Qualitative Spatial Reasoning over Questions

Mohammad Kazemi Beydokhti[1]
Department of Geospatial Science, RMIT University, Melbourne, Australia

Matt Duckham
Department of Geospatial Science, RMIT University, Melbourne, Australia

Yaguang Tao
Department of Geospatial Science, RMIT University, Melbourne, Australia

Maria Vasardani
Department of Geospatial Science, RMIT University, Melbourne, Australia

Amy Griffin
Department of Geospatial Science, RMIT University, Melbourne, Australia

Abstract

Although geospatial question answering systems have received increasing attention in recent years, existing prototype systems struggle to properly answer qualitative spatial questions. In this work, we propose a unique framework for answering qualitative spatial questions, which comprises three main components: a geoparser that takes the input questions and extracts place semantic information from text, a reasoning system which is embedded with a crisp reasoner, and finally, answer extraction, which refines the solution space and generates final answers. We present an experimental design to evaluate our framework for point-based cardinal direction calculus (CDC) relations by developing an automated approach for generating three types of synthetic qualitative spatial questions. The initial evaluations of generated answers in our system are promising because a high proportion of answers were labelled correct.

2012 ACM Subject Classification Theory of computation → Semantics and reasoning

Keywords and phrases Qualitative spatial reasoning, geospatial question answering, Qualitative spatial questions

Digital Object Identifier 10.4230/LIPIcs.COSIT.2022.18

Category Short Paper

Supplementary Material *Software (Source Code)*: `https://github.com/MohammadUT/QSR-QA`
archived at `swh:1:dir:fbe89a2479977a64c2fb15a1a10a7592fe3bd1ab`

1 Introduction

Qualitative spatial reasoning (QSR) in knowledge representation and reasoning (KR) deals with knowledge about the discrete, imprecise, and non-numerical properties of space and time. Humans' common-sense understanding of space is more connected with the concept of qualitative reasoning, as opposed to quantitative reasoning. Qualities are conceptually simpler than quantities (e.g., "tall" versus 1.93m), they can be obtained from quantities (e.g., bearing of 274∘ is "left"), and they generally correspond to discontinuities that are salient to humans (e.g., "left" versus "right," "in front" versus "behind") [4, 10].

Recently, a thorough classification of GeoQA systems based on the type of questions they can answer has been proposed [7]. Answering natural language qualitative spatial questions, which is the focus of the current study, has been studied using approaches that are mainly based on retrieving answers over linked geospatial data sources, such as DBpedia, GADM,

[1] Corresponding author

and OSM [8, 11]. In this paper, we go beyond these approaches by proposing a qualitative spatial reasoning framework for answering such questions, considering certain evidence in qualitative spatial scenes. For example, in the question 'Is A left of C?', if the supporting evidence is certain, then the qualitative spatial inferences may be certain (e.g., if A is left of B and B is left of C, then A is certainly left of C).

In this work, we propose a framework for answering qualitative spatial questions in which a reasoning system is integrated into a GeoQA system. This framework has three main components, including: 1) A semantic geoparser which gets the natural language questions and returns the place entities and spatial relationships in a triple format. 2) A reasoning system which reasons based on an evidence database. 3) Answer extraction which post-processes the generated answers from the reasoning system to return the best possible solutions. In addition, we provide an automated approach to generate three types of synthetic qualitative spatial questions, including Finding Relations (QType-1), Finding Places (QType-2), and Yes/No (QType-3). To make useful inferences over the queries, this study assumes that enough evidence is fed to the reasoning engine.

2 Related works

Expressions of spatial relations serve as the basis of reasoning on spatial data. Researchers have proposed extensive qualitative spatial calculi on the expression of spatial relations. A thorough survey of qualitative spatial and temporal calculi along with their computational properties has been developed by [3]. The inherently imprecise nature of many types of spatial information can be represented and modelled with qualitative spatial reasoning. Qualitative spatial calculi provide the basis for automated qualitative spatial reasoning with complex spatial scenes. For example, the RCC represents the topological relation between two regions without needing their precise quantitative locations or geometries.

