

Some Connections between Qualitative Spatial Reasoning and Machine Learning

Anthony G Cohn^{1,2,*},†

¹School of Computer Science, University of Leeds, LS2 9JT, UK

²The Alan Turing Institute, UK

Abstract As has been remarked on before, *Space is Special*[1, 2]. Tobler’s First Law of Geography [3] captures the notion that all things are related, but close things are more related. Tversky [2] eloquently argues for the special place for spatial representations, and in particular that (living) things must move and act in space to survive, that all thought begins as spatial thought and that spatial thinking comes from and is shaped by perceiving the world and acting in it, be it through learning or through evolution. Artificial Intelligence has thus naturally sought to endow artificial agents with spatial representations and ways of reasoning about space. Amongst these, I will focus on qualitative spatial representations and reasoning mechanisms (henceforth QSR, where the ‘R’ may stand for representation or reasoning or both, depending on the context). There have been many calculi developed for representing and reasoning about space in qualitative ways, covering aspects such as (mereo)topology, orientation/direction, size, distance and shape [4, 5]. Whilst QSR has primarily been concerned with deductive reasoning, there have been and there are increasingly many connections between QSR and machine learning. In this talk I will discuss a number of such connections, ranging from the use of qualitative spatial representations in an inductive logic programming system to learn event classes occurring in video data, to the question of whether large language models (LLMs) are able to make inferences reliably about qualitative spatial relations, and whether they can be supported by symbolic reasoners.

Learning rules for video interpretation: Dubba et al. [6] show how Inductive Logic Programming can be used to learn a set of rules which can be used to recognise event class instances where videos have been abstracted to a set of qualitative spatio-temporal relations. The method is demonstrated in two domains including one which involves recognising the events which are necessary to service an aircraft whilst it is turning around at an airport. Whilst the resulting rules are relatively simple and it might be wondered whether a hand-written set of rules could not be easily written and just as effective, it turns out that in a comparison with such a set of manually written rules, the learned model is more effective, because the latter does not take account of noise in the video data, where as the learned model was already trained on noisy data and was thus more robust in the face of noisy data at classification time. The paper also shows how the inductive process can be interleaved with abduction, using an embedded spatial theory to improve the learned model in the face of noisy training data.

Learning groundings for spatial representations: A key question for QSR is how the relations in the calculus correspond to their use in language and their correspondence to the real world. Whilst relations are usually given plausible names in a relational calculus, there is no guarantee that these correspond to naturally occurring instances. Indeed, McDermott [7] notes the dangers of “wishful naming”. Alomari et al. [8] present a system, named OLAV, which addresses the problem of bootstrapping knowledge in language and vision for autonomous robots. OLAV is able, for the first time, to (1) learn to form discrete concepts from sensory data; (2) ground language (n-grams) to these concepts (which include not only spatial relations, but also object attributes and actions); (3) induce a grammar for the language being used to describe the perceptual world; and moreover to do all this incrementally, without storing all previous data. The resulting grammar can then be used to parse novel commands for downstream action in a robotic system.

Analysing polysemy in spatial prepositions: One challenge in assigning meanings to spatial prepositions is that they can frequently be polysemous, i.e. they can have multiple related senses (the polysemes). As the senses of polysemous terms are so closely intertwined, the theoretical and computational treatment of polysemy presents a difficult challenge for semantic models. To give an example: compare “book on a table”, “balloon on the ceiling” and “picture on the wall”. Richard-Bollans et al. [9] discuss this problem and shows how a model can be built in which these senses can be distinguished using data from human subjects.

Can Large Language Models perform qualitative spatial reasoning reliably? Many claims (e.g. [10, 11, 12]) have been made since the emergence of Large Language Models (LLMs) as to their ability to reason. Spatial

STRL’24: Third International Workshop on Spatio-Temporal Reasoning and Learning, 5 August 2024, Jeju, South Korea

✉ a.g.cohn@leeds.ac.uk (A. G. Cohn)

🌐 <https://tinyurl.com/A-G-Cohn> (A. G. Cohn)

🆔 0000-0002-7652-8907 (A. G. Cohn)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

reasoning is of particular interest since not only does it underlie a human’s ability to operate in the physical world, but also because LLMs are not embodied; so the question arises, have they nonetheless acquired an ability to reason about situations which might occur in the real physical world? I will present the results of a number of experiments in which this ability is tested: for (cardinal)Michael Sioutis <michael.sioutis@lirmm.fr> directions [13, 14], for relational composition and conceptual neighbourhood construction [15] and other notions in spatial reasoning [16]. One challenge for evaluating LLMs in the domain of spatial reasoning (and commonsense more generally [17]) is the paucity of good benchmarks – I will discuss this issue and briefly present a new benchmark which is based on a synthetic generator, able to provide arbitrarily many examples of automatically labelled indoor virtual scenes[18].

