Logical Argumentation, Abduction and Bayesian Decision Theory:

A Bayesian Approach to Logical Arguments and its Application to Legal Evidential Reasoning

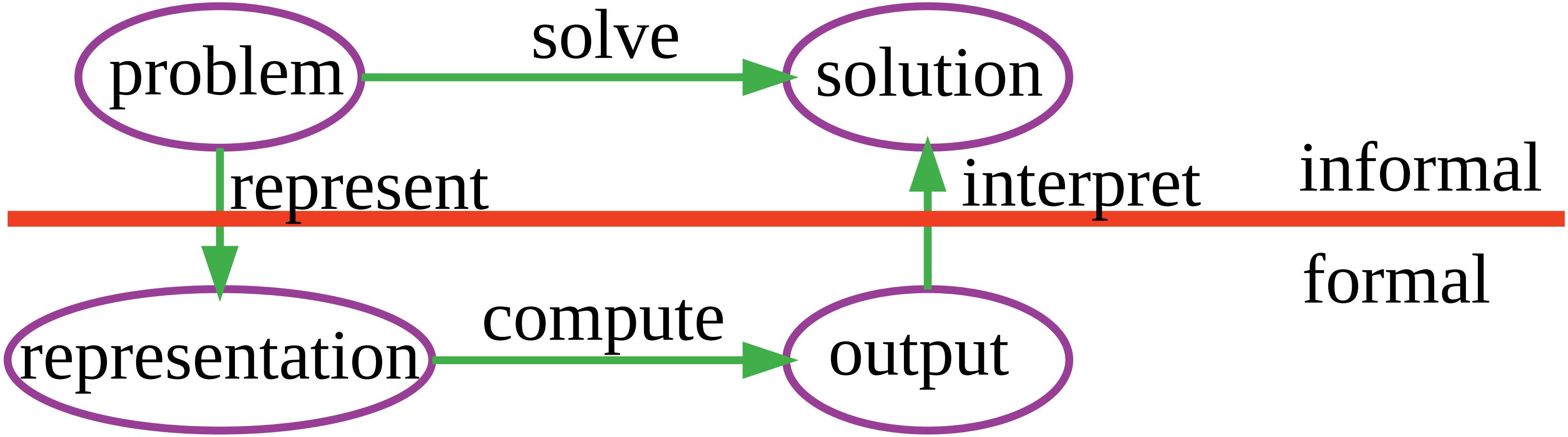
David Poole

University of British Columbia

Overview

- Knowledge representation, logic, decision theory.
- Independent Choice Logic
  - Logic programming + arguments
  - Abduction
  - Belief networks + first-order rule-structured conditional probabilities
- Peter Tillers' Example

Knowledge Representation



- Find compact / natural representations
- Exploit features of representation for computational gain.
- Tradeoff representational adequacy, efficient (approximate) inference and learnability

Normative Traditions

- **Logic**
  - Semantics (symbols have meaning)
  - Sound and complete proof procedures
  - Quantification over variables (relations amongst multiple individuals)
- **Decision Theory**
  - Tradeoffs under uncertainty
  - Probabilities and utilities

Independent Choice Logic

- **C**, the choice space is a set of alternatives.
  An alternative is a set of atomic choices.
  An atomic choice is a ground atomic formula.
  An atomic choice can only appear in one alternative.

- **F**, the facts is an acyclic logic program.
  No atomic choice unifies with the head of a rule.

- P_0 a probability distribution over alternatives:

$$\forall A \in \mathbf{C} \sum_{a \in A} P_0(a) = 1.$$

Meaningless Example

$$\mathbf{C} = \{\{c_1, c_2, c_3\}, \{b_1, b_2\}\}$$

$$\begin{aligned}
\mathbf{F} = \{\; & f \leftarrow c_1 \wedge b_1, && f \leftarrow c_3 \wedge b_2, \\
& d \leftarrow c_1, && d \leftarrow {\sim} c_2 \wedge b_1, \\
& e \leftarrow f, && e \leftarrow {\sim} d\}
\end{aligned}$$

$$P_0(c_1) = 0.5 \quad P_0(c_2) = 0.3 \quad P_0(c_3) = 0.2$$

$$P_0(b_1) = 0.9 \quad P_0(b_2) = 0.1$$

Semantics of ICL

- A **total choice** is a set containing exactly one element of each alternative in **C**.
- For each total choice τ there is a **possible world** w_τ.
- Proposition f is **true** in w_τ (written $w_\tau \models f$) if f is true in the (unique) stable model of $\mathbf{F} \cup \tau$.
- The probability of a possible world w_τ is

$$\prod_{a \in \tau} P_0(a).$$

- The **probability** of a proposition f is the sum of the probabilities of the worlds in which f is true.