Different approaches have been proposed for answering qualitative spatial questions. Early work answered three classes of qualitative and quantitative questions, including proximity, containment, and crossing, by reasoning over DBpedia [11]. This work was further developed by [8], who developed a GeoQA architecture with three main components: identifying the required instances from each question in their own proposed benchmark dataset, translating each formulated question into a SPARQL/GeoSPARQL query, and finally executing each correctly generated query over the linked geospatial data sources. Their proposed system did not perform well as it could only correctly answer 22% of all questions.

Extracting geospatial information that is widely used in queries (e.g., place names, spatial relationships, and place types [9, 6] from unstructured natural language text is one of the main reasons for analyzing questions. [5] proposed a semantic encoding approach in which various semantics such as place names and spatial relationships, among others, are extracted from natural text. This approach, derived from part-of-speech tagging and pretrained named entity recognition (NER) models in the AllenNLP library, has performed well for extracting all their considered entities except events, which are completely missing.

In this study, we integrated the above components into one framework that takes qualitative spatial questions as inputs, extracts the spatial information from the question in the form of triples, performs qualitative spatial reasoning under certain situations, and finally, generates the best possible answer(s) to each question.

3 Methodology: Experimental design

Synthetic question dataset

A question corpus is important to this research because it serves two main purposes: 1) understanding what types of questions are asked and what are their typical answers; and 2) serving as a gold standard for QA system evaluation. The ideal question corpus to use in this study should satisfy two conditions: having a sufficient number of qualitative spatial questions and having an accompanying supporting evidence database including spatial relations between spatial entities in the questions.

In order to prepare a question corpus that met our criteria, we developed an automated approach to generate three types of synthetic qualitative spatial questions: asking for the spatial relation between two features (Q-Type1), asking for feature(s) of a given class that have a particular spatial relation with another feature (Q-Type2), and asking whether a feature has a spatial relation with another feature (Q-Type3) (Table 1). These are three basic and widely used categories that have been also discussed in previous research by [8]. In our question-generation approach, 500 place names in and around the Melbourne CBD area are selected from the Gazetteer of Australia database[2]. Next, the recursive algorithm randomly selects a small number of key places and checks whether the points are well distributed. It is important to have a well distributed set of points as it is more in accordance with the distribution of the rest of the points that we are going to find in our answer.

To generate our corpus, the algorithm uses four randomly selected places, test whether they are well distributed, and if so, generates a simulated query for each question pattern following its corresponding general question template presented in Table 1. For the sake of testing our system, 1000 synthetic queries based on the CDC qualitative spatial logic [1] were generated for each question type.

Table 1 General question template for each question type, illustrated with an example.

Question type	Question template	Example
Q-Type1	What is the spatial relation between <*feature*> and <*feature*>?	What is the spatial relation between London and Manchester?
Q-Type2	Which features of type <X> are <*spatial relation*> of <*feature*>	Which county is east of county Dorset?
Q-Type3	Is <*feature*> <*spatial relation*> of <*feature*>?	Is Hampshire north of Berkshire?

Supporting evidence database

The qualitative spatial reasoning process described in this paper considers a crisp reasoning scenario, where we have certain questions and certain evidence sets, but the answers could be uncertain. In this case where the relations between features are exactly known, we store all the CDC relations between all places in each configuration using the qualification process proposed in [10]. The certain evidence database is finalized by applying the following post-processing steps: 1) Removing mutually inferable relations (composition relations). For example, if we have "A is north of B" and "B is north of C" relations, then the relation between A and C is not stored. 2) Removing converse relations. For instance, if we have the relation between A and B, then the relation of B and A is not stored. 3) Removing any relations that we want to infer in the questions. For example, in the question "Is A southeast of B", the relation between A and B is not included in the evidence database.

