In most work on causal reasoning, an agent’s knowledge assigns one of three values to domain facts: yes, no, or maybe. These values are not sufficient, however, to represent the statistical information available in many interesting domains (arguably including most realistic domains). Thus some recent approaches to causal reasoning have concentrated on representation and inference with probabilistic degrees of belief [DK88, Han88, SH88]. We have found that probabilistic approaches to causality suffer from some of the same hard problems as traditional approaches. In particular the frame and qualification problems arise in subtle ways, and it is important to realize when such profound representational problems exit. The problems implicitly motivate the representational primitives of Dean and Kanazawa’s approach, but we find fault with their choice of primitives. In this paper, we first describe the persistence and qualification problems in a probabilistic setting, then explain and criticize Dean and Kanazawa’s solutions from a more traditional non-probabilistic causal framework.
Familiar Problems in Probabilistic Causal Reasoning

Topic areas: Commonsense Reasoning, Temporal Reasoning, Planning, Evidential Reasoning

Jay C. Weber
jay@cs.rochester.edu
(716)275-1443

Extended Abstract

In most work on causal reasoning, an agent's knowledge assigns one of three values to domain facts: yes, no, or maybe. These values are not sufficient, however, to represent the statistical information available in many interesting domains (arguably including most realistic domains). Thus some recent approaches to causal reasoning have concentrated on representation and inference with probabilistic degrees of belief [DK88, Han88, SH88].

We have found that probabilistic approaches to causality suffer from some of the same hard problems as traditional approaches. In particular the frame [Hay81, McD82, a.k.a. persistence] and qualification [McC77] problems arise in subtle ways, and it is important to realize when such profound representational problems exist. These problems implicitly motivate the representational primitives of Dean and Kanazawa's approach [DK88], but we find fault with their choice of primitives. In this paper, we first describe the persistence and qualification problems in a probabilistic setting, then explain and criticize Dean and Kanazawa's solutions to these problems, and lastly present a probabilistic causal framework that inherits its solutions from a more traditional non-probabilistic causal framework.

To review, suppose a non-probabilistic, reified, discrete time line representation of causality consisted of axioms of the following form:

$$\forall t \left( \bigwedge_{i=1}^{n} \text{Holds}(\phi_i, t) \land \text{Occurs}(e, t) \rightarrow \text{Holds}(\psi, t+1) \right)$$

Here the preconditions ($\phi_i$'s) and the effect ($\psi$) are properties, $e$ is an event, and $t + 1$ is the time point immediately after $t$. A rule of this form might say that if someone switches the light switch...
and the light and circuitry works, then the light will be on immediately afterwards. The persistence problem says that although these axioms provide a mechanism for deciding when properties change, there must also be a succinct representation of information that derives when properties do not change. The qualification problem (at least according to McCarthy [McC77]) says that it is not feasible in general to ascertain the truth of all the $\phi_i$'s.

A probabilistic causal theory will calculate the probability that $\psi$ holds at $t + 1$ given the probabilities that the $\phi_i$'s hold and $\epsilon$ occurs at $t$. For example, given a probability .5 that someone hits the light switch and an independent probability .9 that the light and circuitry works, we may say with probability (at least) .45 that the light is on immediately afterwards. Any number of these rules, however, do not supply an effective way to calculate how the probability of an arbitrary property is updated over time, such as the new probability of “candy on the coffee table” after the light is switched on. This is the persistence problem in a probabilistic setting. The qualification problem appears when one tries to calculate the probability that the light and circuitry works, which is a joint probability of a large number of conjuncts (intact wires, good bulb, working switch, ...) with potentially unknown truth values.

These problems are implicitly addressed in the work by Dean and Kanazawa [DK88] through two representational primitives: survivor functions and projection rules. In their approach, a property $\phi$ is assigned a function which we call $f_\phi$, defined by:

$$f_\phi(\Delta) = P(\text{Holds}(\phi, t) | \text{Holds}(\phi, t - \Delta))$$

Thus each property's survivor function specifies a "default" rate of decay for the probability of the property. In the light switch example, the probability of arbitrary properties such as "candy on the coffee table" can be calculated as long as the appropriate survivor function is specified. This is a solution to the persistence problem. The qualification problem is addressed in a more subtle way by projection rules of the form:

$$P(\text{Holds}(\psi, t + 1) | \bigwedge_{i=1}^{\kappa} \text{Holds}(\phi_i, t) \land \text{Ocurs}(\epsilon, t)) = \kappa$$

The difficulty of discovering and combining the $\phi_i$'s distributions constitutes the severity of the qualification problem. Notice, though, that in specific instances preconditions ($\phi_i$'s) can be removed from the rule as long as $\kappa$ is adjusted appropriately, presumably by making it smaller. If some small number of preconditions still produces a high $\kappa$, then perhaps the remaining preconditions can be safely ignored. This is a solution to the qualification problem.
We have described the issues underlying Dean and Kanazawa's representational primitives. As enlightening as they are, however, we find the following problems in their solutions:

- The reasoner does not have access to the assumptions underlying primitive survivor functions. For example, the survivor function for coffee table candy might be based on the average frequency of people using the room. If there is some information about context, say a dinner party, then it is unclear how much the function should be modified unless the derivation of the function is available.

