the output of existing parsers and to rely on semantic reasoning to resolve ambiguities. Technically, we propose to add additional interpretations in a rule-based manner, i. e., to instantiate additional spatial relation terms such as $\mathrm{on}(X, Y)$ for all close-by noun phrases.

2.2 Semantic Analysis

Question understanding aims to infer a user's demand for knowledge. Even when setting pragmatics underlying a question aside, question understanding may require a comprehensive body of background knowledge [3]. It remains an open question of how required knowledge can be provided to a question understanding system.

Machine Learning Application of machine learning faces at least two severe problems in natural language question understanding. First, there is no sufficient (freely available) corpus of domain-specific data that also provides potentially relevant context information. For example, to train interpreting “What is the name of the big mountain behind the city?” one would need to include the specific position location of the user stating the question, etc. Second, learnt models will only choose the statistically relevant interpretations. It is however not clear how a balanced set of training data should be compiled unless all dependencies underlying the desired interpretation(s) are known. For semantically ambiguous interpretations (e. g., whether "next city" refers to closest distance or succeeding location on a route), machine learning may easily lead to counter-intuitive results if training data is biased.

Reasoning To narrow down the linguistically possible interpretations of a sentence, we apply reasoning. Reasoning allows us to integrate all candidate interpretations and then to prune off those that are not jointly plausible. When processing a question like "How to get to Frankfurt Airport?" we aim for maximizing validity of the input, i. e., we search for possible interpretations of "Frankfurt" that also allow "Airport" to be resolved, discarding several villages also named "Frankfurt". We also try to assess the utility of potential answers. If the intention of asking for route instructions is to actually follow them, an in-car assistance system may disregard candidates from different continents. That is, we aim to consider task context in interpretation. But it also requires some background information about the user and airports in general to determine whether the user is referring to the major airline hub near Frankurt (Main), or a local airfield near Frankfurt (Oder), another major German city named Frankfurt. A clear advantage of applying reasoning over machine learning is that in cases where ambiguity cannot be resolved, alternative interpretations are available and can be presented to the user in a dialogue. Reasoning has to take into account that human spatial thinking is not based on precise metric information, but on qualitative, non-veridical models of the outside world as perceived and categorized by humans [4]. We are thus not aiming to apply classic logic inference, but propose to maximize validity on a continuous scale.

2.3 Background Knowledge

Formalization of background knowledge for empowering reasoning is a long-standing problem which, in the area of GIS, goes back at least to Egenhofer and Mark's "Naive Geography" program [5], which explores the represention of the commonsense geographic world with formal models. A problem with these formalizations, however, is that they appear so simple that they are not regarded as "science" by different communities [5].

Today, much background knowledge can be found in KBs and readily be used. With this, different possible meanings of a word can be discovered. A main issue with identifying concrete semantics of spatial relations is that there exists a complex relationship between the domains of time and space at the conceptual level. Prepositions like "before" or "after" can be used in both domains, and there are several spatial interpretations depending on the frame of reference (cp. [6, 7]) to start with. Looking up a noun in a KB we are able to appoint a certain type and characterization of that type. Currently, we are working on an integration of Cyc[1], WordNet[2] (respectively GermaNet[3] for German), and DBPedia[4]. WordNet/GermaNet supply us with semantics of words on a concept level, whereas Cyc provides formalizations of commonsense knowledge and basic concepts. To evaluate interpretation in context of potential answers

[1] `http://www.cyc.com`
[2] `https://wordnet.princeton.edu/`
[3] `http://www.sfs.uni-tuebingen.de/GermaNet/`
[4] `http://wiki.dbpedia.org`

we relate to further factual KBs, primarily DBPedia and OpenStreetMap[5] as well as GeoNames Ontology[6].

2.4 Core Reasoning

While we argue that manifold sources of information are required to understand questions in the realm of navigation assistance, we expect a rather small collection of reasoning techniques to be sufficient in many cases. As discussed above, we are not aiming for a Boolean logic approach, where we only get "true" or "false" as result, but for continuous truth semantics to improve robustness. On the technical level, we aim to apply qualitative reasoning [8] in conjunction with simple ontological reasoning (e. g. to infer class relationships). Since spatial and temporal knowledge comprises a great variety of aspects, a large repertoire of qualitative spatial reasoning techniques has been developed [9].

