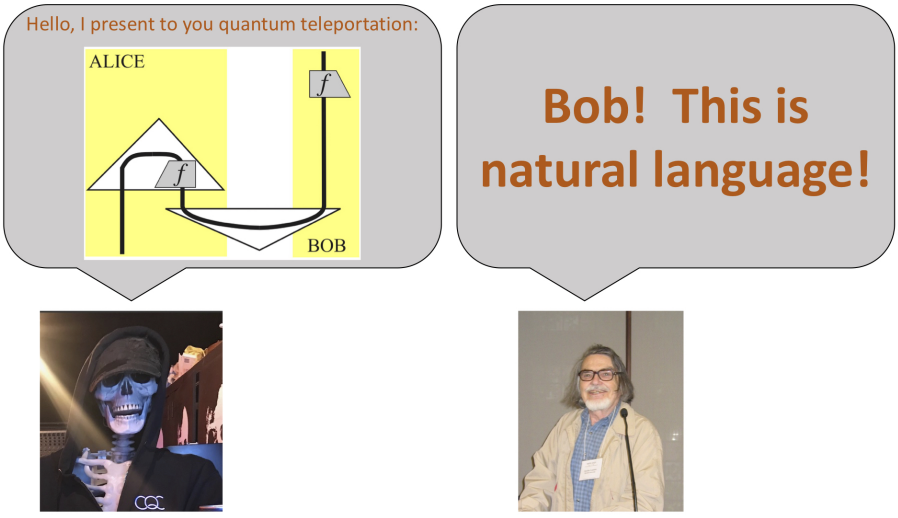
A Tale of Four Disciplines for All Ages and All Languages

Bob Coecke
Quantinuum, 17 Beaumont Street, OX1 2NA Oxford
bob.coecke@cambridgequantum.com

Our tale starts with John von Neumann's rejection of his own quantum mechanical formalism [26], goes via Erwin Schrödinger's conviction that composition is at the heart of quantum theory [29] and Roger Penrose's diagrammatic substitute for tensor notation [25], and a first pit-stop at categorical quantum mechanics (CQM) [5, 6, 11]. First presented as a high-level language for quantum computing [1], it has evolved in an educational tool [12], first university level, and now also high-school level [9], or maybe even earlier. Experiments are underway to verify this.

Going back in time to 2005, when Jim Lambek attended our presentation of quantum teleportation in diagrammatic terms, he yelled: "Bob, this is grammar!"



Hence, when Steve Clark and Steve Pulman put the challenge forward about how to combine meaning and grammar, in a manner that sentence meanings all live in the same space [4], we knew that CQM provided the answer, and did so in a canonical manner [3, 15, 2]. The resulting framework has been referred to as DisCoCat. While in most presentations pregroup grammar [20, 21] has been used, the framework works for any kind of type grammar [10, 34].

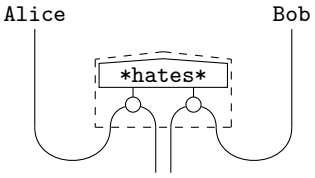
Moreover, DisCoCat does not only consider linguistic structure, but also uses the spiders of CQM to provide 'logical' accounts of words like relative pronouns [27, 28, 18, 13, 7].

But of course, a quantum model's natural habitat is a quantum computer [17], and after a series of theoretical explorations [35, 23, 8], we performed quantum natural language processing on quantum hardware [24, 22], and now you can do so too, making use of our open-source Python library lambeq [19]. What lambeq does, is turning any sentence in a circuit that can then be compiled on any kind of quantum hardware, e.g. via Quantinuum's tket [30].

One limitation of DisCoCat is that it does not allow for sentence meanings to be composed into text, sentence meanings being vectors. On the other hand, circuits are things that do easily compose. Prior

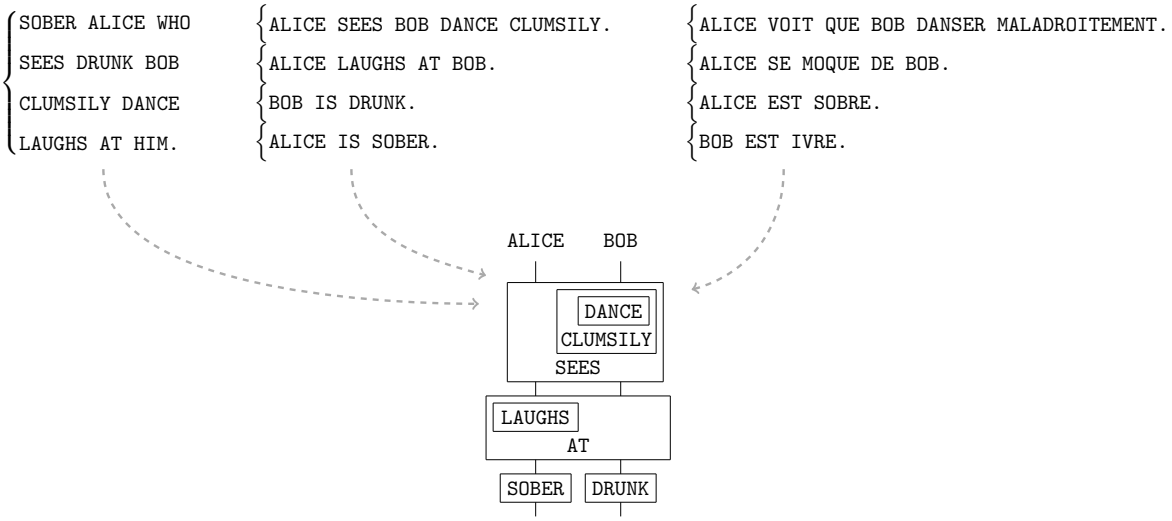
M. Moortgat and G. Wijnholds: End-to-End Compositional Models of Vector-Based Semantics, 2022 (E2ECOMPVEC)
EPTCS ??, 2022, pp. 1–4, doi:10.4204/EPTCS.??.1

