A Corpus of Spatial Reasoning Problems

Adam Pease, Paulo Santos, and Alexandre Rademaker

[1] Articulate Software, San Jose, CA, USA
apease@articulatesoftware.com
[2] Flinders University, Adelaide, South Australia
paulo.santos@flinders.edu.au
[3] IBM and EMAp/FGV, Rio de Janeiro, Brazil
alexrad@br.ibm.com

1 Introduction

Spatial reasoning is an important part of common-sense reasoning but has most often been examined in isolation from other areas of common sense knowledge representation. Spatial reasoning also has been a prominent area of research in linguistic semantics. In this work, we aim to create a corpus of common sense reasoning problems formalized in the context of a large and comprehensive theory of world knowledge, in the hopes that it will be more generally reusable on a wide variety of practical reasoning problems. An additional goal is to demonstrate the computational sufficiency of this work by employing automated reasoning to solve the problems. Finally, we also state the problems first in natural language, to provide a set of test cases for theories of linguistic semantics and computational linguistics, with a computational representation that can help to validate whether any method or process of interpreting language into a computational logic representation is correct or sufficient.

This is an ambitious set of goals, and it will not be possible to provide solutions for all the problems that could be formulated in natural language. In particular, we do not attempt to create an implementation, at least at this time, for a system that can translate all of the stated problems from language to logic. Many of the problems also appear to require a logic and reasoning system beyond first order logic. It is important however, to have challenges that are unsolved in order to motivate research.

We utilize the Suggested Upper Merged Ontology (SUMO)[5, 7][1], a comprehensive ontology of around 20,000 concepts and 80,000 hand-authored logical statements in a higher-order logic, that has an associated integrated development environment[9] integrated with leading theorem provers such as Eprover [10] Vampire [4] and LEO-II [1], and manually-created links[6] to the WordNet lexico-semantic database[3]. We described [9] elsewhere how to translate SUMO to the strictly first order language of TPTP [12], as well as TF0 [8] and THF[2].

2 Reasoning Problems

The current set of problems with solutions is available on GitHub[2]. Many of the problems are simple in representation and reasoning. For example

(26) *The road goes from MV to MP. Does the road go from MP to MV?* Yes.

We have the notion of a `BidirectionalTransitway` and Vampire easily solves the problem as formulated in SUO-KIF with SUMO terms in 16 steps, most of which are transformation

[1] https://www.ontologyportal.org
[2] https://github.com/ontologyportal/sumo/blob/master/tests/SpatialQs.txt

of the required axioms into conjunctive normal form. But an NLP system must interpret that 'go' means the road traverses a path, rather than moves, which in this case we have done by manually creating the SUMO-based formalization.

(4) *John is carrying a vase. There is a flower in the vase. Is John carrying a flower?* Yes

Problem 4 is solved in 26 steps and using four axioms from SUMO and the problem statement. In each case, we have common formalized terms to reuse from SUMO. This ensures that each problem has a degree of compatible and comparable semantics to anchor the semantics of terms. In some problems, the primary challenge is linguistic, as in

(37) *The painting went from the first to the second floor. Did the painting move?* Most likely.

where once it is decided that the painting actually moved as opposed to spanning two floors on an atrium wall, for example, the problem is easily formalized and solved in 15 steps by Vampire.

In other problems such as (28) below, although the statements can easily be formulated in SUO-KIF/SUMO (following the question, below), the challenge is performing higher-order reasoning with modals or temporal qualification of situations, using SUMO's `holdsDuring` construct. One can also reasonably have different pragmatic interpretations of the text depending upon whether one captures an implied legal prohibition (that one `holdsObligation` as a law-abiding driver not to use the road), or a practical one (maybe there's an actual gap in the bridge being fixed), and for what vehicles (they're tearing up the pavement but a motorcycle or bicycle could still get by if ridden by a permitted member of the road crew). That is an additional interesting area of study.

(28) *From Monday to Friday the Bay Bridge will be unusable from Yerba Buena Island to Oakland. Will you be able to use the road to go from Yerba Buena to Oakland between Monday and Friday?* No.

```
(=>
  (holdsDuring ?TIME
    (attribute ?T TransitwayClosed))
  (exists (?P)
    (holdsObligation ?P
      (holdsDuring ?TIME
        (not
          (exists (?TP)
            (and
              (instance ?TP Translocation)
              (located ?TP ?T)))))))
```

We hope that the range of different challenges for NLP, representation and reasoning, within a common ontology and tool set helps to unify and motivate some separate threads of research. For instance, the formalization of such spatial problems could aid the evidence analysis in digital forensics, and the combination of spatial reasoning capabilities could take Visual Question Answering Challenges from the current simple pattern recognition in images to the actual interpretation of scenes and the actions and actors depicted in them. Additional contributed problems (and solutions) are welcome. We expect that it should be possible to add some of these problems to the yearly CASC [11] competition as well.

We thank Dr. Annie Zaenen for contributing a number of the problems in the corpus.

References

[1] Christoph Benzmüller, Laurence Paulson, Frank Theiss, and A. Fietzke. (2008). *LEO-II - A Cooperative Automatic Theorem Prover for Higher-Order Logic. In Proceedings of the Fourth International Joint Conference on Automated Reasoning (IJCAR'08), LNAI volume*, 5195:162–170, 2008.

[2] Christoph Benzmüller and Adam Pease. Progress in automating higher-order ontology reasoning. In Boris Konev, Renate Schmidt, and Stephan Schulz, editors, *Workshop on Practical Aspects of Automated Reasoning (PAAR-2010)*. CEUR Workshop Proceedings, Edinburgh, UK, 2010.

[3] Christiane Fellbaum, editor. *WordNet: An Electronic Lexical Database*. MIT Press, 1998.

[4] Laura Kovács and Andrei Voronkov. First-order theorem proving and vampire. In *Proceedings of the 25th International Conference on Computer Aided Verification*, volume 8044 of *CAV 2013*, pages 1–35, New York, NY, USA, 2013. Springer-Verlag New York, Inc.

[5] Ian Niles and Adam Pease. Toward a Standard Upper Ontology. In Chris Welty and Barry Smith, editors, *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001)*, pages 2–9, 2001.

[6] Ian Niles and Adam Pease. Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In *Proceedings of the IEEE International Conference on Information and Knowledge Engineering*, pages 412–416, 2003.

[7] Adam Pease. *Ontology: A Practical Guide*. Articulate Software Press, Angwin, CA, 2011.

[8] Adam Pease. Arithmetic and inference in a large theory. In *AI in Theorem Proving*, 2019.

[9] Adam Pease and Stephan Schulz. Knowledge Engineering for Large Ontologies with Sigma KEE 3.0. In *The International Joint Conference on Automated Reasoning*, 2014.

[10] S. Schulz. E – A Brainiac Theorem Prover. *Journal of AI Communications*, 15(2/3):111–126, 2002.

[11] Geoff Sutcliffe and Christian Suttner. The state of casc. *AI Commun.*, 19(1):35–48, January 2006.

[12] Steven Trac, Geoff Sutcliffe, and Adam Pease. Integration of the TPTPWorld into SigmaKEE. In *Proceedings of IJCAR '08 Workshop on Practical Aspects of Automated Reasoning (PAAR-2008)*. CEUR Workshop Proceedings, 2008.