Talking Space: inference from spatial linguistic meanings (abstract)

Vincent Wang-Maścianica and Bob Coecke

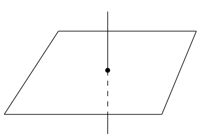
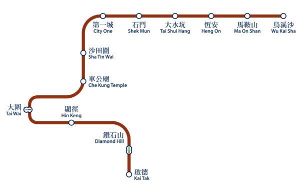
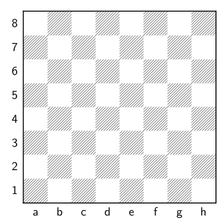
Abstract

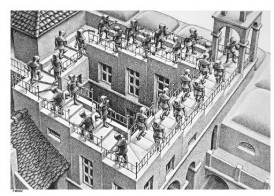
This is an **extended abstract** associated with the published paper with the same title that appeared in *Journal of Cognitive Science*, **22**(3), 421–463, 2021. arXiv:2109.06554. It concerns the intersection of natural language and the physical space around us in which we live, that we observe and/or imagine things within. Many important features of language have spatial connotations, e.g. prepositions (in, next to, after, on, etc.) are fundamentally spatial. Space is also a key factor of the meanings of many words/phrases/sentences/text (e.g. being above, the next station, chasing), and space is the main context for referencing (e.g. pointing) and embodiment.

We propose a mechanism for how space and linguistic structure can be made to interact in a matching compositional fashion. Examples include Cartesian space, subway stations, chesspieces on a chess-board, and Penrose's staircase. The starting point for our construction is the quantum-inspired DisCoCat model of compositional natural language meaning, which we relax to accommodate physical space. We address the issue of having multiple agents/objects in a space, including the case that each agent has different capabilities with respect to that space, e.g., the specific moves each chesspiece can make, or the different velocities one may be able to reach.

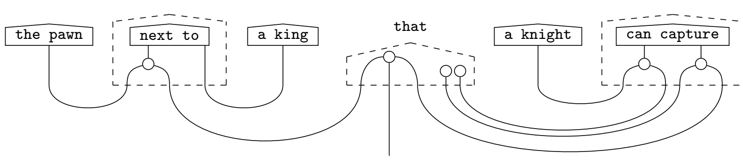
Once our model is in place, we show how inferences drawing from the structure of physical space can be made. We also how how linguistic model of space can interact with other such models related to our senses and/or embodiment, such as the conceptual spaces of colour, taste and smell, resulting in a rich compositional model of meaning that is close to human experience and embodiment in the world. The correspondence of this 'relational' space-time model with categorical quantum mechanics is also of particular interest for foundational physics.

Space plays a central role in the use and development of language [6], and as Gärdenfors [4, 5] points out, physical space is a natural mediator between sensory data and linguistic content: for many of us, interpreting linguistic content is spatial visualisation. Let's consider 3D Cartesian space, a subway line, a chess board, and something weird:

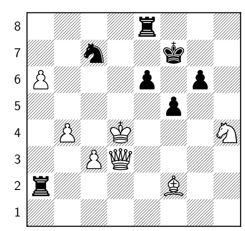






One kind of thing that we would like to to, is from `The painting is above the chest` & `The light is above the painting` infer that `The light is above the chest`, and from `Alice chases Bob` & `Alice is in Paris` infer that `Bob is in Paris`. We also want a language diagram like:

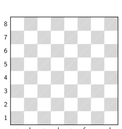
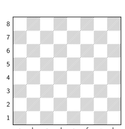


w.r.t.

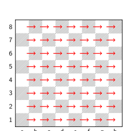
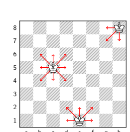


to specify a specific piece on the chessboard, hence exposing explicit 'spatial reasoning capability'.

In order to combine physical space and linguistic structure we will take the DisCoCat [3] and DisCoCirc [1] models of compositional natural language meaning as our starting point. DisCoCirc is needed here as in the above examples we deal with multiple sentences. In its most general interpretation, it assumes a compositional model of linguistic structure, and meaning spaces organised in a matching structure. We are going for this model in the case of the chess-board, with the obvious analogue for other spaces: Monoidal subcategory of **Rel**:

-  $\times \ldots \times$ 
- "spatial relations" $R \subseteq X \times \ldots \times X$ including cups, caps, spiders and standard composition [2]

Example relations for chess are:

`move right` :=    `king's moves` := 

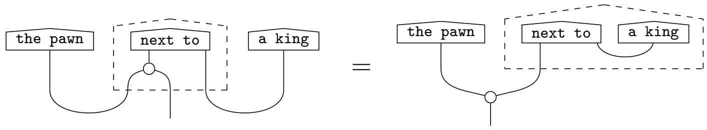
for subway:

$$\texttt{next stop} := \{(\texttt{Kai Tak}, \texttt{D. H.}), (\texttt{D. H.}, \texttt{Hin Keng}), \ldots\} \qquad \texttt{in-between} := \{(\texttt{Kai Tak}, \texttt{D. H.}, \texttt{Hin Keng}), \ldots\}$$