to our quantum implementations, we had already proposed DisCoCirc as an improvement on DisCoCat [7]. By taking nouns as inputs rather than as vectors, and substituting sentence types by noun-types, we obtain composable circuits:



Now, noun-meanings aren't fixed, but evolve as text progresses [14, 16], and they will also entangle as the corresponding agents are interacting, like in the above picture.

When we attempted to realise DisCoCirc for all of English [32], we were up for a surprise: we discovered that in the passage to circuits differences between languages in terms of word-ordering vanished [33], and so did stylistic differences:



These circuits are moreover generative: any circuit that one can put together always corresponds to some text, although of course that text won't be unique.

One may hope that therefore DisCoCirc brings us closer to what meaning actually is, independent from the bureaucracy that comes into play when forcing language to live on a line, as we humans do. For example, it is much closer to spatial reasoning [31].

References

[1] S. Abramsky and B. Coecke. A categorical semantics of quantum protocols. In *Proceedings of the 19th Annual IEEE Symposium on Logic in. Computer Science (LICS)*, pages 415–425, 2004. arXiv:quant-ph/0402130. doi:10.48550/arXiv.quant-ph/0402130

[2] S. Clark, B. Coecke, E. Grefenstette, S. Pulman, and M. Sadrzadeh. A quantum teleportation inspired algorithm produces sentence meaning from word meaning and grammatical structure. *Malaysian Journal of Mathematical Sciences*, 8:15–25, 2014. arXiv:1305.0556. doi:10.48550/arXiv.1305.0556

[3] S. Clark, B. Coecke, and M. Sadrzadeh. A compositional distributional model of meaning. In *Proceedings of the Second Quantum Interaction Symposium (QI-2008)*, pages 133–140, 2008. url:http://www.cs.ox.ac.uk/people/stephen.clark/papers/qai08.pdf

[4] S. Clark and S. Pulman. Combining symbolic and distributional models of meaning. In *Proceedings of AAAI Spring Symposium on Quantum Interaction*. AAAI Press, 2007. url:https://www.aaai.org/Papers/Symposia/Spring/2007/SS-07-08/SS07-08-008.pdf

[5] B. Coecke. Kindergarten quantum mechanics. In A. Khrennikov, editor, *Quantum Theory: Reconsiderations of the Foundations III*, pages 81–98. AIP Press, 2005. arXiv:quant-ph/0510032. doi:10.48550/arXiv.quant-ph/0510032

[6] B. Coecke. Quantum picturalism. *Contemporary Physics*, 51:59–83, 2009. arXiv:0908.1787. doi:10.48550/arXiv.0908.1787

[7] B. Coecke. *The Mathematics of Text Structure*, pages 181–217. Springer International Publishing, 2021. arXiv:1904.03478. doi:10.48550/arXiv.1904.03478

[8] B. Coecke, G. de Felice, K. Meichanetzidis, and A. Toumi. Foundations for near-term quantum natural language processing, 2020. arXiv:2012.03755. doi:10.48550/arXiv.2012.03755

[9] B. Coecke and S. Gogioso. *Quantum in Pictures*. Quantinuum, September 2022.

[10] B. Coecke, E. Grefenstette, and M. Sadrzadeh. Lambek vs. Lambek: Functorial vector space semantics and string diagrams for Lambek calculus. *Annals of Pure and Applied Logic*, 164:1079–1100, 2013. arXiv:1302.0393. doi:10.1016/j.apal.2013.05.009

[11] B. Coecke, D. Horsman, A. Kissinger, and Q. Wang. Kindergarden quantum mechanics graduates... or how i learned to stop gluing LEGO together and love the ZX-calculus. *Theoretical Computer Science*, 897:1–22, 2022. arXiv:2102.10984. doi:10.48550/arXiv.2102.10984

[12] B. Coecke and A. Kissinger. *Picturing Quantum Processes. A First Course in Quantum Theory and Diagrammatic Reasoning*. Cambridge University Press, 2017. doi:10.1017/9781316219317

