Symbolic Causal Networks

Adnan Darwiche
Rockwell International Science Center
444 High Street
Palo Alto, CA 94301
darwiche@rpal.rockwell.com

Judea Pearl
Computer Science Department
University of California
Los Angeles, CA 90024
pearl@cs.ucla.edu

Abstract

For a logical database to faithfully represent our beliefs about the world, one should not only insist on its logical consistency but also on its causal consistency. Intuitively, a database is causally inconsistent if it supports belief changes that contradict with our perceptions of causal influences — for example, coming to conclude that it must have rained only because the sprinkler was observed to be on. In this paper, we (1) suggest the notion of a causal structure to represent our perceptions of causal influences; (2) provide a formal definition of when a database is causally consistent with a given causal structure; (3) introduce symbolic causal networks as a tool for constructing databases that are guaranteed to be causally consistent; and (4) discuss various applications of causal consistency and symbolic causal networks, including nonmonotonic reasoning, Dempster–Shafer reasoning, truth maintenance, and reasoning about actions.

Introduction

Consider the database,

$$\Delta = \begin{array}{lcl} \mathsf{wet\_ground} & \supset & \mathsf{it\_rained} \\ \mathsf{sprinkler\_was\_on} & \supset & \mathsf{wet\_ground}, \end{array}$$

which entails no beliefs about whether it rained last night: $\Delta \not\models \mathsf{it\_rained}$ and $\Delta \not\models \neg\mathsf{it\_rained}$. If we tell this database that the sprinkler was on, it surprisingly jumps to the conclusion that it must have rained last night: $\Delta \cup \{\mathsf{sprinkler\_was\_on}\} \models \mathsf{it\_rained}$. This change in belief is counterintuitive! Given that we perceive no causal connection between the sprinkler and rain, we would not come to believe that it rained only because we observed the sprinkler on. That is, database Δ supports a belief change that contradicts common perceptions of causal influences, hence, it will be labeled *causally inconsistent.*

For another example of causal inconsistency, consider the database,

$$\Gamma = \mathsf{kind} \supset \mathsf{popular}, \ \mathsf{fat} \supset \neg\mathsf{popular}.$$

Initially, this database is ignorant about whether the person is kind: $\Gamma \not\models \mathsf{kind}$ and $\Delta \not\models \neg\mathsf{kind}$. However, once we tell the database that John is fat, it jumps to the strange result that John must be unkind: $\Gamma \cup \{\mathsf{fat}\} \models \neg\mathsf{kind}$. Here also, the database contradicts common perceptions of causal influences according to which no causal connection exists between kindness and weight. Therefore, database Γ is also causally inconsistent.

As it turns out, it is not uncommon for domain experts to construct databases that contradict with their own perceptions of causal influences, especially when the database is large enough and has multiple authors. The reason is that domain experts tend to focus on the plausibility of individual sentences rather than on the interactions among these sentences or how they would respond to future information.

But even when an expert is careful enough to construct a causally consistent database, it is not uncommon to turn it into a causally inconsistent one in the process of augmenting it with default assumptions. For example, an expert could have constructed the following database,

$$\Delta = \begin{array}{lcl} \mathsf{wet\_ground} \wedge \neg\mathsf{ab}_1 & \supset & \mathsf{it\_rained} \\ \mathsf{sprinkler\_was\_on} \wedge \neg\mathsf{ab}_2 & \supset & \mathsf{wet\_ground}, \end{array}$$

which is causally consistent because it remains ignorant about rain given information about the sprinkler. However, a nonmonotonic formalism that minimizes abnormalities would turn Δ into the database,

$$\Delta' = \begin{array}{lcl} \neg\mathsf{ab}_1 \wedge \neg\mathsf{ab}_2 & & \\ \mathsf{wet\_ground} \wedge \neg\mathsf{ab}_1 & \supset & \mathsf{it\_rained} \\ \mathsf{sprinkler\_was\_on} \wedge \neg\mathsf{ab}_2 & \supset & \mathsf{wet\_ground}, \end{array}$$

which is causally inconsistent because it finds $\mathsf{sprinkler\_was\_on}$ a sufficient evidence for $\mathsf{it\_rained}$: $\Delta' \not\models \mathsf{it\_rained}$ and $\Delta' \cup \{\mathsf{sprinkler\_was\_on}\} \models \mathsf{it\_rained}$.