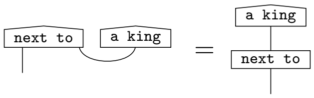
and for Cartesian space:

$$\texttt{above} := \{((x,y,z),(x',y',z')) | x = x' \ , \ y = y' \ , \ z > z'\} \qquad \texttt{chases}_{\delta t > 0} := \{((x,y,z,t),(x,y,z,t')) | t = t' + \delta t\}$$

Starting with the chess example, the noun-phrase `pawn next to a king` has the following diagram:



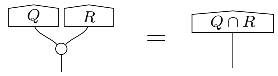
where `next to a king` is :



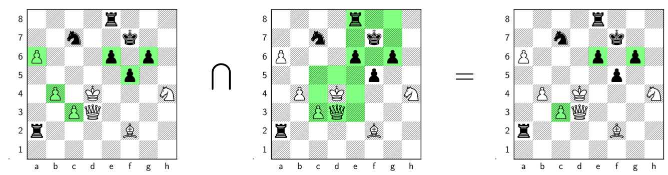
$$= \{(\alpha x, \beta y) | |\alpha - \beta| = 1 \text{ or } |x - y| = 1\} \circ \{\texttt{d4, f7}\}$$

$$= \{\texttt{c3,c4,c5,d3,d5,e3,e4,e5,e6,e7,e8,f6,f8,g6,g7,g8}\}$$

Using:



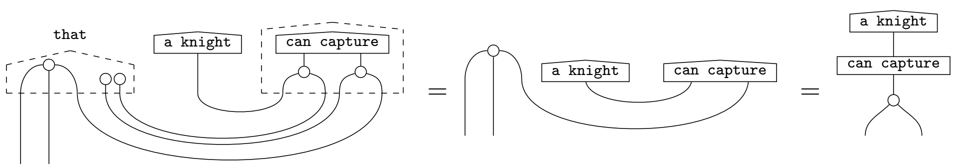
we get the intersection of `the pawn` and `next to a king`:



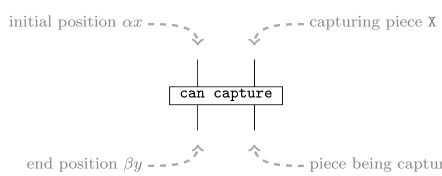
Now, in an example like chess, when we say `can capture`, this will mean something very different depending on the chesspiece doing the capturing. Generally, different inhabitants have different capabilities and properties in the space they inhabit. In order to make this a part of our relations model, we will first augment the representation of the space, besides the locations of the chess board, also accounting for the spatially-related features an inhabitant may have:

$$\texttt{chessboard} := (\texttt{a-h} \times \texttt{1-8}) \times \{♙, ♖, ♗, ♘, ♕, ♔\} \qquad (1)$$

and e.g. $\texttt{pawn} = \{\texttt{a6,b4,c3,e6,f5,g6}\} \times \{♙\}$ and $\texttt{king} := \{\texttt{d4,f7}\} \times \{♔\}$. In order to deal with:



we fine-grain each wire of the `can capture` box as two wires: left wire for spatial locations, and right wire for chesspiece labels. The result is a box with four wires as follows:

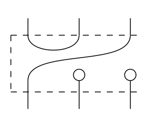


Here the data of different piece movement rules is encoded within `can capture`. One may argue that it is more natural that, rather than the meaning of `can capture` having to carry the data concerning which pieces can do what, that the chesspieces themselves should carry that data. This is also possible. In that case, rather than the space (1) we have as the space:

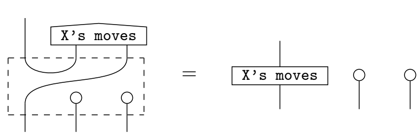
$$\texttt{chessboard} := (\texttt{a-h} \times \texttt{1-8}) \times \{♙'\texttt{s moves}, ♖'\texttt{s moves}, \ldots, ♔'\texttt{s moves}\}$$

and:

`can capture` :=



since then we have:



As our long paper shows, this takes us to the right result, and we also provide an example in it of a chase in the animal kingdom, where speed and endurance matter. For inferences we also refer to our long paper, as well as examples combining space with features such as colour and smell.

References

[1] B. Coecke. *The Mathematics of Text Structure*, pages 181–217. Springer International Publishing, 2021. arXiv:1904.03478.

[2] B. Coecke and A. Kissinger. *Picturing Quantum Processes. A First Course in Quantum Theory and Diagrammatic Reasoning*. Cambridge University Press, 2017.

[3] B. Coecke, M. Sadrzadeh, and S. Clark. Mathematical foundations for a compositional distributional model of meaning. In J. van Benthem, M. Moortgat, and W. Buszkowski, editors, *A Festschrift for Jim Lambek*, volume 36 of *Linguistic Analysis*, pages 345–384. 2010. arxiv:1003.4394.

[4] P. Gärdenfors. *Conceptual spaces: the geometry of thought*. MIT Press, 2000.

[5] P. Gärdenfors. *The Geometry of Meaning: Semantics Based on Conceptual Spaces*. MIT Press, 2014.

[6] G. Lakoff and M. Johnson. *Metaphors We Live By*. University of Chicago Press, Chicago, new edition edition, 2003.