Beyond spatial reasoning: Challenges for ecological problem solving

Christian Freksa

Bremen Spatial Cognition Center, University of Bremen, Germany

Abstract: This vision piece reflects upon virtues of early computer science due to scarcity and high cost of computational resources. It critically assesses divergences between real-world problems and their computational counterparts in commonsense problem solving. The paper points out the different objectives of commonsense versus scientific approaches to problem solving. It describes how natural cognitive systems exploit space and time without explicitly representing their properties and why purely computational approaches are less efficient than their natural role models, as they depend on explicit representations. We argue for investigating spatio-temporally integrated methods to spatial problem solving. We contrast these methods to sequential computational approaches that require *digital twins* of the environment and cannot make direct use of simultaneous spatio-temporal interactions. The paper concludes with predicting future developments in problem solving, praising the relative merits of different routes to be taken. It advocates the translation of fundamental cognitive principles into technical robotic solutions.

Keywords: artificial intelligence, cognitive system, commonsense problem solving, learning from nature, perception and action, spatial organization, spatial cognition, affordance, coarse to fine processing

1 Past

Once upon a time, when computers were big, their processor words short, their cycle times long, and their memory capacities tiny and expensive, computer scientists and artificial intelligence (AI) researchers strove for understanding the structural essence of problems to be solved. They investigated ways of solving problems in order to develop efficient and

effective algorithms that could be run on these computers within reasonable time frames. They asked questions like:

- Which are the essential pieces of information needed to solve the problem?
- What spatial and temporal resolution is required for tackling the problem?
- What precision is needed for the different attributes?
- Can we form clusters of concepts to operate on a more abstract level?
- Can we find meaningful categories rather than working on quantities too fine to be meaningful?
- Do we need to introduce absolute reference systems to obtain objective observations? or
- How does my algorithm scale with the size of the problem?

Such questions helped protagonists tremendously to understand problems in depth and to identify problem-adequate solutions. Cognitive scientists looked for inspiration from nature where similar problems seemed to be solved in smart ways. They learned about the environmental conditions in which problems need to be solved. Over the years, computers became cheaper and ubiquitous, their memory capacities and speed became enormous, the existential need to conserve resources disappeared, the virtues of ecological computing were forgotten, and the sense of the nature of a problem often got lost. We now can implement amazing neural networks and general purpose deep learning algorithms which generate solutions to big problems without deeply understanding them. Unfortunately, in many cases nobody knows to what extent the problems they solve in the computer correspond to the problems in the world that they were supposed to solve. One reason is that the appropriateness and correctness of the problem representation usually is not thoroughly scrutinized.

2   Present

2.1   Natural cognitive systems

We can learn much about smart approaches from nature by investigating spatial and temporal problem solving [14]. So, what is special about space and time? Natural cognitive agents (humans, animals, and even plants) first and foremost must deal with their spatio-temporal environments [1,7,16]. They benefit particularly from resources and social interactions in their vicinities, as these can be managed most economically in terms of perception, action, and energy. To get around in their environments and to make use of them, cognitive agents initially depend completely on local perception and action that provide information about the spatial and temporal organization of their surroundings [10]. They have neither an abstract model of the world nor ways of reasoning on the basis of such a model. Therefore these agents first of all develop a spatial and temporal understanding of the world.

The spatial conception of the world develops from coarse to fine, as initially there is no perception for fine distinctions and consequently there are no concepts for them [8]. In the course of interacting with the environment finer distinctions successively develop, for example, when a toddler discovers that not every four-legged moving object is a dog. In developing an understanding of spatio-temporal environments cognitive agents receive

strong structural support from their own anatomy as well as from the environment [18]. Our perception and motor organs as well as the neural maps connecting them are spatially organized such that near things in the perceived environment correspond to near things in perceptual organs (e.g., the retina) and often to near things in the mental representation of the cortex [9, 12].