In fact, we shall see later that causally inconsistent databases are also not uncommon in ATMS implementations of diagnosis systems and Dempster-Shafer reasoning, thus leading to counterintuitive results (Laskey & Lehner 1989; Pearl 1990).

Given the importance of causal consistency, and given the tendency to generate causally inconsistent databases, we shall concern ourselves in this paper with formalizing this notion in order to support do-

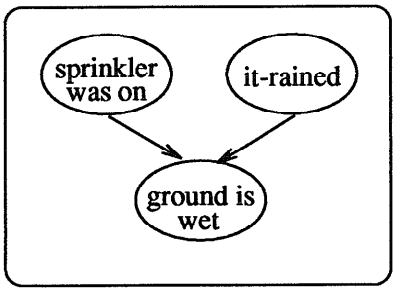

Figure 1: A causal structure.

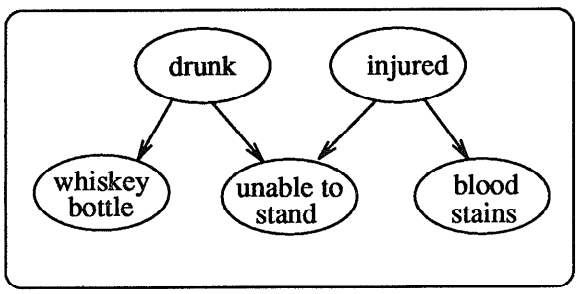

Figure 2: A causal structure.

main experts and commonsense formalisms in avoiding causally inconsistent databases.[1] In particular, we shall suggest the notion of a *causal structure* to represent perceptions of causal influences; provide a formal definition of when a database is causally consistent with a given causal structure; introduce *symbolic causal networks* as a tool for constructing causally consistent databases; and, finally, discuss various applications of symbolic causal networks, including nonmonotonic reasoning, truth maintenance, and reasoning about actions.

Causality and Belief Change

Since causal consistency is relative to specific perceptions of causal influences, formalizing causal consistency requires one to represent these perceptions formally. For this purpose, we will adopt causal structures, which are directed acyclic graphs that have been used extensively in the probabilistic literature for the same purpose (Pearl 1988b; Spirtes, Glymour, & Schienes 1993) — see Figures 1 and 2.[2]

The parents of a proposition p in a causal structure are those perceived to be its **direct causes.** The descendants of p are called its **effects** and the non-descendants of p are called its **non-effects.** In the structure of Figure 2, drunk and injured are the direct causes of unable_to_stand; unable_to_stand and whiskey_bottle are the effects of drunk; while injured and bloodstains are its non-effects.

Propositions that are relevant to the domain under consideration but do not appear in a causal structure are called the **exogenous propositions** of that structure. Exogenous propositions are the source of uncertainty in causal influences. They appear as *abnormality predicates* in nonmonotonic reasoning (Reiter 1987), as *assumption symbols* in ATMSs (de Kleer 1986), and as *random disturbances* in probabilistic models of causality (Pearl & Verma 1991). Each state of exogenous propositions will be referred to as an **extension.** For example, if ab_1 and ab_2 are the exogenous propositions of the structure in Figure 1, then $\neg\mathsf{ab}_1 \wedge \mathsf{ab}_2$ is an extension of that structure. Given an extension of a causal structure, there would be no longer any uncertainty about the causal influences it portrays.

The basic premise of this paper is that changes in our beliefs are typically constrained by the causal structures we perceive (Pearl 1988a). And the purpose of this section is to make these constraints precise so that a database is said to be consistent with a causal structure precisely when it does not contradict such constraints. The key to formalizing these constraints is the following interpretation of causal structures:

> The truth of each proposition in a causal structure is functionally determined by (a) the truth of its direct causes and (b) the truth of all exogenous propositions.

Following are some constraints on belief changes that are suggested by the above interpretation of causal structures (Pearl 1988b):

Common causes: In Figure 3a, the belief in c_2 should be independent of information about c_1 assuming that no information is available about e and that all exogenous propositions are known.

Indirect effects: In Figure 3b, the belief in e should be independent of information about c whenever proposition m and all exogenous propositions are known.

Common effects: In Figure 3c, the belief in e_2 should be independent of information about e_1 given that proposition c and all exogenous propositions are known.