Meaningless Example: Semantics

There are 6 possible worlds:

$$
\begin{array}{llllllll}
w_1 & \models & c_1 & b_1 & f & d & e & P(w_1) = 0.45 \\
w_2 & \models & c_2 & b_1 & \sim f & \sim d & e & P(w_2) = 0.27 \\
w_3 & \models & c_3 & b_1 & \sim f & d & \sim e & P(w_3) = 0.18 \\
w_4 & \models & c_1 & b_2 & \sim f & d & \sim e & P(w_4) = 0.05 \\
w_5 & \models & c_2 & b_2 & \sim f & \sim d & e & P(w_5) = 0.03 \\
w_6 & \models & c_3 & b_2 & f & \sim d & e & P(w_6) = 0.02
\end{array}
$$

$$P(e) = 0.45 + 0.27 + 0.03 + 0.02 = 0.77$$

Assumption-based reasoning

- Given background knowledge / facts F and assumables / possible hypotheses H,
- An explanation of g is a set D of assumables such that

  $F \cup D$ is consistent

  $F \cup D \models g$
- abduction is when g is given and you want D
- default reasoning / prediction is when g is unknown

Abductive Characterization of ICL

- The atomic choices are assumable.
- The elements of an alternative are mutually exclusive.

Suppose the rules are disjoint

$$\text{rules for } a \left\{ \begin{array}{l} a \leftarrow b_1 \\ \cdots \\ a \leftarrow b_k \end{array} \right. \qquad b_i \wedge b_j \text{ for } i \neq j \text{ can't be true}$$

$$P(g) = \sum_{E \text{ is a minimal explanation of } g} P(E)$$

$$P(E) = \prod_{h \in E} P_0(h)$$

Conditional Probabilities

$$P(g|e) = \frac{P(g \wedge e)}{P(e)} \quad \begin{array}{l} \longleftarrow \text{explain } g \wedge e \\ \longleftarrow \text{explain } e \end{array}$$

- Given evidence e, explain e then try to explain g from these explanations.
- The explanations of $g \wedge e$ are the explanations of e extended to also explain g.
- Probabilistic conditioning is abduction + prediction.