- Survivor functions predict the persistence of properties holding but not the persistence of properties not holding. Projection rules predict only when properties begin to hold and not when they stop holding. It is unclear why these asymmetries appear. The absent directions could be easily supported by adding a definition of property negation and then specifying survivor functions and projection rules for the negations, but then it seems strange to use survivor functions instead of reasoning about causes for property negations.

- Because the definition of a survivor function is just in terms of duration, the rate of decay must be constant. This amounts to assuming that the distribution of causes for the negation is uniform. This assumption may not be correct for many influences, such as the fact that cars tend to get stolen more frequently at night.

- Survivor functions apply only until the property changes. The definition of conditional probabilities (derived from Bayes's rule) tells us that:

\[
P(\text{Holds}(\phi, t)) = \frac{P(\text{Holds}(\phi, t)|\text{Holds}(\phi, t - \Delta))P(\text{Holds}(\phi, t - \Delta))}{P(\text{Holds}(\phi, t - \Delta)|\text{Holds}(\phi, t))}
\] (1)

This rule tells us how to calculate future probabilities. The numerator is simply \( f_\phi \) multiplied by the past probability. The denominator is a new quantity, showing that the survivor function is not itself sufficient to update probabilities. We must assume that the denominator is always equal to one, which amounts to assuming that once a property ceases to hold it stays that way.

- If the reasoner makes the reasonable assumption that the domain is deterministic, then probabilities emerge from a lack of certain knowledge and not from randomness in the domain (quantum mechanics notwithstanding). Adding preconditions to a projection rule can bring \( \kappa \) arbitrarily close to one; a smaller \( \kappa \) therefore characterizes default assumptions about the
probabilities of omitted preconditions. Because of this, different projection rules can make
different implicit assumptions about the probabilities of the same properties. We feel that an
agent's assumptions about probabilities should not depend on what it wants to use them for.

We will now present a probabilistic causal reasoning approach of our own that answers these
criticisms by its choice of specific causal rules as representational primitives. Survivor functions
and projection rules are derived from these primitives, thereby making explicit the assumptions
involved in the derivations. This provides a mechanism for amending survivor functions and
projection rules based on contextual information and interacting assumptions.

A feature of our approach is that the probabilistic causal rules are a direct generalization
of non-probabilistic causal rules. Because of this, we incorporate developments from traditional
causal reasoning which solve common problems with probabilistic models. The rationale behind
the particular non-probabilistic theory we use is described elsewhere [Web88]; here we simply
identify four types of causal rules:

\[ \forall t \left[ \bigvee_{i=1}^{n} \text{OCCURS}(\epsilon_i, t) \rightarrow (\neg \text{HOLDS}(\phi, t) \land \text{HOLDS}(\phi, t + 1)) \right] \quad (2) \]

\[ \forall t \left[ (\neg \text{HOLDS}(\phi, t) \land \text{HOLDS}(\phi, t + 1)) \rightarrow \bigvee_{i=1}^{n} \text{OCCURS}(\epsilon_i, t) \right] \quad (3) \]

\[ \forall t \left[ \text{OCCURS}(\epsilon, t) \rightarrow \bigwedge_{i=1}^{n} \text{HOLDS}(\phi_i, t) \right] \quad (4) \]

\[ \forall t \left[ \bigwedge_{i=1}^{n} \text{HOLDS}(\phi_i, t) \rightarrow \text{OCCURS}(\epsilon, t) \right] \quad (5) \]

These formulas are templates of domain axioms which have constants instead of greek letters. Type
(2) axioms specify necessary effects of events; type (3) axioms specify what events could have been
the cause for a change; type (4) axioms specify necessary conditions for event occurrence; type (5)
axioms specify sufficient conditions for events. We can represent causes both for a property and for
its negation by adding a definition of negation, e.g. \( \text{HOLDS}(p, t) \equiv \neg \text{HOLDS}(\overline{p}, t) \). The persistence
problem is solved by using type (4) axioms to infer that none of the possible causes for the change
in a type (3) axiom could have occurred, thereby inferring that the property involved does not
change. The qualification problem is solved by adding default assumptions about properties that
appear in the antecedent of type (5) axioms.

We generalize these causal rules to probabilistic belief by interpreting the implications as partial
orders on the probabilities, i.e. the following schema templates:

\[
P(\bigvee_{i=1}^{\nu} \text{OCCURS}(\epsilon_i, t)) \leq P(\neg \text{HOLDS}(\phi, t) \land \text{HOLDS}(\phi, t + 1)) \tag{6}
\]

\[
P(\neg \text{HOLDS}(\phi, t) \land \text{HOLDS}(\phi, t + 1)) \leq P(\bigvee_{i=1}^{\nu} \text{OCCURS}(\epsilon_i, t)) \tag{7}
\]

\[
P(\text{OCCURS}(\epsilon, t)) \leq P(\bigwedge_{i=1}^{\nu} \text{HOLDS}(\phi_i, t)) \tag{8}
\]

\[
P(\bigwedge_{i=1}^{\nu} \text{HOLDS}(\phi_i, t)) \leq P(\text{OCCURS}(\epsilon, t)) \tag{9}
\]

This generalization reflects our stance that probabilities arise from the agent’s lack of certain information and not non-determinism in the domain. This interpretation of implication as bounds on probabilities actually follows from an approach to assigning probabilities to formulas based on distributions over models [Nil86], since the consequent is true in at least those models in which the antecedent is true.