These constraints on belief changes and others are summarized by **the principle of causal independence,** which is a version of the Markovian Condition in the probabilistic literature (Pearl 1988b; Spirtes, Glymour, & Schienes 1993). In a nutshell, the principle says that "once exogenous propositions and the direct causes of a proposition are known, the belief in that proposition should become independent

[1] The discussion in this paper is restricted to propositional databases.

[2] The results in this paper do not depend on a graphical representation of causal influences. For example, one can introduce a predicate *Direct_Cause* and proceed to axiomatize the contents of a causal structure.

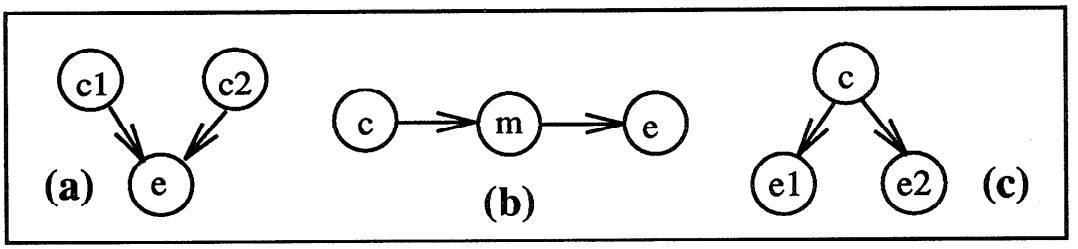

Figure 3: Causal interactions.

of information about its non–effects."[3]

We will formalize this principle in the remainder of this section and then use it later in defining causal consistency. But first, we need to define when database Δ finds X conditionally independent of Y given Z, that is, when adding information about Y to Δ does not change its belief in any information about X[4] given that Δ has full information about Z.[5] The following definition captures exactly this intuition:

Definition 1 (Conditional Independence) *Let X, Y, and Z be disjoint sets of atomic propositions and let $\hat{X}, \hat{Y}$, and $\hat{Z}$ be instantiations of these propositions, respectively. Database Δ finds X independent of Y given Z precisely when the logical consistency of $\Delta \cup \{\hat{Z}, \hat{X}\}$ and $\Delta \cup \{\hat{Z}, \hat{Y}\}$ implies the logical consistency of $\Delta \cup \{\hat{Z}, \hat{Y}, \hat{X}\}$.*

This is equivalent to saying that if Δ has full information about Z, then the addition of information about Y to Δ will not change its belief in any information about X.[6] For example, the database

$$\mathsf{it\_rained} \vee \mathsf{sprinkler\_was\_on} \equiv \mathsf{wet\_ground}$$

finds {it_rained} independent of {sprinkler_was_on}, but finds them dependent given {wet_ground}. Similarly, the database,

$$\mathsf{it\_rained} \supset \mathsf{wet\_ground}, \quad \mathsf{wet\_ground} \supset \mathsf{slippery\_ground},$$

finds {slippery_ground} dependent on {it_rained}, but finds them independent given {wet_ground}. Finally, the database,

$$\mathsf{battery\_is\_ok} \supset \mathsf{lights\_on}, \quad \neg\mathsf{battery\_is\_ok} \supset \neg\mathsf{car\_starts},$$

finds {lights_on} dependent on {car_starts}, but finds them independent given {battery_is_ok}.

We are now ready to state the principle of causal independence formally:

Definition 2 (Causal Independence) *Database Δ satisfies the principle of causal independence with respect to causal structure $\mathcal{G}$ precisely when (a) Δ is logically consistent and (b) for every extension $\mathcal{E}$ of $\mathcal{G}$ that is logically consistent with Δ, the database $\Delta \cup \{\mathcal{E}\}$ finds each proposition in $\mathcal{G}$ conditionally independent of its non–effects given its direct causes.*

Causal Consistency

Consider the following database,

$$\Delta = \begin{array}{lcl} \mathsf{wet\_ground} \wedge \neg\mathsf{ab}_1 & \supset & \mathsf{it\_rained} \\ \mathsf{sprinkler\_was\_on} \wedge \neg\mathsf{ab}_2 & \supset & \mathsf{wet\_ground}. \end{array}$$

This database does not satisfy the principle of causal independence with respect to the structure in Figure 1 because it finds sprinkler_was_on a sufficient evidence for it_rained under the extension $\neg\mathsf{ab}_1 \wedge \neg\mathsf{ab}_2$.