Using LLMs as a natural language interface to symbolic spatial reasoners: Given the deficiencies in the robustness of LLMs in performing qualitative spatial reasoning, it is worth asking the question whether an LLM and a more traditional symbolic reasoner in combination could be more effective than either on their own. An LLM has strengths in analysing language, but not so much in more complex reasoning, whilst an LLM on its own has no ability to comprehend natural language. The combination of the two can be particularly effective, for example as demonstrated in the StepGame benchmark [19, 14].

Acknowledgements This work was supported by: the Fundamental Research priority area of The Alan Turing Institute; Microsoft Research - Accelerating Foundation Models Research program; the Economic and Social Research Council (ESRC) under grant ES/W003473/1. I also wish to give heartfelt thanks to all my co-authors in the papers [6, 8, 9, 14, 16, 18] I will discuss in the talk, and with whom it has been such a pleasure to interact with.

References

- [1] M. F. Goodchild, Challenges in geographical information science, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 467 (2011) 2431–2443.
- [2] B. Tversky, *Mind in Motion: How Action Shapes Thought*, Basic Books, 2019.
- [3] H. J. Miller, Tobler’s first law and spatial analysis, *Annals of the association of American geographers* 94 (2004) 284–289.
- [4] A. G. Cohn, J. Renz, Qualitative spatial representation and reasoning, in: F. Van Harmelen, V. Lifschitz, B. Porter (Eds.), *Handbook of knowledge representation*, Elsevier, 2008, pp. 551–596.
- [5] J. Chen, A. G. Cohn, D. Liu, S. Wang, J. Ouyang, Q. Yu, A survey of qualitative spatial representations, *The Knowledge Engineering Review* 30 (2015) 106–136.
- [6] K. S. Dubba, A. G. Cohn, D. C. Hogg, M. Bhatt, F. Dylla, Learning relational event models from video, *Journal of Artificial Intelligence Research* 53 (2015) 41–90.
- [7] D. McDermott, Artificial intelligence meets natural stupidity, *ACM SIGART Bulletin* (1976) 4–9.
- [8] M. Alomari, F. Li, D. C. Hogg, A. G. Cohn, Online perceptual learning and natural language acquisition for autonomous robots, *Artificial Intelligence* 303 (2022) 103637.
- [9] A. Richard-Bollans, L. G. Álvarez, A. G. Cohn, Identifying and modelling polysemous senses of spatial prepositions in referring expressions, *Cognitive Systems Research* 77 (2023) 45–61.
- [10] A. Creswell, M. Shanahan, Faithful reasoning using large language models, 2022. [arXiv:2208.14271](https://arxiv.org/abs/2208.14271).
- [11] J. Huang, K. C.-C. Chang, Towards reasoning in large language models: A survey, 2023. [arXiv:2212.10403](https://arxiv.org/abs/2212.10403).
- [12] T. Kojima, S. S. Gu, M. Reid, Y. Matsuo, Y. Iwasawa, Large language models are zero-shot reasoners, *Advances in neural information processing systems* 35 (2022) 22199–22213. [arXiv:2205.11916](https://arxiv.org/abs/2205.11916).
- [13] A. G. Cohn, R. E. Blackwell, Evaluating the ability of large language models to reason about cardinal directions, in: *Proc. COSIT-24 (to appear)*, [arXiv:2406.16528](https://arxiv.org/abs/2406.16528), 2024.
- [14] F. Li, D. C. Hogg, A. G. Cohn, Advancing spatial reasoning in large language models: An in-depth evaluation and enhancement using the StepGame benchmark, in: *Proc. AAAI*, 2024.
- [15] A. G. Cohn, An evaluation of ChatGPT-4’s qualitative spatial reasoning capabilities in RCC-8, 2023. URL: <https://arxiv.org/abs/2309.15577>. [arXiv:2309.15577](https://arxiv.org/abs/2309.15577), appears in *Working Notes of QR-23*.
- [16] A. G. Cohn, J. Hernandez-Orallo, Dialectical language model evaluation: An initial appraisal of the commonsense spatial reasoning abilities of LLMs, *arXiv preprint arXiv:2304.11164* (2023).
- [17] E. Davis, Benchmarks for automated commonsense reasoning: A survey, *ACM Computing Surveys* 56 (2023) 1–41.
- [18] F. Li, D. C. Hogg, A. G. Cohn, Reframing spatial reasoning evaluation in language models: A real-world simulation benchmark for qualitative reasoning, in: *Proc. IJCAI*, 2024.
- [19] Z. Shi, Q. Zhang, A. Lipani, StepGame: A new benchmark for robust multi-hop spatial reasoning in texts, in: *Proc. AAAI*, volume 36, 2022, pp. 11321–11329.