

ESSLLI 2021 Course Proposal

Causal Models in Linguistics, Philosophy, and Psychology

General Information

Title: Causal Models in Linguistics, Philosophy, and Psychology

Areas: Logic and Language

Category: Foundational Course

Instructors

Thomas Icard

Associate Professor, Department of Philosophy

Building 90

450 Jane Stanford Way

Stanford, CA, USA

Email: icard@stanford.edu

Url: <http://www.stanford.edu/~icard/>

Daniel Lassiter

Associate Professor, Department of Linguistics

Building 460

450 Jane Stanford Way

Stanford, CA, USA

Email: danlassiter@stanford.edu

Url: <http://web.stanford.edu/~danlass/>

Abstract

This course explains and motivates formal models of causation built around Bayes nets and structural equation models, a topic of increasing interest across multiple cognitive science fields, and describes their application to select problems in psychology, philosophy, and linguistics.

Motivation and Description

Theories of causation—long the provenance mainly of metaphysics, philosophy of science, and inferential statistics—have emerged recently as a crucial topic in the cognitive sciences. Much of the momentum for the recent surge of interest in causation has come from technical innovations in causal modeling due to Spirtes et al. (1993); Pearl (2000), among others. These models provide a formal semantics for theories of causation which have been applied fruitfully to numerous issues in philosophy, psychology, and linguistics. This course is designed to provide a unifying perspective on causal models, allowing students to access the literatures on causal models, to identify points of connection between the problems and solutions explored in the somewhat fragmented subliterations in different fields. Day 1 will begin with an overview, some general motivation for the importance of causation across multiple domains, and a brief description of two of the most influential kinds of causal models—causal Bayes nets and structural equation models—including comparison of these models with the familiar world-ordering, or “system of spheres”, models. These model classes will also be compared to causal modeling paradigms grounded in the modern practice of probabilistic

programming (e.g., Goodman et al. 2015), which offer a more procedural characterization of causal knowledge in a higher-order setting.

Much of the technical literature on causation has taken place in statistics and computer science, by theorists interested in problems of causal inference and reasoning (Spirtes et al., 1993; Pearl, 2000; Peters et al., 2017; Bareinboim et al., 2021). Day 2 will present, in a broadly accessible way, some of the key ideas and results from this literature. When can we in fact draw causal conclusions from observational or experimental data? What kinds of formal languages are suitable for encoding substantive causal assumptions? In the second half of this lecture we will transition to discuss the kinds of *psychological* biases and inferential strategies that infants and adult humans bring to causal learning and inference tasks (Glymour, 2001; Gopnik et al., 2004; Gopnik and Schultz, 2007). Some of the prominent ideas here have their origin in the earlier philosophical and computational work, but we also see genuinely new ways of thinking about causal induction. In particular we will highlight the idea—inspired by ideas from probabilistic programming—that human causal knowledge may be implemented in a more procedural manner, making use of ideas like *simulation* (Gerstenberg et al., 2021), a theme that will continue through the course.

Day 3 will survey issues around “actual” or token-level causation (which events actually cause which others), which stands in opposition to the more common application of causal models to type-level or generic causation (see, e.g., Halpern and Pearl 2005a,b). This opens up broader issues about *explanation* and related social concepts like *blame* and *responsibility*. We will discuss the relationship between type and token causation, and then survey formal and psychological results on the impact of statistical norms and other kinds of norms, or defaults, to judgments of causation (Halpern and Hitchcock, 2015; Icard et al., 2017).

The literatures on causation in philosophy, psychology, computer science, and linguistics still remain somewhat fragmented. Linguists and philosophers of language, in particular, are just beginning to explore the potential of formal models of causation to inform linguistic semantics and pragmatics. Days 4 and 5 of the course will concentrate on natural language. Day 4 will present the causal-modeling semantics for counterfactual conditionals due to Pearl which is increasingly influential in linguistics semantics (e.g., Pearl 2000; Schulz 2007; Sloman and Lagnado 2005; Dehghani et al. 2012; Kaufmann 2013; Lassiter 2017a,b), comparing its formal properties and empirical coverage to the classic approaches of Lewis and Stalnaker. Day 5 will consider the lexical semantics of items such as “cause, make, enable, prevent” (Wolff and Song, 2003; Sloman et al., 2009; Nadathur and Lauer, 2020; Gerstenberg et al., 2021), and the relationship between causal models and the semantics and pragmatics of causative verbs such as “kill” and “burn” (McCawley, 1978; Levin et al., 1995). We will also address pragmatic issues in some detail, focusing on the way that causal knowledge may influence the interpretation of conditionals, and several instances in which attention to natural language pragmatics is crucial in interpreting the results of psychological experiments which ask people to make explicit judgments involving counterfactuals or linguistic items such as “cause”.

Tentative Outline

Summarizing the schedule described above, the tentative outline is as follows:

1. Background, Motivation, and Technical Basics.
2. Causal Learning and Reasoning.
3. Actual Causation: Causes, Statistics, and Norms.
4. Causality and Conditionals.
5. Causal Verbs and the Meaning of ‘Cause’.