Note, however, that although database Δ does not satisfy the principle of causal independence, it does not contradict it either. Specifically, the extended database $\Delta \cup \{\mathsf{ab}_1 \vee \mathsf{ab}_2\}$ satisfies the principle because the added sentence $\mathsf{ab}_1 \vee \mathsf{ab}_2$ rules out the only extension, $\neg\mathsf{ab}_1 \wedge \neg\mathsf{ab}_2$, under which the database violates the principle. This suggests the following definition:

Definition 3 (Causal Consistency) *Let Δ be a database and let $\mathcal{G}$ be a causal structure. An extension $\mathcal{E}$ of $\mathcal{G}$ is causally consistent with Δ precisely when $\Delta \cup \{\mathcal{E}\}$ satisfies the principle of causal independence with respect to $\mathcal{G}$.*

That is, the extension $\neg\mathsf{ab}_1 \wedge \neg\mathsf{ab}_2$ above is causally inconsistent with database Δ, while the remaining extensions $\neg\mathsf{ab}_1 \wedge \mathsf{ab}_2$, $\mathsf{ab}_1 \wedge \neg\mathsf{ab}_2$, and $\mathsf{ab}_1 \wedge \mathsf{ab}_2$ are causally consistent with it.

We can further define a database as causally consistent precisely when it has at least one causally consistent extension. A definition of causal consistency was given in (Goldszmidt & Pearl 1992) for databases

[3] The requirement of knowing exogenous propositions renders this principle applicable to incomplete causal structures and constitutes, in fact, a definition of what information ought to be summarized by the exogenous variables.

[4] By information about a set of atomic propositions we mean a logical sentence constructed from these propositions.

[5] Database Δ has full information about atomic propositions Z if for each p in Z, Δ entails p or entails $\neg p$.

[6] This notion of independence is isomorphic to the one known in the literature on relational databases as multivalued embedded dependencies. Therefore, this notion of independence satisfies the graphoid axioms (Pearl 1988b), which include symmetry (X is independent of Y given Z precisely when Y is independent of X given Z).

containing defeasible conditionals (defaults) and was based on a probabilistic semantics of defeasible conditionals. Definition 3 is based on standard propositional semantics, which makes it directly applicable to commonsense formalisms based on classical logic. Moreover, causal consistency as we have defined it here is a semantical notion, independent of database syntax.

As we mentioned earlier, even when a domain expert is careful enough to construct a causally consistent database, it is not uncommon for a nonmonotonic formalism to turn it into a causally inconsistent one. Consider, for example, the database

$$\Delta = \begin{array}{lcl} \mathsf{whiskey\_bottle} \wedge \neg\mathsf{ab}_1 & \supset & \mathsf{drunk} \\ \mathsf{drunk} \wedge \neg\mathsf{ab}_2 & \supset & \mathsf{unable\_to\_stand} \\ \mathsf{unable\_to\_stand} \wedge \neg\mathsf{ab}_3 & \supset & \mathsf{injured} \\ \mathsf{injured} \wedge \neg\mathsf{ab}_4 & \supset & \mathsf{bloodstains}, \end{array}$$

which could easily be authored by a person perceiving the causal structure in Figure 2. When this database is fed to a nonmonotonic formalism that minimizes abnormalities, the formalism ends up augmenting it with the extension $\mathcal{E} = \{\neg\mathsf{ab}_1, \neg\mathsf{ab}_2, \neg\mathsf{ab}_3, \neg\mathsf{ab}_4\}$, which is causally inconsistent with Δ. Specifically, $\Delta \cup \mathcal{E}$ jumps to the conclusion bloodstains only because whiskey_bottle was observed: $\Delta \cup \mathcal{E} \not\models \mathsf{bloodstains}$ and $\Delta \cup \mathcal{E} \cup \{\mathsf{whiskey\_bottle}\} \models \mathsf{bloodstains}$.

With respect to the structure in Figure 2, database Δ has sixteen extensions corresponding to the different instantiations of the four abnormality predicates. As it turns out, all extensions containing $\neg\mathsf{ab}_2 \wedge \neg\mathsf{ab}_3$ are causally inconsistent, including the extension in which all abnormalities are minimized.

