

CAUSALITY

Models, Reasoning, and Inference

Judea Pearl

Development of Western science is based on two great achievements: the invention of the formal logical system (in Euclidean geometry) by the Greek philosophers, and the discovery of the possibility to find out causal relationships by systematic experiment (during the Renaissance).

— A. Einstein, April 23, 1953

Forthcoming, Cambridge University Press

Contents

1	INTRODUCTION TO PROBABILITIES, GRAPHS, AND CAUSAL MODELS	23
1.1	Introduction to Probability Theory	24
1.1.1	Why probabilities	24
1.1.2	Basic concepts in probability theory	26
1.1.3	Combining predictive and diagnostic supports . .	35
1.1.4	Random Variables and Expectations	38
1.1.5	Conditional independence and graphoids	42
1.2	Graphs and Probabilities	44
1.2.1	Graphical Notation and Terminology	44
1.2.2	Bayesian Networks	47
1.2.3	The d -separation criterion	53
1.2.4	Inference with Bayesian networks	60
1.3	Causal Bayesian Networks	63
1.3.1	Causal networks as oracles for interventions . . .	66
1.3.2	Causal relationships and their stability	70
1.4	Functional Causal Models	73

1.4.1	Structural Equations	75
1.4.2	Probabilistic predictions in causal models	80
1.4.3	Interventions and causal effects in functional models	85
1.4.4	Counterfactuals in functional models	88
1.5	Causal vs. Statistical Terminology	99
2	A THEORY OF INFERRED CAUSATION	103
2.1	Introduction	106
2.2	The Causal Modeling Framework	109
2.3	Model Preference (Occam's razor)	112
2.4	Stable Distributions	118
2.5	Recovering DAG Structures	121
2.6	Recovering Latent Structures	125
2.7	Local Criteria for Causal Relations	131
2.8	Non-Temporal Causation and Statistical Time	137
2.9	Conclusions	141
2.9.1	On minimality, Markov, and stability	145
3	CAUSAL DIAGRAMS AND THE IDENTIFICATION OF CAUSAL EFFECTS	153
3.1	Introduction	156
3.2	Intervention in Markovian Models	160
3.2.1	Graphs as models of interventions	160
3.2.2	Interventions as variables	165

3.2.3	Computing the effect of interventions	168
3.2.4	Identification in causal models	177
3.3	Controlling Confounding Bias	180
3.3.1	The back-door criterion	182
3.3.2	The front-door criterion	185
3.3.3	Example: Smoking and the genotype theory . . .	189
3.4	A Calculus of Intervention	193
3.4.1	Preliminary notation	193
3.4.2	Inference rules	194
3.4.3	Symbolic derivation of causal effects: An example	196
3.4.4	Causal inference by surrogate experiments	200
3.5	Graphical Tests of Identifiability	202
3.5.1	Identifying models	205
3.5.2	Nonidentifying models	209
3.6	Discussion	211
3.6.1	Qualifications and extensions	211
3.6.2	Diagrams as a mathematical language	215
3.6.3	Translation from Graphs to Potential Outcomes .	218
3.6.4	Relations to Robin's G -estimation	228
4	ACTIONS, PLANS, AND DIRECT EFFECTS	233
4.1	Introduction	235
4.1.1	Actions, acts and probabilities	235
4.1.2	Actions in decision analysis	240
4.1.3	Actions and counterfactuals	243

4.2	Conditional Actions and Stochastic Policies	245
4.3	When is the Effect of an Action Identifiable?	248
4.3.1	Graphical conditions for identification	248
4.3.2	Remarks on efficiency	252
4.3.3	Deriving a closed-form expression for control queries	254
4.3.4	Summary	255
4.4	The Identification of Plans	256
4.4.1	Motivation	257
4.4.2	Plan identification: Notation and assumptions . .	261
4.4.3	Plan identification: A general criterion	263
4.4.4	Plan identification: A procedure	267
4.5	Direct Effects and their Identification	272
4.5.1	Direct vs. Total Effects:	272
4.5.2	Direct effects, definition and identification	274
4.5.3	Example: Sex Discrimination in College Admis- sion	276
4.5.4	Average direct effects	280
5	CAUSALITY AND STRUCTURAL MODELS IN THE SOCIAL SCIENCES	283
5.1	Introduction	286
5.1.1	Causality in search of a language	286
5.1.2	SEM: How its meaning got obscured	288
5.1.3	Graphs as a mathematical language: An example	295

5.2	Graphs and Model Testing	298
5.2.1	The testable implications of structural models . .	299
5.2.2	Testing the testable	306
5.2.3	Model equivalence	309
5.3	Graphs and Identifiability	318
5.3.1	Parameter identification in linear models	318
5.3.2	Comparison to nonparametric identification . . .	328
5.3.3	Causal effects: The interventional interpretation of structural equation models	332
5.4	Some Conceptual Underpinnings	337
5.4.1	What do structural parameters really mean? . . .	337
5.4.2	Interpretation of effect decomposition	347
5.4.3	Exogeneity, superexogeneity and other frills . . .	351
5.5	Conclusion	361
6	SIMPSON'S PARADOX, CONFOUNDING, AND COLLAPSIBILITY	365
6.1	Simpson's Paradox: An Anatomy	367
6.1.1	A tale of a non-paradox	367
6.1.2	A tale of statistical agony	371
6.1.3	Causality vs. exchangeability	374
6.1.4	A paradox resolved (or what kind of machine is man)	381
6.2	Why there is no statistical test for confounding, why many think there is, and why they are almost right . . .	385