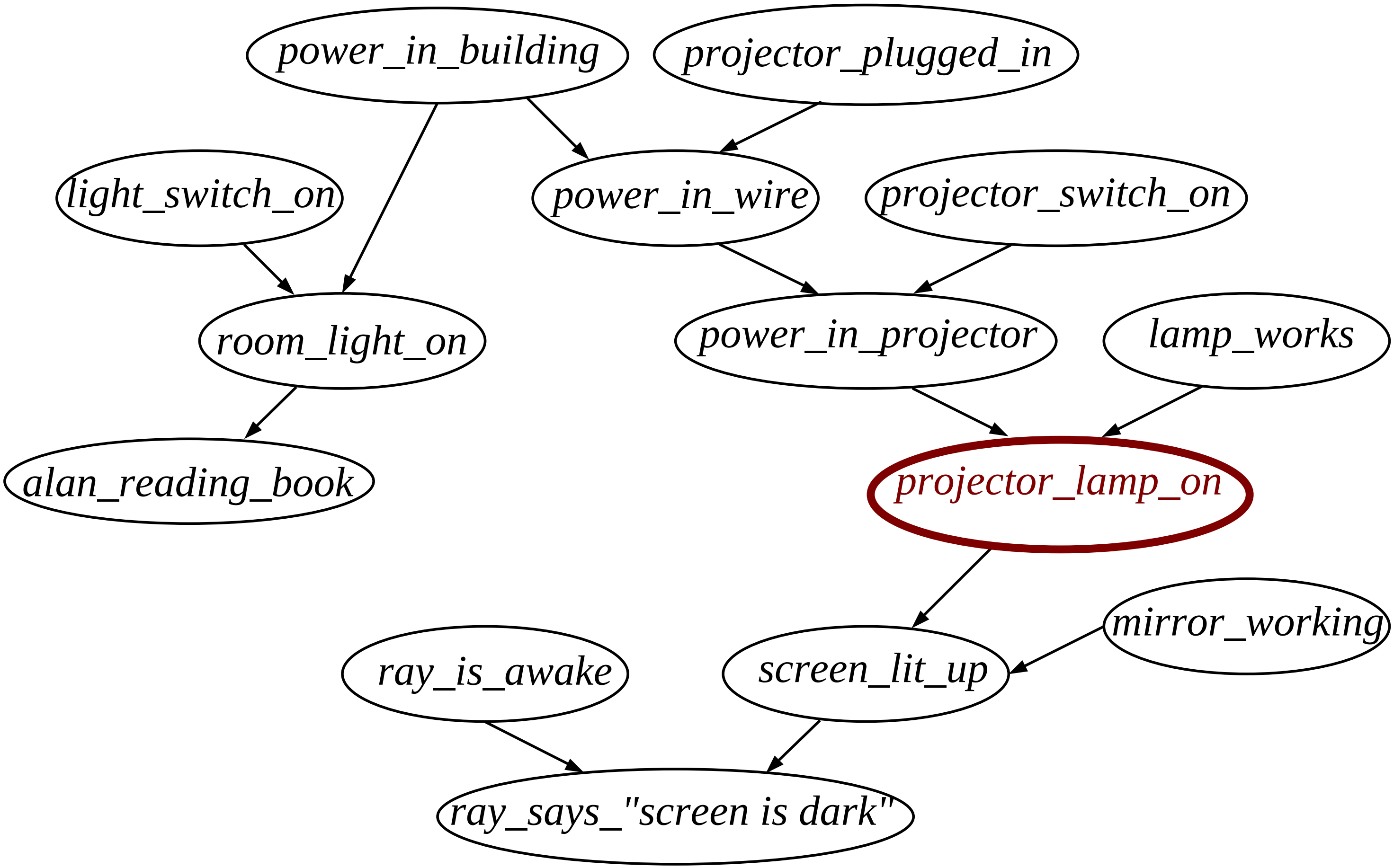
(Bayesian) Belief Networks

- Graphical representation of dependence.
- DAGs with nodes representing random variables.
- Arcs from parents of a node into the node.
- If $b_1, \cdots, b_k$ are the parents of a, we have an associated conditional probability table

$$P(a|b_1, \cdots, b_k)$$

- Doesn't specify how a variable depends on its parents.

Belief Network for Overhead Projector



Belief networks as logic programs

$$projector\_lamp\_on \leftarrow$$

$$power\_in\_projector \wedge$$

$$lamp\_works \wedge$$

$$projector\_working\_ok.$$ $\longleftarrow$ atomic choice

$$projector\_lamp\_on \leftarrow$$

$$power\_in\_projector \wedge$$

$$\sim lamp\_works \wedge$$

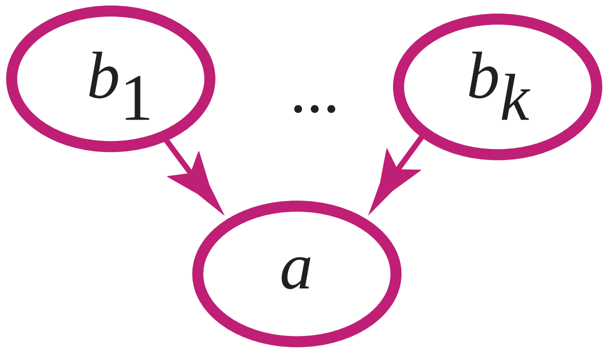
$$working\_with\_faulty\_lamp.$$

Probabilities of hypotheses

$$
\begin{aligned}
&P_0(\mathit{projector\_working\_ok}) \\
&\quad = P(\mathit{projector\_lamp\_on} \mid \\
&\qquad\qquad \mathit{power\_in\_projector} \wedge \mathit{lamp\_works})
\end{aligned}
$$

— provided as part of belief network

Mapping Belief networks into ICL



- Translated into the rules

$$a(V) \leftarrow b_1(V_1) \wedge \cdots \wedge b_k(V_k) \wedge h(V, V_1, \ldots, V_k)$$

- and the alternatives

$$\forall v_1 \cdots \forall v_k \{h(v, v_1, \ldots, v_k) | v \in domain(a)\} \in \mathbf{C}$$

Belief networks and the ICL

- The probabilities for the belief network and the ICL translation are identical.
- In the translation, the ICL requires the same number of probabilities as the belief network.
- Often the ICL theory is more compact than the corresponding conditional probability table.
- The probabilistic part of the ICL can be seen as a representation for the independence of belief networks.

What can we learn from the mapping?