This holds for most of our senses, including visual, auditory, haptic, olfactory, and gustatory sense. In other words, perception, action, and their neural representations are largely spatially organized. Through perception, action, and mental processes they also become temporally organized, with a strong correspondence between spatial and temporal structure [16]. As perception and action operate on the external world (the spatial environment), the spatial structure of the environment participates in the spatial organization of our internal representations of the perceived world; these, in turn, form the basis of concepts about the world including dynamics and interactions.

I put forward the hypothesis that concepts acquired through perception and action in the spatial world form the basis of more abstract concepts such as strength, consumption, accessibility, wealth, etc. that cannot be perceived directly but can easily be related to perceivable features. In this way, the spatial organization of our bodies and our environments forms the foundation of what we perceive, how we think, and how we act. A notion of time evolves as a direct correlate to spatial distance as our motion modalities ensure that nearby locations will be reached more quickly and directly than more distant locations, both in locomotion and perception.

With this design we have cognitive systems that are spatio-temporally organized and whose concepts develop from coarse to fine. Various perceptual relations are intimately related to one another and therefore are strongly intertwined. For example, in a 2-dimensional image of a 3-dimensional landscape scene, the vertical dimension represents height and depth information in a geometrically systematic way, distances and orientations are interrelated in systematic ways, time becomes accessible through speed which in turn is strongly related to distance and perceivable through its relation to sizes of moving objects. In other words, in the realm of our perception and actions, aspects of space and time are inherently connected; they become separated only when we focus our attention from integrated configurations and events to individual aspects which tend to be less meaningful in terms of our overall experience.

A highly interesting aspect of this kind of organization of cognitive systems is that they can operate on the world level early on by means of their aspectually integrated systems. They may not (yet) be able to distinguish strongly related notions like "amount", "height", and "volume" [11] and therefore may not be able to answer correctly, whether an amount of fluid has increased or decreased when poured from a narrow tall container into a wide flat container. But, nevertheless, they may be able to perform the pouring operation and obtain a valid result ("amount remains equal"). The valid result is guaranteed by space-inherent dependencies that cannot be violated in the real world. In such ways, integrated cognitive systems correctly solve spatial problems without representing all aspects of the world or understanding all aspects of the result.

In the process of learning about the world spatially organized systems can refine their concepts in a top-down manner, focusing on individual aspects of the integrated overall systems rather than on the wholes [4]. For example, they may distinguish between dogs and cats or between "amount", "height", and "volume". Such top-down processes from coarse to finer can proceed indefinitely, as new methods or new distinctions become avail-

able. They can yield new insights by connecting new distinctions to their established global world model.

2.2 Formal AI systems

Formal representations and conceptions of the world in AI are structured differently: they are copies of aspects of the world ("digital twins") designed to describe the world and solutions to problems from an external (scientist's) rather than from an internal or developmental perspective. A major discovery was that we can solve problems by reasoning about assertions about the world. This turns out to be such a powerful method not only to solving problems but also to proving the correctness of solutions (assuming the formal representation of the problems is correct!) that reasoning about problems became a universal approach to solving them. It is the standard approach of weak AI [13].

In reasoning successfully, we may have neglected that commonsense cognitive agents have different objectives than scientists. When humans (or animals) want to meet their friend, they want to establish certain spatio-temporal relations with the person, and are happy when they recognize and meet the friend. Commonsense agents do not verify a world model that leads to the conclusion that the person they encountered actually is the targeted friend (although this may be an observing scientist's objective). In other words, the commonsense agent's perspective is an internal perspective that focuses on the primary objective (establishing a relation); it does not care about how and why the target was met. On the other hand, the scientist's perspective is external and focuses on the validity of inference steps that lead to the final result; the established result itself may not have a particular significance. It is this scientist's perspective and not the commonsense reasoner's perspective that is implemented in typical AI "commonsense reasoning" systems.

