Causal Spaces and String Diagrams

Siddharth Setlur

November 2, 2025

1 Introduction

References

[BPS24] Simon Buchholz, Junhyung Park, and Bernhard Schölkopf. *Products, Abstractions and Inclusions of Causal Spaces*. 2024. DOI: 10.48550/ARXIV.2406.00388. URL: https://arxiv.org/abs/2406.00388 (visited on 05/10/2025).

[FK23] Tobias Fritz and Andreas Klingler. "The d-Separation Criterion in Categorical Probability". In: *Journal of Machine Learning Research* 24.46 (2023), pp. 1–49. ISSN: 1533-7928. URL: http://jmlr.org/papers/v24/22-0916.html (visited on 08/01/2025).

[Gei+21] Atticus Geiger et al. *Causal Abstractions of Neural Networks*. 2021. DOI: 10.48550/ARXIV.2106.02997. URL: https://arxiv.org/abs/2106.02997.

[Gei+25] Atticus Geiger et al. "Causal Abstraction: A Theoretical Foundation for Mechanistic Interpretability". In: *Journal of Machine Learning Research* 26.83 (2025), pp. 1–64. ISSN: 1533-7928. URL: http://jmlr.org/papers/v26/23-0058.html.

[Hub24] Martin Huber. "An introduction to causal discovery". In: *Swiss Journal of Economics and Statistics* 160.1 (Oct. 29, 2024), p. 14. ISSN: 2235-6282. DOI: 10.1186/s41937-024-00131-4. URL: https://sjes.springeropen.com/articles/10.1186/s41937-024-00131-4.

[JKZ19] Bart Jacobs, Aleks Kissinger, and Fabio Zanasi. "Causal Inference by String Diagram Surgery". In: *Foundations of Software Science and Computation Structures*. Ed. by Mikołaj Bojańczyk and Alex Simpson. Vol. 11425. Cham: Springer International Publishing, 2019, pp. 313–329. ISBN: 9783030171261 9783030171278. DOI: 10.1007/978-3-030-17127-8_18. URL: https://link.springer.com/10.1007/978-3-030-17127-8_18 (visited on 08/01/2025).

[Law62] William Lawvere. "The Category of Probabilistic Maps". In: (1962). URL: https://ncatlab.org/nlab/files/lawvereprobability1962.pdf.

[LT23] Robin Lorenz and Sean Tull. *Causal models in string diagrams*. arXiv:2304.07638. Apr. 2023. DOI: 10.48550/arXiv.2304.07638. URL: http://arxiv.org/abs/2304.07638.

[ML78] Saunders Mac Lane. *Categories for the Working Mathematician*. Vol. 5. Graduate Texts in Mathematics. New York, NY: Springer New York, 1978. ISBN: 9781441931238 9781475747218. DOI: 10.1007/978-1-4757-4721-8. URL: http://link.springer.com/10.1007/978-1-4757-4721-8.

[Par+24] Junhyung Park et al. *A Measure-Theoretic Axiomatisation of Causality*. arXiv:2305.17139. June 2024. DOI: 10.48550/arXiv.2305.17139. URL: http://arxiv.org/abs/2305.17139.

[Pet17] Jonas Peters. *Lectures on Causality: Jonas Peters, Part 1*. Mar. 2017. URL: https://www.youtube.com/watch?v=zvrcyqcN9Wo&list=PLW01hpWnEtbTcuY0a0jhZyanHX3GPImAy&index=1.

[PJS17] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference: foundations and learning algorithms*. eng. Adaptive computation and machine learning. Cambridge, Mass: The MIT press, 2017. ISBN: 9780262037310.

[Tem+24] Adly Templeton et al. "Scaling Monosemanticity: Extracting Interpretable Features from Claude 3 Sonnet". In: *Transformer Circuits Thread* (2024). URL: https://transformer-circuits.pub/2024/scaling-monosemanticity/index.html.

[Tri+17] Sofia Triantafillou et al. "Predicting Causal Relationships from Biological Data: Applying Automated Causal Discovery on Mass Cytometry Data of Human Immune Cells". In: *Scientific Reports* 7.1 (Oct. 5, 2017), p. 12724. ISSN: 2045-2322. DOI: 10.1038/s41598-017-08582-x. URL: https://www.nature.com/articles/s41598-017-08582-x (visited on 08/01/2025).

[Tul+24] Sean Tull et al. *Towards Compositional Interpretability for XAI*. arXiv:2406.17583. June 2024. DOI: 10.48550/arXiv.2406.17583. URL: http://arxiv.org/abs/2406.17583.

[Çi11] Erhan Çinlar. *Probability and Stochastics*. en. Vol. 261. Graduate Texts in Mathematics. New York, NY: Springer New York, 2011. ISBN: 9780387878584 9780387878591. DOI: 10.1007/978-0-387-87859-1. URL: https://link.springer.com/10.1007/978-0-387-87859-1.