One way to avoid selecting these counterintuitive extensions is to inform nonmonotonic formalisms about causal consistency and provide them access to causal structures. A nonmonotonic formalism would then insist that only causally consistent extensions are selected. This approach should then lead to causal versions of existing nonmonotonic formalisms — for example, causal circumscription, causal default logic, and so on — in which each theory has a causal structure associated with it (Goldszmidt & Pearl 1992). Another solution is to feed nonmonotonic formalisms databases that already satisfy the principle of causal independence. In this case, nonmonotonic formalisms need not know about causality; they are guaranteed to stay out of danger because any extension that is logically consistent with the database is also causally consistent with it. For example, the database $\Delta \cup \{\mathsf{ab}_2 \vee \mathsf{ab}_3\}$ satisfies the principle of causal independence because the four extensions that are causally inconsistent are also logically inconsistent with the database. We will provide in the next section a systematic method for constructing databases that satisfy the principle of causal independence, thus eliminating the need to test for the causal consistency of extensions.

The connection between causality and nonmonotonic reasoning was discussed in (Pearl 1988a), where a system, called C-E, has been proposed for ensuring the faithfulness of default inferences to causal perceptions. In the C-E system, causality is represented by classifying defaults into either causal or evidential rules. We observe, however, that by classifying rules as such, one is indirectly communicating a causal structure. For example, the C-E rules wet_ground $\longrightarrow_E$ it_rained and sprinkler_was_on $\longrightarrow_C$ wet_ground implicitly encode the causal structure in Figure 1. In the C-E system, a causal structure is used procedurally to block default inferences that contradict with the structure. In the approach described here, a causal structure is used declaratively to classify extensions into those consistent with the structure and those inconsistent with it.

Separating the causal structure from the rule syntax has a number of merits. First, it keeps one within the realm of classical logic, which makes the approach applicable to logic-based formalisms such as circumscription and ATMSs. Next, using rules to communicate a causal structure may overburden domain experts. For example, to express the sentence $\neg$sprinkler_was_on $\wedge$ $\neg$it_rained $\supset$ $\neg$wet_ground, one needs three C-E rules: $\neg$sprinkler_was_on$\wedge\neg$it_rained $\longrightarrow_C$ $\neg$wet_ground, wet_ground $\wedge$ $\neg$it_rained $\longrightarrow_E$ sprinkler_was_on, and wet_ground $\wedge$ $\neg$sprinkler_was_on $\longrightarrow_E$ it_rained. Finally, it is not always clear whether a rule is causal or evidential. For example, given the rules it_rained $\longrightarrow_C$ $\neg$sprinkler_was_on[7] and wet_ground $\longrightarrow_E$ it_rained, it is not clear whether the rule $\neg$it_rained $\wedge$ wet_ground $\longrightarrow$ sprinkler_was_on is evidential or causal.

Symbolic Causal Networks

We will concern ourselves in this section with providing a systematic procedure that is guaranteed to generate databases that are not only causally consistent, but also satisfy the principle of causal independence. Therefore, the generated databases can be given to nonmonotonic formalisms without having to worry about whether adding assumptions will make them causally inconsistent.

The procedure we propose is that of constructing a symbolic causal network: a propositional database that is annotated by a causal structure.

A symbolic causal network has two components: A **causal structure** $\mathcal{G}$ that captures perceptions of causal influences and a set of **micro theories** capturing logical relationships between propositions and their direct causes — see Figures 4 and 5.

A micro theory for p is a set of clauses δ, where

1. each clause in δ refers only to p, its direct causes, and to exogenous propositions;
2. if δ entails a clause that does not mention p, then that clause must be vacuous.

Condition 1 ensures the locality of a micro theory to a proposition and its direct causes, while Condition 2

[7] We have a device that deactivates the sprinkler when it detects rain.