6.2.1	Introduction	385
6.2.2	Causal and associational definitions	388
6.3	How the Associational Criterion Fails	390
6.3.1	Failing sufficiency due to marginality	390
6.3.2	Failing sufficiency due to closed-world assumptions	391
6.3.3	Failing necessity due to barren proxies	392
6.3.4	Failing necessity due to incidental cancelations	396
6.4	Stable vs. Incidental Unbiasedness	397
6.4.1	Motivation	397
6.4.2	Formal definitions	401
6.4.3	Operational test for stable no-confounding	404
6.5	Confounding, Collapsibility, and Exchangeability	406
6.5.1	Confounding and collapsibility	406
6.5.2	Confounding vs. confounders	408
6.5.3	Exchangeability vs. structural analysis of confounding	412
6.6	Conclusions	419
6.7	Acknowledgment	421
7	STRUCTURAL AND COUNTERFACTUAL MODELS	423
7.1	Structural Model Semantics	425
7.1.1	Definitions: Causal models, actions and counterfactuals	427

7.1.2	Evaluating counterfactuals: Deterministic analysis	435
7.1.3	Evaluating counterfactuals: Probabilistic analysis	445
7.1.4	The twin-network method	447
7.2	Applications and Interpretation of Structural Models . .	451
7.2.1	Policy analysis in linear econometric models: An example	451
7.2.2	The empirical content of counterfactuals	456
7.2.3	Causal explanations, utterances, and their inter- pretation	463
7.2.4	From mechanisms to actions to causation	468
7.2.5	Simon's causal ordering	474
7.3	Axiomatic Characterization	479
7.3.1	The axioms of structural counterfactuals	479
7.3.2	Causal effects from counterfactual logic: An ex- ample	485
7.3.3	Axioms of causal relevance	489
7.4	Structural and Similarity-based Counterfactuals	496
7.4.1	Relations to Lewis' counterfactuals	496
7.4.2	Axiomatic comparison	501
7.4.3	Imaging vs. conditioning	505
7.4.4	Relations to Neyman-Rubin framework	507
7.4.5	Exogeneity revisited: counterfactual and graphi- cal definitions	513
7.5	Structural vs. Probabilistic Causality	518
7.5.1	The reliance on temporal ordering	520

7.5.2	The perils of circularity	521
7.5.3	The closed-world assumption	526
7.5.4	Singular vs. general causes	529
7.5.5	Summary	535
8	IMPERFECT EXPERIMENTS: BOUNDS AND COUNTERFACTUALS	539
8.1	Introduction	540
8.1.1	Imperfect and indirect experiments	540
8.1.2	Noncompliance and intent-to-treat	544
8.2	Bounding Causal Effects	546
8.2.1	Problem formulation	546
8.2.2	The evolution of potential-response variables . . .	549
8.2.3	Linear programming formulation	553
8.2.4	The natural bounds	558
8.2.5	Effect of treatment on the treated	559
8.2.6	Example: The effect of cholestyramine	561
8.3	Counterfactuals and Legal Responsibility	563
8.4	A Test for Instruments	568
8.5	Causal Inference From Finite Sample	571
8.5.1	Gibbs sampling	571
8.5.2	The effects of sample size and prior distribution .	574
8.5.3	Causal effects from clinical data with imperfect compliance	576
8.5.4	Bayesian estimate of single-event causation	578

8.6	Conclusion	581
9	PROBABILITY OF CAUSATION: INTERPRETA- TION AND IDENTIFICATION	583
9.1	Introduction	584
9.2	Necessary and Sufficient Causes: Conditions of Identifi- cation	589
9.2.1	Definitions, notation, and basic relationships . . .	589
9.2.2	Bounds and basic relationships under exogeneity .	596
9.2.3	Identifiability under monotonicity and exogeneity	599
9.2.4	Identifiability under monotonicity and non- exogeneity	603
9.3	Examples and Applications	608
9.3.1	Example-1: Betting against a fair coin	608
9.3.2	Example-2: The firing squad	611
9.3.3	Example-3: The effect of radiation on leukemia .	614
9.3.4	Example-4: Legal responsibility from experimen- tal and non-experimental data	618
9.3.5	Summary of results	622
9.4	Identification in Non-monotonic Models	624
9.5	Conclusions	629
10	THE ACTUAL CAUSE	633
10.1	Introduction: The Insufficiency of Necessary Causation .	635
10.1.1	Singular causes revisited	635

10.1.2	Preemption and the role of structural information	638
10.1.3	Overdetermination and quasi-dependence	642
10.1.4	Mackie's INUS condition	643
10.2	Production, dependence, and sustenance	648
10.3	Causal Beams and Sustenance-Based Causation	652
10.3.1	Causal beams: Definitions and implications . . .	652
10.3.2	Examples: From disjunction to general formulas .	657
10.3.3	Beams, preemption, and the probability of single- event causation	661
10.3.4	Path-switching causation	665
10.3.5	Temporal preemption	666
10.4	Conclusions	669

11 EPILOGUE: THE ART AND SCIENCE OF CAUSE AND EFFECT

(A public lecture delivered November 1996 as part of
the UCLA Faculty Research Lectureship Program.) 673