A big win of this generalization is that we have inherited solutions to the persistence and qualification problems. We can calculate an upper bound on the probability of a property changing from the probabilities of event occurrences in axioms of type (7), and a lower bound from probabilities of event occurrences of type (6). The probability of these occurrences can be bounded by axioms of types (8) and (9). This solves the persistence problem. For example, suppose we generalize the following causal rule:

\[
\neg \text{HOLDS}(\text{LIT}, t) \land \text{HOLDS}(\text{LIT}, t + 1) \rightarrow \text{OCCURS}(\text{SWITCHED}, t) \lor \text{OCCURS}(\text{BRIDGED}, t)
\]

Here BRIDGED means that the circuit is completed somewhere along the wiring (say bare wires touch in the wall). The probabilistic version of this is:

\[
P(\neg \text{HOLDS}(\text{LIT}, t) \land \text{HOLDS}(\text{LIT}, t + 1)) \leq P(\text{OCCURS}(\text{SWITCHED}, t) \lor \text{OCCURS}(\text{BRIDGED}, t))
\]

Thus between any two time points the probability of the light becoming LIT can be bounded from above solely by the disjunctive probability of SWITCHED and BRIDGED occurrences.

We can derive projection rules from our causal rules by making explicit assumptions about the probabilities of properties, and in so doing solve the qualification problem. For example, consider adding the following probabilistic causal rules:

\[
P(\text{OCCURS}(\text{SWITCHED}, t)) \leq P(\text{HOLDS}(\text{LIT}, t + 1))
\]

\[
P(\text{OCCURS}(\text{SWITCHED}, t) \land \text{HOLDS}(\text{WIRING WORKS}, t)) \leq P(\text{OCCURS}(\text{SWITCHED}, t))
\]
The latter rule is of type (5) when \texttt{OCCURS(SWITCHPUSHED, t)} is expanded out to its necessary and sufficient properties (forces on the switch, etc.) These formulas imply the following projection rule:

$$P(\texttt{OCCURS(LIT, t)} | \texttt{OCCURS(SWITCHPUSHED, t)} \land \texttt{HOLDS(WIRINGWORKS, t)}) = 1$$

If house wiring is in general uniformly reliable across contexts, it will be more efficient to make an \textit{a priori} assumption about $P(\texttt{HOLDS(WIRINGWORKS, t)})$ for arbitrary $t$. If we assume independence of the conditions, the above causal rule becomes:

$$P(\texttt{OCCURS(SWITCHPUSHED, t)}) \cdot \kappa \leq P(\texttt{OCCURS(SWITCHED, t)})$$

where $\kappa$ is the assumed probability. This derives the projection rule:

$$P(\texttt{OCCURS(LIT, t + 1)} | \texttt{OCCURS(SWITCHPUSHED, t)}) = \kappa$$  \hspace{1cm} (10)

Thus we have derived the projection rule by \textit{explicitly} assigning a uniform distribution to the reliability of \texttt{WIRINGWORKS}. This is more desirable than asserting (10) as primitive, which does not even mention \texttt{WIRINGWORKS}.

We can derive survivor functions by first considering that since $f_p(\Delta) = f_p(1)^\Delta$, it suffices to derive $f_p(1)$. By definition, this quantity is $P(\texttt{HOLDS(p, t + 1)} | \texttt{HOLDS(p, t)})$, which equals $1 - P(\neg \texttt{HOLDS(p, t + 1)} | \texttt{HOLDS(p, t)})$. In our approach, this quantity is bounded by the probabilities of event occurrences in the following two types of causal axioms:

$$P(\bigvee_{i=1}^\nu \texttt{OCCURS(\epsilon_i, t)}) \leq P(\texttt{HOLDS(\phi, t + 1)})$$

$$P(\texttt{HOLDS(p, t)} \land \neg \texttt{HOLDS(p, t + 1)}) \leq P(\bigvee_{i=1}^\nu \texttt{OCCURS(\epsilon_i, t)})$$

The survivor function assumption that $P(\neg \texttt{HOLDS(p, t + 1)} | \texttt{HOLDS(p, t)})$ is constant over all $t$ amounts to the assumption that the probability of the event occurrences is the same for all $t$. In other words, survivor functions are exponential in character precisely because of the assumption that causes have a uniform distribution.

Lastly, in the full paper we consider the difficult task of calculating the joint distributions of properties holding and events occurring. Since we employ Baysian updating of probabilities, we can use the techniques of Pearl [Pea85] to represent and reason with priors and independence information.
References