ICL adds

- rule-structured conditional probability tables
- logical variables and negation as failure in rules
- arbitrary computation in the network
- choices by other agents
- algorithms

Belief networks add

- theory of causation
- algorithms
- ties to MDPs, Neural networks, …

Representing a domain in the ICL

- Axiomatize background knowledge causally
- Hypothesize what is going on in the world
- Condition on the observations of the specific case
  - Most observations have trivial explanations
  - Explanations with coherent story become more likely than those that assume independent coincidences

Tillers' Example: Observations

- $says(peter, wentto(peter, hvstore))$

  Peter says that he went to the Happy Valley Store.
- $says(peter, clerk\_at(harry, hvstore))$

  Peter says Harry was a clerk at the Happy Valley Store
- $says(peter, vicious\_sob(harry))$

  Peter says that Harry is a vicious SOB.
- $says(peter, observed(peter, blinding\_flash))$

  Peter says that he observed a blinding flash.
- $says(peter, says(doctor, shot(peter)))$

  Peter said that the doctor said he was shot.
- $says(peter, says(newspaper, disappeared(harry)))$

  Peter said that the newspaper said Harry disappeared.

Witness Honesty

$$
\begin{aligned}
says(P, F) &\leftarrow thinks\_true(P, F) \wedge \\
&\quad honest(P) \wedge \\
&\quad tr\_h(P, F). \\
says(P, F) &\leftarrow thinks\_true(P, F) \wedge \\
&\quad dishonest(P) \wedge \\
&\quad tr\_h(P, F).
\end{aligned}
$$

$$random([honest(P) : 0.999, dishonest(P) : 0.001]).$$

$$random([tr\_h(P, F) : 0.9999, untr\_h(P, F) : 0.0001]).$$

$$random([tr\_d(P, F) : 0.998, untr\_d(P, F) : 0.002]).$$

Peter May be Mistaken

$$thinks\_true(P,F) \leftarrow true(F) \land notmistaken\_t(P,F).$$

$$thinks\_true(P,F) \leftarrow false(F) \land mistaken\_f(P,F).$$

$$random([mistaken\_t(P,F) : 0.02, notmistaken\_t(P,F) : 0.98]).$$

$$random([mistaken\_f(P,F) : 0.06, notmistaken\_f(P,F) : 0.94]).$$

Why did he disappear?

$$true(disappeared(X)) \leftarrow$$
$$\quad left\_for\_no\_reason(X).$$
$$true(disappeared(X)) \leftarrow$$
$$\quad disappeared\_when\_criminal(X) \wedge$$
$$\quad committed\_crime(X).$$
$$random([disappeared\_when\_criminal(X) : 0.8,$$
$$\quad stayed\_when\_criminal(X) : 0.2]).$$
$$random([left\_for\_no\_reason(P) : 0.001,$$
$$\quad open\_in\_whereabouts(P) : 0.999]).$$

Shooting Explains Multiple Propositions

$$true(shot(P)) \leftarrow$$
$$\quad shot(X, P).$$
$$true(observed(P, blinding\_flash)) \leftarrow$$
$$\quad picture\_taken(P).$$
$$true(observed(P, blinding\_flash)) \leftarrow$$
$$\quad shot(X, P).$$
$$committed\_crime(X) \leftarrow$$
$$\quad shot(X, P).$$
$$random([picture\_taken(X) : 0.06, no\_picture(X) : 0.94]$$

Explaining why X shot P

$$
\begin{aligned}
& shot(X, P) \leftarrow \\
& \qquad means\_opportunity\_to\_shoot(X, P) \wedge \\
& \qquad motive\_to\_shoot(X, P) \wedge \\
& \qquad actually\_shot(X, P). \\
& means\_opportunity\_to\_shoot(X, P) \leftarrow \\
& \qquad at(X, L) \wedge at(P, L). \\
& at(X, L) \leftarrow true(clerk\_at(X, L)). \\
& at(X, L) \leftarrow true(wentto(X, L)). \\
& random([actually\_shot(X, P) : 0.01, \\
& \qquad didnt\_actually\_shoot(X, P) : 0.99]).
\end{aligned}
$$

Simplifications

- Reasonable probabilities
- Time
- Modalities
- Populations
- Subtleties of Language
- Utilities
- …

Conclusions

- ICL is a representation that combines logic and Bayesian decision theory.
- Inference is by variable elimination (marginalization, summing out a variable) and/or by enumerating the most likely explanations and bounding the error.
- Bayesian conditioning (abduction) gets dynamics of reasoning right.
- First-order rules let us reason about multiple individuals.
- Still many problems.