Formal approaches to representing space start in an assumption-free abstract space in which we can introduce three independent (!) spatial dimensions and an additional independent temporal dimension. From the independent dimensions we build an absolute reference system that uses scales at a resolution that will cover the foreseen relevant distinctions. We represent atomic entities from which we construct compound systems in a bottom-up manner. A feature of this approach is that we may be able to derive properties of the compound from properties of its parts. Knowledge about dependencies between the compound elements (e.g., between dimensions) can be added as they become known or relevant for solving specific problems. But once the aspects to be represented and the reference system and scales to be used have been decided upon, it becomes difficult to conceptually modify or grow the system. In contemporary big data approaches, we also feed our systems with data collected under pre-selected reference frames and scales; demand-driven interactive data acquisition generally is not supported.

2.3 Towards cognitively oriented AI systems

Qualitative spatial and temporal reasoning (QSTR) [2] offers approaches to account for important aspects of natural cognitive systems in the framework of formal AI settings. In particular, QSTR supports internal reference frames, coarse-to-fine representation and processing, aggregation of neighboring relations, as well as adapting to problem-centered granularities (instead of using absolute reference frames). In this way, QSTR supports conceptual neighborhood [6] in horizontal (similarity) and vertical (abstraction) dimensions.

QSTR presents a first step towards problem solving on the basis of spatial structures and supports efficient spatial and temporal reasoning. It also points the way towards a more drastic exploration of spatial structures for spatial problem solving. In [5] we describe an approach that avoids the generation of digital twins in favor of employing spatially faithful structures that maintain inherent spatial relations and interactions rather than formally describing them. We named this approach "strong spatial cognition", as it fulfills essential requirements of strong AI [13].

3 Future

I predict that we will continue to virtualize real-world structures by building "digital twins" of everything we can get hold of. From a scientific point of view, this will be quite rewarding, as we will learn a lot about how our environments and our agents are structured and perceived. Solving problems by means of digital twins, however, requires much more computational effort than mental effort in solving comparable everyday problems, as each piece of knowledge must be made explicit and must be computationally addressed. Presently, we can estimate and compare the effort between natural and artificial problem solutions in terms of "intelligence per Kilowatt hour" [15, 17], but as research proceeds, we will learn more about structural differences concerning hard- and software in neural and artificial systems and we may become more aware of unwarranted power consumption for computational processing. This will eventually lead to technical problem-solving approaches that will integrate the spatio-temporal real world in a way similar to nature.

A highly significant scientific issue we have not really tackled is the fact that with current computer architectures "digital twins" operate on linear descriptions of space. These do not maintain essential spatial structures that make spatial problem solving effective and efficient. There are two ways of addressing this issue: (1) developing computer hardware that maintains interacting spatial relations according to affordances and constraints as individual relations change (it appears conceivable that human minds incorporate such structures capable of modest imagery processes as part of their working memory); and (2) combining physical spatial structures with sequential computational control structures to directly benefit from spatial updating capabilities. The latter is the approach most commonly found in biological problem-solving systems that immerse their perception and actuation apparatus in spatio-temporal environments to make use of spatial affordances and constraints. In this way, spatial structure needs not to be described for reasoning, it simply can be used.

My vision is that besides logic-based computational reasoning engines that are particularly useful for processing sophisticated scientific knowledge we develop spatial relation-based inference engines that structurally support perception and action processes for everyday problem solving [3]. This should become a very natural approach to autonomous mobile robotics, as robots need to interact directly with their environments. I hope we will return to addressing fundamental principles of cognitive processing and understanding their ecological roles for intelligent problem solving. At times when big data is en vogue we continue to look into the power of small data.

Acknowledgments

References

[1] CHAPPELL, J., AND KACELNIK, A. Tool selectivity in a non-primate, the New Caledonian crow *Corvus moneduloides*. *Animal Cognition 5* (2002), 71–78. doi:10.1007/s10071-002-0130-2.