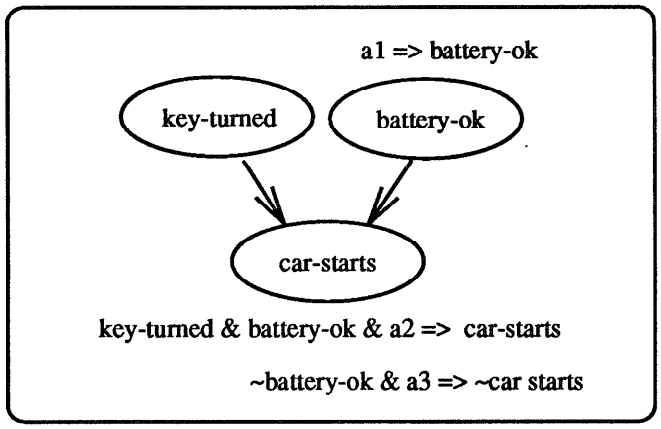

Figure 4: A symbolic causal network.

prohibits a micro theory for p from specifying a relationship between the direct causes of p.[8]

One can ensure the previous conditions by adhering to micro theories that contain only two types of material implications: $\psi \wedge \alpha \supset p$ and $\phi \wedge \beta \supset \neg p$, where

1. ψ and ϕ are constructed from the direct causes of p;
2. α and β are constructed from exogenous propositions; and
3. $\alpha \wedge \beta$ is unsatisfiable whenever $\psi \wedge \phi$ is satisfiable.

For example, the sentences $\mathsf{kind} \wedge \neg\mathsf{ab}_1 \supset \mathsf{popular}$ and $\mathsf{fat} \wedge \neg\mathsf{ab}_2 \supset \neg\mathsf{popular}$ do not constitute a micro theory for popular since $\neg\mathsf{ab}_1 \wedge \neg\mathsf{ab}_2$ and $\mathsf{kind} \wedge \mathsf{fat}$ are both satisfiable. This leads to the relationship $\neg\mathsf{ab}_1 \wedge \neg\mathsf{ab}_2 \wedge \mathsf{fat} \supset \neg\mathsf{kind}$ between weight and kindness, thus violating Condition 2 of micro theories.

We stress here that micro theories do not appeal to the distinction between evidential and causal rules. Formally, a micro theory contains standard propositional sentences and is constrained only by its locality (to specific propositions) and by what it can express about these propositions, both are characteristic of causal modeling. For example, one typically does not specify a relationship between the inputs to a digital gate by stating that certain input combinations would lead to conflicting predictions about the output.

If one induces a propositional database using a symbolic causal network — that is, by associating micro theories with the propositions of a causal structure — then one is guaranteed the following:

Theorem 1 *Let Δ be a database induced by a symbolic causal network having causal structure $\mathcal{G}$. Then Δ satisfies the principle of causal independence with respect to $\mathcal{G}$.*

[8]The satisfaction of such local conditions permits us to predict feasible scenarios in a backtrack-free manner (Dechter & Pearl 1991).

As a representational language, symbolic causal networks are complete with respect to databases that do not constrain the state of exogenous propositions:

Theorem 2 *Let Δ be a database satisfying the principle of causal independence with respect to a causal structure $\mathcal{G}$. If Δ is logically consistent with every extension of $\mathcal{G}$, then Δ can be induced by a symbolic causal network that has $\mathcal{G}$ as its causal structure.*

Applications of Symbolic Causal Networks

The basic motivation behind symbolic causal networks has been their ability to guarantee causal consistency. But symbolic causal networks can be viewed as the logical analogue of probabilistic causal networks; see Table 1. Therefore, many of the applications of probabilistic causal networks have counterparts in symbolic causal networks. We elaborate on some of these applications in this section. Other applications, such as diagnosis, are discussed elsewhere (Darwiche 1993).

Logical consistency One of the celebrated features of probabilistic causal networks is their ability to ensure the global consistency of the probability distribution they represent as long as the probabilities associated with each proposition in a causal structure are locally consistent. Symbolic causal networks provide a similar guarantee: As long as the micro theories associated with individual propositions satisfy their local conditions, the global database is guaranteed to be logically consistent. This is a corollary of Theorem 1.

Causal truth maintenance In the same way that probabilistic causal networks are supported by algorithms that compute probabilities (Pearl 1988b), symbolic causal networks are supported by algorithms that compute ATMS labels (Darwiche 1993; de Kleer 1986;

	Probabilistic causal network	Symbolic causal network
Represents	probability distribution + effects of actions	propositional database + effects of actions
Graphically Encodes	probabilistic independences + causal structure	logical independences + causal structure
Guarantees	probabilistic consistency Markovian condition	logical consistency causal independence
Computes	probabilities	arguments (ATMS labels)

Table 1: Analogous notions in probabilistic and symbolic causal networks.