Expected Level and Prerequisites

No specific background in linguistics, psychology, or philosophy will be assumed. However, facility with logic, formal semantics, and probability theory will be useful in allowing students to absorb and focus on the concepts explored.

Background on Instructors

Thomas Icard has taught two ESSLLI courses (in 2014 and 2016) and organized three ESSLLI workshops (2014, 2015, and the Student Session in 2009). For the partner summer school in North America, NASSLLI, he has taught twice (2012 and 2018) and served as co-chair of the program committee twice (2014 and 2016). At Stanford he teaches a range of courses in logic and cognitive science to interdisciplinary audiences, where he won the Dean’s Award for Distinguished Teaching.

Daniel Lassiter has taught three ESSLLI courses (two in 2013 and one in 2014) and organized a workshop in 2021. He also taught NASSLLI courses in 2012 and 2014, and was chair of the ESSLLI Student Session in 2011. He teaches linguistics and cognitive science at Stanford.

References

- Bareinboim, E., Correa, J. D., Ibeling, D., and Icard, T. (2021). On Pearl’s hierarchy and the foundations of causal inference. In Dechter, R., Geffner, H., and Halpern, J., editors, *Probabilistic and Causal Inference: The Works of Judea Pearl*. ACM Turing Award Series.
- Dehghani, M., Iliev, R., and Kaufmann, S. (2012). Causal explanation and fact mutability in counterfactual reasoning. *Mind and Language*, 27(1):55–85.
- Gerstenberg, T., Goodman, N. D., Lagnado, D., and Tenenbaum, J. B. (2021). A counterfactual simulation model of causal judgment. *Psychological Review*.
- Glymour, C. N. (2001). *The Mind’s Arrows: Bayes Nets and Graphical Causal Models in Psychology*. MIT press.
- Goodman, N. D., Gerstenberg, T., and Tenenbaum, J. B. (2015). Concepts in a probabilistic language of thought.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., and Danks, D. (2004). A theory of causal learning in children: causal maps and bayes nets. *Psychological review*, 111(1):3.
- Gopnik, A. and Schultz, L., editors (2007). *Causal Learning: Psychology, Philosophy, and Computation*. Oxford University Press.
- Halpern, J. Y. and Hitchcock, C. (2015). Graded causation and defaults. *British Journal for the Philosophy of Science*, 66:413–457.
- Halpern, J. Y. and Pearl, J. (2005a). Causes and explanations: A structural-model approach. part i: Causes. *The British journal for the philosophy of science*, 56(4):843–887.
- Halpern, J. Y. and Pearl, J. (2005b). Causes and explanations: A structural-model approach. part ii: Explanations. *The British journal for the philosophy of science*, 56(4):889–911.
- Icard, T. F., Kominsky, J. F., and Knobe, J. (2017). Normality and actual causal strength. *Cognition*, 161:80–93.

- Kaufmann, S. (2013). Causal premise semantics. *Cognitive science*, 37(6):1136–1170.
- Lassiter, D. (2017a). Complex antecedents and probabilities in causal counterfactuals. In Cremers, A., van Gessel, T., and Roelofsen, F., editors, *Proceedings of the 21st Amsterdam Colloquium*, pages 45–54.
- Lassiter, D. (2017b). Probabilistic language in indicative and counterfactual conditionals. In Burgdorf, D., Collard, J., Maspong, S., and Stefánsdóttir, B., editors, *Semantics & Linguistic Theory (SALT) 27*, pages 525–546. CLC Publications.
- Levin, B., Hovav, M. R., and Keyser, S. J. (1995). *Unaccusativity: At the syntax-lexical semantics interface*, volume 26. MIT press.
- McCawley, J. D. (1978). Conversational implicature and the lexicon. In Cole, P., editor, *Syntax and Semantics: Pragmatics*, pages 245–259.
- Nadathur, P. and Lauer, S. (2020). Causal necessity, causal sufficiency, and the implications of causative verbs. *Glossa: a journal of general linguistics*, 5(1).
- Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge University Press.
- Peters, J., Janzing, D., and Schölkopf, B. (2017). *Elements of Causal Inference: Foundations and Learning Algorithms*. MIT Press.
- Schulz, K. (2007). *Minimal Models in Semantics and Pragmatics: Free choice, exhaustivity, and conditionals*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Sloman, S., Barbey, A. K., and Hotaling, J. M. (2009). A causal model theory of the meaning of cause, enable, and prevent. *Cognitive Science*, 33(1):21–50.
- Sloman, S. A. and Lagnado, D. A. (2005). Do we “do”? *Cognitive Science*, 29(1):5–39.
- Spirtes, P., Glymour, C., and Scheines, R. (1993). *Causation, prediction, and search*. MIT press.
- Wolff, P. and Song, G. (2003). Models of causation and the semantics of causal verbs. *Cognitive Psychology*, 47(3):276–332.