[2] DYLLA, F., LEE, J. H., MOSSAKOWSKI, T., SCHNEIDER, T., VAN DELDEN, A., VAN DE VEN, J., AND WOLTER, D. A survey of qualitative spatial and temporal calculi: Algebraic and computational properties. *ACM Computing Surveys 50*, 1 (2017), 7:1–7:39. doi:10.1145/3038927.

[3] FREKSA, C. Qualitative spatial reasoning. In *Cognitive and Linguistic Aspects of Geographic Space*, D. M. Mark and A. U. Frank, Eds. Kluwer, Dordrecht, 1991, pp. 361–372. doi:10.1007/978-94-011-2606-9_20.

[4] FREKSA, C. Using orientation information for qualitative spatial reasoning. In *Theories and methods of spatio-temporal reasoning in geographic space*, A. U. Frank, I. Campari, and U. Formentini, Eds. Springer, Berlin, 1992, pp. 162–178. doi:10.1007/3-540-55966-3-10.

[5] FREKSA, C. Strong spatial cognition. In *Spatial Information Theory* (2015), S. I. Fabrikant, M. Raubal, M. Bertolotto, C. Davies, S. M. Freundschuh, and S. Bell, Eds., vol. 9368 of *Lecture Notes in Computer Science*, Springer, pp. 65–86. doi:10.1007/978-3-319-23374-1_4.

[6] FREKSA, C., AND KREUTZMANN, A. Neighborhood, conceptual. In *The International Encyclopedia of Geography*. Wiley-Blackwell, Hoboken, 2017, pp. 1–12. doi:10.1002/9781118786352.wbieg0935.

[7] MAZZOLAI, B., LASCHI, C., DARIO, P., MUGNAI, S., AND MANCUSO, S. The plant as a biomechatronic system. *Plant Signaling & Behavior 5*, 2 (2010), 90–93. doi:10.4161/psb.5.2.10457.

[8] MERMILLOD, M., GUYADER, N., AND CHAUVIN, A. The coarse-to-fine hypothesis revisited: evidence from neuro-computational modeling. *Brain and Cognition 57*, 2 (2005), 151–157. doi:10.1016/j.bandc.2004.08.035.

[9] PENFIELD, W., AND BOLDREY, E. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain 60*, 4 (1937), 389–443. doi:10.1093/brain/60.4.389.

[10] PIAGET, J., AND INHELDER, B. *The child's conception of space*. Routledge & Kegan Paul, London, 1948/1956.

[11] PIAGET, J., INHELDER, B., AND SZEMINSKA, A. *The child's conception of geometry*. International library of psychology, philosophy, and scientific method. Basic Books, New York, 1960.

[12] ROSA, M. Visual maps in the adult primate cerebral cortex: some implications for brain development and evolution. *Brazilian Journal of Medical and Biological Research 35*, 12 (2002), 1485–1498. doi:10.1590/s0100-879x2002001200008.

[13] SEARLE, J. R. Minds, brains, and programs. *Behavioral and Brain Sciences 3* (1980), 417–457.

[14] SHAFFER, D. M., KRAUCHUNAS, S. M., EDDY, M., AND MCBEATH, M. K. How dogs navigate to catch frisbees. *Psychological Science 15*, 7 (2004), 437–441. doi:10.1111/j.0956-7976.2004.00698.x.

[15] STRUBELL, E., GANESH, A., AND MCCALLUM, A. Energy and policy considerations for deep learning in NLP. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (2019), pp. 3645–3650.

[16] TVERSKY, B. *Mind in motion: how action shapes thought*. Basic Books, New York, 2019.

[17] WELLING, M. Intelligence per Kilowatt hour. Invited talk at 35th International Conference on Machine Learning (ICML), Stockholm, 2018.

[18] WILSON, M. Six views of embodied cognition. *Psychonomic Bulletin & Review 9*, 4 (2002), 625–636. doi:10.3758/BF03196322.