# Notes on Causality, Prediction and Search by Peter Spirtes

D. Gueorguiev, 9/13/23

## Notation and Basic Definitions:

### Graphs:

Notation for Edges:

**undirected edge** :

**directed edge** :

**non-directed edge** :

**partially directed edge** :

Note on Inducing Path Graph: contains both directed edges (e.g., ) , bi-directed edges (e.g., ), non-directed edges (e.g., ), and partially directed edges (e.g., ).

**Graph** (*traditional definition*): ordered pair where is a set of vertices and is a set of edges. The members of are pairs of vertices (an ordered pair in a directed graph and an unordered pair in an undirected graph). For example, the edge is represented by the ordered pair . We need to specify variables and **marks** at each end. In general, we will allow that the end of an edge can be unmarked, can be marked with an arrowhead , or can be marked with an . For example, the left end of can be represented as ordered pair , while the right end can be represented as the ordered pair . The entire edge is a set of ordered pairs representing the endpoints . The edge is the same as .

Note that a directed edge such as has no mark at the endpoint; we consider the mark at the A endpoint to be empty, but when we write out the ordered pair, we will use the notation to stand for the empty mark e.g.,

**Graph** (our definition): an ordered triple where is a non-empty set of vertices, is a non-empty set of marks, and is a set of sets of ordered pairs of the form , where and are in , , and and are in . Except in our discussion of systems with feedback we will always assume that in any graph, any pair of vertices and occur in at most one set in , or, in other words, that there is at most one edge between any two vertices. If we say that is over .

Figure 1: Example of directed graph

For example, the directed graph on Figure 1 can be represented as:

**edge**: any member of .

**edge-end**: each ordered pair

**endpoint**: each vertex in an edge

**adjacent endpoints:** vertices are adjacent iff there is an edge with endpoints

**undirected graph**: a graph in which the set of marks .

**directed graph**: a graph in which the set of marks .

**directed edge** from to : an edge

**edge into** : any edge

**edge out** of **:** any edge

**parent/child**: is parent of which is child of if there is a directed edge from to

**indegree** of vertex : the number of the parents of

**outdegree** of vertex : the number of children of

**undirected path** between and in graph : a sequence of vertices beginning with and ending with such that for every pair of vertices and that are adjacent in the sequence there is an edge in .

**edge is in path**: is in the path iff and are adjacent to each other (in either order) in .

vertices **adjacent on path**: if an edge between and is in the path we say that and are adjacent on .

**path is out of vertex**: if the edge containing in **an undirected** **path** between and **is** **out of** then we say that **path is out** of .

**path is into vertex**: if the edge containing in a path between and is into we say that **the path is into** .

**empty path**: sequence which consists of a single vertex.

**acyclic path**: a path which contains no vertex more than once; otherwise, it is **cyclic path**.

## Introductory Notes

*Fundamental questions in this thesis*: understanding the systematic connections between causal dependency and stochastic dependency. What are the limits to reliable causal inference given the presence of such connections? Under what conditions the causal inference will produce reliable results? What are the limitations for causal inference given our theoretical understanding between causality and probability? Within these limits we can investigate rigorously the shortcomings of the algorithms that search for causal structures from statistical properties of the datasets.

The asymptotic reliabilities of well-defined procedures can be determined mathematically, while short run behavior can be estimated/bounded through well-chosen simulation experiments. This thesis investigates into the theory of model specification search adopting results from the theory of estimation.

*Additional investigations*:

Understanding latent variables and their impact on causality.

# Bibliography

Eells, E. (1991). *Probabilistic Causality.* Cambridge UK: Cambridge University Press.

G.E. Hughes, M. C. (1996). *A New Introduction To Modal Logic.* London: Routledge.

Good, I. J. (1961). A Causal Calculus (I and II). *The British Journal for the Philosophy of Science*, 305-318, 43-51.

Hume, D. (1748). *Philosophical Essays Concerning Human Understanding.* London: Printed for A. Millar, opposite Katharine-Street in the Strand. MDCCXLVII.

Lewis, D. (1973). *Counterfactuals.* Malden, Massachusetts: Blackwell Publishers.

Lewis, D. (1974). Causation. *Journal of Philosophy*, 556-567.

Mackie, J. (1980). *The Cement of The Universe: A Study of Causation.* New York: Oxford University Press, New York, United States.

Otte, R. E. (1982). *Probability and Causality, PhD Thesis.* Ann Arbor , MI, 48106: University of Arizona Graduate College, University Microfilm International.

Reichenbach, H. (1956). *The Direction of Time.* Berkeley and Los Angeles, California: University of California Press.

Salmon, W. C. (1980). Probabilistic Causality. *Pacific Philosophical Quarterly*, pp. 137-153.

Salmon, W. C. (1998). *Causality and Explanation.* Pittsburgh, Pennsylvania: Oxford University Press.

Spirtes, P., Glymour, C., & Sheines, R. (1993). *Causation, Prediction and Search.* New York: Springer Verlag.

# Downloadable Links for the Bibliography

(Eells, 1991): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/books/eells_probabilistic_causality_1991.pdf)

(Reichenbach, 1956): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/books/the-direction-of-time-hans-reichenbach-ucal-press-1971.pdf)

(Spirtes, Glymour, & Sheines, 1993): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/books/CausationPredictionandSearch_Spirtes_CMU_2000.pdf)

(Otte, 1982): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/Probability_and_Causality_PhD_Thesis_Otte_1982.pdf)

(G.E. Hughes, 1996): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/ModelChecking/huges_cresswell_modal_logic.pdf)

(Salmon, Causality and Explanation, 1998): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/books/Causality_and_Explanation_Wesley_Salmon_1997.pdf)

(Salmon, Probabilistic Causality, 1980): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/Probabilistic_Causality_Salmon_1980.pdf)

(Good, 1961): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/CausalCalculus_part_I_and_II_Good_1960.pdf)

(Mackie, 1980): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/books/cement-of-the-universe-a-study-of-causation-JL-Mackie-1980.pdf)

(Lewis, Counterfactuals, 1973): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/books/Counterfactuals-lewis-1973.pdf)

(Lewis, Causation, 1974): [here](https://github.com/dimitarpg13/root_cause_analysis_and_model_checking/blob/main/literature/Lewis-Causation_1974.pdf)