[13] B. Coecke, M. Lewis, and D. Marsden. Internal wiring of cartesian verbs and prepositions. In M. Lewis, B. Coecke, J. Hedges, D. Kartsaklis, and D. Marsden, editors, Procs. of the 2018 Workshop on *Compositional Approaches in Physics, NLP, and Social Sciences*, volume 283 of *Electronic Proceedings in Theoretical Computer Science*, pages 75–88, 2018. doi:10.4204/EPTCS.283.6

[14] B. Coecke and K. Meichanetzidis. Meaning updating of density matrices. *Journal of Applied Logics*, 2631(5):745, 2020. doi:10.48550/arXiv.2001.00862

[15] 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. doi:10.48550/arXiv.1003.4394

[16] G. De las Cuevas, A. Klingler, M. Lewis, and T. Netzer. Cats climb entails mammals move: preserving hyponymy in compositional distributional semantics. arXiv:2005.14134, 2020. doi:10.48550/arXiv.2005.14134

[17] R. P. Feynman. Simulating physics with computers. *International journal of theoretical physics*, 21:467–488, 1982.

[18] D. Kartsaklis. Coordination in categorical compositional distributional semantics. In *Semantic Spaces at the Intersection of NLP, Physics and Cognitive Science*, 2016. arXiv:1606.01515. doi:10.4204/EPTCS.221.4

[19] D. Kartsaklis, I. Fan, R. Yeung, A. Pearson, R. Lorenz, A. Toumi, G. de Felice, K. Meichanetzidis, S. Clark, and B. Coecke. lambeq: An efficient high-level Python library for quantum NLP. arXiv:2110.04236, 2021. doi:10.48550/arXiv.2110.04236

[20] J. Lambek. Type grammar revisited. *Logical Aspects of Computational Linguistics*, 1582, 1999. doi:10.1007/3-540-48975-4_1

[21] J. Lambek. From word to sentence. *Polimetrica, Milan*, 2008. url:https://www.math.mcgill.ca/barr/lambek/pdffiles/2008lambek.pdf

[22] R. Lorenz, A. Pearson, K. Meichanetzidis, D. Kartsalkis, and B. Coecke. Qnlp in practice: Running compositional models of meaning on a quantum computer. arXiv:2102.12846, 2021. doi:10.48550/arXiv.2102.12846

[23] K. Meichanetzidis, S. Gogioso, G. de Felice, A. Toumi, N. Chiappori, and B. Coecke. Quantum natural language processing on near-term quantum computers. *Accepted for QPL 2020*, 2020. arXiv:2005.04147. doi:10.4204/EPTCS.340.11

[24] K. Meichanetzidis, A. Toumi, G. de Felice, and B. Coecke. Grammar-aware question-answering on quantum computers. arXiv:2012.03756, 2020. doi:10.48550/arXiv.2012.03756

[25] R. Penrose. Applications of negative dimensional tensors. In *Combinatorial Mathematics and its Applications*, pages 221–244. Academic Press, 1971. url:http://homepages.math.uic.edu/ kauffman/Penrose.pdf

[26] M. Redei. Why John von Neumann did not like the Hilbert space formalism of quantum mechanics (and what he liked instead). *Studies in History and Philosophy of Modern Physics*, 27(4):493–510, 1996. url:citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.35.763&rep=rep1&type=pdf

[27] M. Sadrzadeh, S. Clark, and B. Coecke. The Frobenius anatomy of word meanings I: subject and object relative pronouns. *Journal of Logic and Computation*, 23:1293–1317, 2013. arXiv:1404.5278. doi:10.1093/logcom/ext044

[28] M. Sadrzadeh, S. Clark, and B. Coecke. The Frobenius anatomy of word meanings II: possessive relative pronouns. *Journal of Logic and Computation*, 26:785–815, 2016. arXiv:1406.4690. doi:10.1093/logcom/exu027

[29] E. Schrödinger. Discussion of probability relations between separated systems. *Cambridge Philosophical Society*, 31:555–563, 1935. doi:10.1017/S0305004100013554

[30] S. Sivarajah, S. Dilkes, A. Cowtan, W. Simmons, A. Edgington, and R. Duncan. t|ket>: A retargetable compiler for NISQ devices. arXiv:2003.10611, 2020. doi:10.48550/arXiv.2003.10611

[31] V. Wang-mańcianica and B. Coecke. Talking space: Inference from spatial linguistic meanings. *Journal of Cognitive Science*, 22(3):421–463, 2021. doi:10.48550/arXiv.2109.06554

[32] V. Wang-Mascianica, J. Liu, and B. Coecke. Distilling text into circuits. 2022. arXiv:soon.

[33] J. Waseem, M. H. Liu, V. Wang-Mascianica, and B. Coecke. Language-independence of DisCoCirc's text circuits: English and Urdu. 2022. ESSLLI-EPTCS.

[34] R. Yeung and D. Kartsaklis. A CCG-based version of the DisCoCat framework. arXiv:2105.07720, 2021. doi:10.48550/arXiv.2105.07720

[35] W. Zeng and B. Coecke. Quantum algorithms for compositional natural language processing. *Electronic Proceedings in Theoretical Computer Science*, 221, 2016. arXiv:1608.01406. doi:10.4204/EPTCS.221.8