[2] `https://placenames.fsdf.org.au/`

Extracting spatial semantics from text

Here, we extract the required spatial information from questions and then presented them in a triple format, including geographical objects and the spatial relation between them. Place names and CDC relations are the required semantics that need to be extracted for each question. There are several available pre-trained NER models developed by open-source NLP libraries. We used the BERT-based model from the DeepPavlov library, which can recognize up to 19 different entities [2].

Qualitative spatial reasoning

To reason over the synthetic queries and make inferences for each of them, we used the open-source SparQ toolbox. for the crisp reasoning scenario when both questions and evidence sets are certain. As the SparQ shell commands are directly callable and executable in Python, this allowed us to integrate it with our GeoQA system. The SparQ reasoning system infers any possible relation among each pair of entities in each configuration when their relations are not known in the evidence database.

Answer extraction

The inference results in the previous step need to be refined in order to identify the best possible answer(s) for each question pattern. In addition, this stage enables us to better evaluate the results of our system for each question type with the benchmark question corpus containing actual answers. For QType-1 where the two place names that we want to infer are mentioned in the questions, we extract the generated relations between the place names in SparQ. For QType-2 where one place name and the cardinal relation are known in the question, all possible relations between the known place name and each of the key place names in each configuration are generated by SparQ. Then, any place names contained in the known relation in the question are retrieved. For QType-3, where two place names and their cardinal relation are known in the question, we extract the generated relations between the place names in SparQ and then check whether the known relation in the question is in agreement with them, and if yes, it returns 'YES', and vice versa.

4 Results and discussions

Because we have stored all the relations between place names for each configuration in the evidence database, answering any of the three types of qualitative spatial questions for the place names in the database is an information retrieval (IR) task, which is not in line with the goal of this work, making inferences over qualitative spatial queries. To address this, we add a new place name (Place P) in the study area which does not exist in our place name database. In addition, its cardinal relation is only known to one or two existing place names in each configuration and no information is available about the relation of other place names with Place P. The aim then is to infer these unknown relations based on the known evidence database. By including the new Place P, all of the simulated questions require a qualitative spatial reasoning process in order to answer them, and the answers are not directly retrievable from the evidence database. These simulated questions are freely available in our GitHub repository[3].

[3] `https://github.com/MohammadUT/QSR-QA`

Next, we extract place names and spatial relations from text and then structure them in a triple format as *(place name 1, place name 2, relation)* and do the following analysis. We refer to this triple as a 'query sentence' in equation 1. The evaluation of the semantic encoding tool's accuracy is conducted at the sentence level and is defined as the ratio of the number of correctly extracted query sentences from each question to the total number of questions.

Here, we only address the accuracy for QType-1 and QType-3 questions, because we do not have any place names in QType-2 because that question type is about finding places that are in a known spatial relation with the unknown Place P. We used the NER model to extract a query sentence for each question, then matched them with true labels to measure the extraction accuracy. Our model performs well for correctly detection of query sentences, where the accuracies obtained for QType-1 and QType-3 are 79.10% and 82.30%, respectively.

SparQ requires the certain evidence sets for each configuration from the evidence database as well as the extracted query sentence from the NER models to generate possible answers to each qualitative spatial question in each question pattern. We would expect certain and uncertain answers from SparQ. The excerpt outputs of SparQ for the first five configurations are provided in Tables 3 to 5 for QType-1, QType-2, and QType-3, respectively. For the sake of simplicity, place IDs have been utilized instead of their corresponding place names.

Table 2 shows that for QType-1, SparQ predicted all possible relations for the requested query sentence with equal level of importance. For example, the generated relations between places 499 and 501 in C4 could be south, southeast, or southwest and no priority ordering has been made here. By comparing these results with the corresponding actual answer, we get southeast, meaning the correct answer is a subset of the answers identified by SparQ. Blank rows indicate that the NER model was unable to successfully extract the query sentence, which results in null values for the reasoning step of the methodology as well.

