Causal learning in humans

Alison Gopnik
Dept. of Psychology
UC-Berkeley

Knowledge as an inverse problem

- **East Coast Solutions**
- Structured
- Abstract
- Innate
- Domain-specific
- Modularity

- **West Coast Solutions**
- Distributed
- Concrete
- Learned
- Domain-General
- Connectionism

The inverse problem of causal knowledge

- Structured
- Coherent
- Abstract
- Complex

- Novel
- Learned
- Related to conditional probabilities

- Solution: Mechanism

- Solution: Associationism

The Mechanistic Solution

Substantive Assumptions and Perceptual Causality

In adults: Michotte, 1962; Scholl & Tremoulet, 2002; Heider, 1958,

In children: Bullock et al. 1982; Shultz, 1982

In infants: Leslie, 1987; Oakes and Cohen, 1990; Watson, 1987; Meltzoff, 1988

The Associationist Solution

- **Formal Assumptions and Contingencies**
- In animals

Rescorla-Wagner, Classical and Operant Conditioning

- In adults

Shanks, 1985: Shanks & Dickinson, 1987; Cheng & Novick, 1992

In children?

Limits of Earlier Work

- Causes and effects specified beforehand
- Causal strength rather than causal structure
- No distinction or integration between intervention and observation
- No intermediate causes
- No unobserved causes

The theory theory

- In adult categorization - Murphy and Medin, 1985
- In infants and children

  Folk Psychology - Gopnik, 1988, Gopnik & Wellman, 1994

  Folk Biology- Carey, 1986, Gelman & Wellman, 1992,

  Folk Physics - Carey et al. 1988, Gopnik, 1988

Features of the Theory Theory

- **Static Features**
- Abstract, coherent, causal entities and rules, including unobserved entities
- **Functional Features**
- Provides predictions, interpretations, and explanations
- **Dynamic Features**
- Changes in the light of new evidence and experimentation

- . “Far too often in the past psychologists have been willing to abandon their own autonomous theorizing because of some infatuation with the current account of computation and neurology. We wake up one morning and discover that the account that looked so promising and scientific – S-R connections, gestaltist field theory, Hebbian cell assemblies – has vanished and we have spent another couple of decades trying to accommodate our psychological theories to it. We should summon up our self-esteem and be more stand-offish in future” (Gopnik & Meltzoff, 1997).

Bayes Nets to the Rescue

- **Representation**
- Coherent, complex causal entities and rules including unobserved entities, and integrating intervention and observation.
- **Inference**
- Normative algorithms for prediction, intervention and explanation
- **Learning**
- Normative algorithms for learning causal structure from data and experimentation

Normative Mathematical Solutions to Inverse Problems

- **Vision**
- 3-d representations of objects
- Projections of objects onto 2-d surfaces
- 2-d surfaces are the result of 3-d projections

- **Causation**
- Acyclic directed graphs
- Projections from graphs to probability distributions (Markov)
- Probability distributions are the result of projections (Faithfulness)

Cheng's Causal Powers

- **Patricia Cheng Psychological Review 1997.**
- Causal powers rather than association
- Representation equivalent to a Bayes-net noisy-or or noisy-and gate with no unobserved common causes
- Normative learning procedure for deriving causal power from conditional probability

Limitations of Cheng's Theory

- Causes and effects specified beforehand
- Causal strength rather than causal structure
- No intermediate causes (causal chains)
- No inference of unobserved causes
- No distinction between intervention and observation

Recent Empirical Work with Bayes Nets in Adults

- **Extensions of Cheng**
- Inhibitory causes and interactive causes (Novick & Cheng, in press).
- Chains and unobserved common causes

  Glymour 2002

Recent Empirical Work With Bayes Nets in Adults

- **Prediction and categorization in adults**
- Waldmann & Martignon, 1998, Waldmann & Hagmeyer, 2001, Rheder & Hastie, 2001
- Bayes Net representations to describe adult causal predictions and categorizations
- Different complex causal structures lead to different predictions and categorizations

Recent Empirical Work With Bayes Nets in Adults

- **Learning complex causal structure from observation in adults?**
- Steyvers, Tenenbaum et al (submitted)
- Danks (submitted)
- Lagnado & Sloman, 2002
- Individual variability

Recent Empirical Work With Bayes Nets in Adults

- **Learning complex causal structure from interventions and observations in adults.**
- Schulz, 2002; Gopnik et al. In press.
- Steyvers et al. submitted

Empirical Work with Bayes Nets in 3-4 Year Old Children

Gopnik, Glymour, Sobel, Schulz, Kushnir & Danks, Psychological Review, in press.

- Categorization - Gopnik & Sobel 2000
- Learning from conditional dependence - Gopnik, Sobel Schulz & Glymour, 2001; Sobel, Tenenbaum & Gopnik, submitted
- Learning from intervention and conditional dependence - Gopnik et al. In press
- Inferring unobserved causes - Gopnik et al. In press.

New Empirical Questions

- Probabilistic learning
- Statistical representativeness
- Interactions and Boolean combinations
- Unobserved variables
- Interactions with other types of knowledge

New Computational Questions

- Processing and memory constraints
- Data-mining versus bootstrapping
- Determining and reorganizing variables
- Conceptual change