Let us outline how qualitative reasoning fosters interpretation of the question "Can I drive my truck through this underpass or is it too low?". The system has to reason, if the user asks whether the truck or the underpass is too low. For humans interpretation is clear due to their background knowledge. For a computer system, this knowledge has to be encoded, otherwise both interpretations appear plausible. Hence, a definition under$(truck, underpass)$ must constrain the height of the truck to be lower than the headroom of the underpass. From that definition it follows that a truck cannot be "too low" to inhibit passing the underpass, yet the headroom (a height measure associated with an underpass) can be too low. It follows that only the latter interpretation yields a question for which an informative answer can be computed.

3 Conclusion

We agree with Egenhofer and Mark's [5] view that mastering Naive Geography is essential for intelligent user applications. Grasping fundamentals underlying such intelligent GIS interfaces is not only crucial for the development of these systems, but is also crucial for GI science education. While we refer to navigation systems, this application can be seen as a microcosm for all knowledge questions in question answering GIS interfaces. The implicit location references that occur in a discourse context are comparable to the location reference by the position of the user in a navigation system. We regard natural language question understanding to offer good means for studying how intuitive questions map to concrete computations.

In this paper we argue that core computations required for grasping semantics of natural language questions comprise logical reasoning; at least ontological and qualitative reasoning techniques over temporal, spatial, and spatio-temporal instances. Computations also rely on a comprehensive record of background knowledge. As a consequence, core computations in the realm of GIS should include cognitively motivated reasoning techniques such as qualitative spatial reasoning.

[5] `https://www.openstreetmap.org`

[6] `http://www.geonames.org/ontology/`

References

[1] W. Kuhn, "Core concepts of spatial information for transdisciplinary research," *Journal International Journal of Geographical Information Science*, vol. 26, no. 12, pp. 2267–2276, 2012, reflections on Geographic Information Science: special issue in honor of Michael Goodchild.

[2] E. Giesbrecht and S. Evert, "Is part-of-speech tagging a solved task? an evaluation of pos taggers for the german web as corpus," *Proceedings of the 5th Web as Corpus Workshop (WAC5)*, pp. 27–35, 2009.

[3] E. Davis and G. Marcus, "Commonsense reasoning and commonsense knowledge in artificial intelligence," *Communications of the ACM*, vol. 58, no. 9, pp. 92–103, 2015.

[4] K. R. Coventry, *Spatial language and dialogue*, 1st ed., ser. Explorations in language and space. Oxford: Oxford Univ. Press, 2009.

[5] M. J. Egenhofer and D. M. Mark, "Naive geography," in *Spatial Information Theory*, A. U. Frank, W. Kuhn, M. J. Egenhofer, and D. M. Mark, Eds. Springer Berlin Heidelberg, 1995.

[6] T. Tenbrink and W. Kuhn, "A model of spatial reference frames in language," in *Spatial Information Theory, Proceedings of COSIT 2011*, ser. Lecture Notes in Computer Science, M. Egenhofer, N. Giudice, R. Moratz, and M. Worboys, Eds., vol. 6899. Springer, 2011.

[7] S. C. Levinson, "Frames of reference and molyneux's question: Crosslinguistic evidence," in *Language and Space*, P. Bloom and M. Peterson, Eds. Cambridge (MA), USA: MIT press, 1996, pp. 109–169.

[8] K. D. Forbus, "Qualitative reasoning," in *CRC Handbook of Computer Science and Engineering*. CRC Press, 1996.

[9] F. Dylla, J. H. Lee, T. Mossakowski, T. Schneider, A. van Delden, J. van de Ven, and D. Wolter, "A survey of qualitative spatial and temporal calculi – algebraic and computational properties," *ACM Computing Surveys (CSUR)*, vol. 50, no. 1, 2017, article no. 7, 53 pages.