Table 2 Inference results for the first five configurations of QType-1 for the SparQ system along with their actual answers and extracted query sentences.

Configurations	Extracted query sentence	SparQ inferences	Actual answers
C1	(18, 501,?relation)	(18, 501, E EQ N NE NW S SE SW W)	(18, 501, NW)
C2	–	–	(415, 501, SE)
C3	–	–	(343, 501, SE)
C4	(499, 501,?relation)	(499, 501, S SE SW)	(499, 501, SE)
C5	(484, 501,?relation)	(484, 501, E EQ N NE NW S SE SW W)	(484, 501, NW)

Considering the results of QType-2 in Table 3, SparQ generates all possible places that could be in a specific relation with the unknown place 501. Taking the C2 configuration as an example, 415 and 360 are possible places inferred from SparQ that could be southeast of 501.

Based on the QType-3 results presented in Table 4, in some cases, SparQ infers all possible relations for each extracted query sentence. For example, to check whether 365 is northwest of 501 in C4, SparQ infers with north, northeast, and northwest relations. Cross checking these outputs with 'YES' as the actual answer, the final answer from the reasoner could be YES, as northwest is one of the generated relations.

In the final stage, we evaluated the accuracy of the generated answers obtained from the SparQ system for the three question patterns in terms of their closeness to the actual answers. To accomplish this, we have defined three categories to characterize the correctness

Table 3 Inferred places from SparQ for the first five configurations of QType-2, along with the actual answers and the extracted query sentence.

Configurations	Extracted query sentence	SparQ inferences	Actual answers
C1	(?Places, 501,NW)	(18, 501, NW) (379, 501, NW)	(18, 501, NW)
C2	(?Places, 501,SE)	(415, 501, SE) (360, 501, SE)	(415, 501, SE)
C3	(?Places, 501, NW)	(343, 501, NW) (326, 501, NW)	(326, 501, NW)
C4	(?Places, 501,SE)	(499, 501, SE)	(499, 501, SE)
C5	(?Places, 501,SE)	(484, 501, SE) (343, 501, SE)	(343, 501, SE)

Table 4 I Inference results from the first five configurations of QType-3 for the SparQ system along with the actual answers and extracted query sentences.

Configurations	Extracted query sentence	SparQ inferences	Actual answers
C1	(379, 501, NE)	(379, 501, E EQ N NE NW S SE SW W)	(379, 501, NO)
C2	–	–	(360, 501, NO)
C3	(326, 501, SW)	(326, 501, E EQ N NE NW S SE SW W)	(326, 501, NO))
C4	(365, 501, NW)	(365, 501, N NE NW)	(365, 501, YES)
C5	–	–	(343, 501, NO)

of generated answers, including Correct, Incorrect, and Uninformative answers. A Correct answer is defined when the answer generated for each question is either a complete or partial match with the relevant actual answer. Taking C4 in Table 3 as an example, the SparQ inference results are labeled as correct as the SparQ-generated relations partially match with the actual relation. An Incorrect answer is tallied when the actual answer is neither a complete match with generated answer nor is found among the generated answers, for instance, when the SparQ inference result for two places is (Place1, Place 2, 'n ne nw') but its corresponding actual answer is se. Finally, an uninformative answer is those cases in which the generated answers do not provide any useful information about whether they are correct or incorrect. Considering C1 in Table 3 as an example, SparQ infers that the relation between the queried places could be any of the cardinal directions.

We represent the number of questions in each type that fall under the three answer correctness categories for the DeepPavlov-based NER model[4]. By considering Figure 3 in general, a high proportion of generated answers in all question types is labelled correct, meaning that the overall performance of each was acceptable. By focusing on each question pattern in particular, in QType-1, where the answers are relations between an unknown place and known places, SparQ generated answers that are either correct or uninformative, with no incorrect answers generated. For QType-2 where the places retrieved from the reasoners are checked with the corresponding actual places, SparQ answered all questions correctly. Finally, in QType-3 where the inferred YES/NO answer is checked with the actual answer, three forms of answers are found here, but the incorrect category contained the fewest answers. The overall performance of the system depends highly on the place semantics extraction step, which means the more questions from which we could successfully extract place semantics, the greater the chance that the reasoner will return a correct answer.

[4] https://github.com/MohammadUT/QSR-QA/blob/main/Figure%203.jpg

5 Conclusion and Future Work

This paper has addressed the problem of answering qualitative spatial questions by presenting a GeoQA system based on deductive spatial reasoning. To achieve this goal, this system begins by taking three types of qualitive spatial questions, extracting toponyms and spatial relations from the question text, applies crisp qualitative spatial reasoner to each question, and finally, generating final answers for each question type. To evaluate our system, we have compared the results obtained from all simulated question types with a benchmark including correct answers. The results have shown that our system performed well with in all three types of questions when the SparQ reasoner has been fed by a sufficient number of evidence sets. Initial results showed that there is the possibility of adapting our approach to addressing other spatial relation logics, a result of which is that more diverse types qualitative spatial questions can be answered.

References

1 UF Andrew, D Mark, and D White. Qualitative spatial reasoning about cardinal directions. In *Proc. of the 7th Austrian Conf. on Artificial Intelligence. Baltimore: Morgan Kaufmann*, pages 157–167, 1991.

2 Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint*, 2018. `arXiv:1810.04805`.

3 Frank Dylla, Jae Hee Lee, Till Mossakowski, Thomas Schneider, André Van Delden, Jasper Van De Ven, and Diedrich Wolter. A survey of qualitative spatial and temporal calculi: algebraic and computational properties. *ACM Computing Surveys (CSUR)*, 50(1):1–39, 2017.

4 Antony Galton et al. *Qualitative spatial change*. Oxford University Press on Demand, 2000.

5 Ehsan Hamzei, Haonan Li, Maria Vasardani, Timothy Baldwin, Stephan Winter, and Martin Tomko. Place questions and human-generated answers: A data analysis approach. In *International Conference on Geographic Information Science*, pages 3–19. Springer, 2019.

6 Mohammad Kazemi Beydokhti, Matt Duckham, Amy Griffin, and Vedran Kasalica. Geo-event question answering systems: A preliminary research study. In *Proceedings of the 11th International Conference on Geographic Information Science (GIScience 2021)*, page 6, 2021.

7 Gengchen Mai, Krzysztof Janowicz, Rui Zhu, Ling Cai, and Ni Lao. Geographic question answering: Challenges, uniqueness, classification, and future directions. *AGILE: GIScience Series*, 2:1–21, 2021.

8 Dharmen Punjani, Kuldeep Singh, Andreas Both, Manolis Koubarakis, Iosif Angelidis, Konstantina Bereta, Themis Beris, Dimitris Bilidas, Theofilos Ioannidis, Nikolaos Karalis, et al. Template-based question answering over linked geospatial data. In *Proceedings of the 12th Workshop on Geographic Information Retrieval*, pages 1–10, 2018.

9 Mark Sanderson and Janet Kohler. Analyzing geographic queries. In *SIGIR workshop on geographic information retrieval*, volume 2, pages 8–10, 2004.

10 Diedrich Wolter and Jan Oliver Wallgrün. Qualitative spatial reasoning for applications: New challenges and the sparq toolbox. In *Geographic Information Systems: Concepts, Methodologies, Tools, and Applications*, pages 1639–1664. IGI Global, 2013.

11 Eman MG Younis, Christopher B Jones, Vlad Tanasescu, and Alia I Abdelmoty. Hybrid geo-spatial query methods on the semantic web with a spatially-enhanced index of dbpedia. In *International Conference on Geographic Information Science*, pages 340–353. Springer, 2012.