Reiter & de Kleer 1987).[9] Therefore, symbolic causal networks inherit the applications of ATMSs. The important difference with traditional ATMSs, however, is that the database formed by a symbolic causal network is guaranteed (by satisfying causal independence) to protect us from conclusions that clash with our causal understanding of the domain. The importance of this property is best illustrated by an example that uses ATMSs to implement Dempster–Shafer reasoning (Laskey & Lehner 1989; Pearl 1990). Specifically, the Dempster–Shafer rules, $\mathsf{wet\_ground} \xrightarrow{.7} \mathsf{it\_rained}$ and $\mathsf{sprinkler\_was\_on} \xrightarrow{.9} \mathsf{wet\_ground}$, are typically reasoned about in an ATMS framework by constructing the database,

$$\Delta = \begin{array}{ll} \mathsf{wet\_ground} \wedge \mathsf{a}_1 & \supset \mathsf{it\_rained} \\ \mathsf{sprinkler\_was\_on} \wedge \mathsf{a}_2 & \supset \mathsf{wet\_ground}, \end{array}$$

and attaching probabilities .7 and .9 to the assumptions a_1 and a_2 (Laskey & Lehner 1989). Initially, the ATMS label of it_rained is empty and, hence, the belief in it_rained is zero. After observing sprinkler_was_on, however, the ATMS label of it_rained becomes $\mathsf{a}_1 \wedge \mathsf{a}_2$, which raises the belief in it_rained to $.7 * .9 = .63$. That is, the belief in it_rained increased from zero to .63 only because sprinkler_was_on was observed; see (Pearl 1990) for more related examples.

We get this counterintuitive behavior here because database Δ does not satisfy the principle of causal independence with respect to the causal structure in Figure 1. If the database satisfies this principle, the ATMS label of it_rained is guaranteed not to change as a result of adding sprinkler_was_on to the database — see (Darwiche 1993) for more details on this guarantee. For example, the database $\Delta \cup \{\neg \mathsf{a}_1 \vee \neg \mathsf{a}_2\}$ satisfies the principle of causal independence with respect to the causal structure in Figure 1. Therefore, the ATMS label it assigns to it_rained is empty and so is the label that $\Delta \cup \{\mathsf{sprinkler\_was\_on}\}$ assigns to it_rained. This guarantees that Dempster–Shafer belief in it_rained remains zero after sprinkler_was_on is observed.

[9] More precisely, symbolic causal networks compute *arguments*, which are logically equivalent to ATMS labels but are not necessarily put in canonical form (Darwiche 1993). Computing arguments is easier than computing ATMS labels. In fact, the complexity of computing arguments in symbolic causal networks is symmetric to the complexity of computing probabilities in probabilistic causal networks (Darwiche 1992; 1993).

Reasoning about actions Our focus so far has been the enforcement of causal constraints on belief changes that result from observations (belief revisions). But perceptions of causal influences also constrain (and in fact are often defined by) belief changes that result from interventions (named belief updates in (Katsuno & Mendelzon 1991)). For example, if we connect C in the first circuit of Figure 5 to a high voltage, we would come to believe that A and D will be set to OFF. But if we perform the same action on the second circuit, we would not come to the same belief. Note, however, that the two circuits have the same logical description, and do not mention external interventions explicitly, which means that they would lead to equivalent belief changes under all observations.

The reason why the same action leads to different results from two logically equivalent descriptions is that the descriptions are accompanied by different causal structures. It is the constraints encoded by these structures that govern our expectations regarding interventions in these circuits (Goldszmidt & Pearl 1992). A related paper (Darwiche & Pearl 1994) provides a specific proposal for predicting the effect of action when domain knowledge is represented using a symbolic causal network, showing also how the frame, ramification and concurrency problems can be handled effectively in this context. The key idea is that micro theories allow one to organize causal knowledge efficiently in terms of just a few basic mechanisms, each involving a relatively small number of propositions. Each external elementary action overrules just one mechanism leaving the others unaltered. The specification of an action then requires only the identification of the mechanism which is overruled by that action. Once this is identified, the overall effect of the action (or combinations thereof) can be computed from the immediate effect of the action, combined with the constraints imposed by the remaining mechanisms. Thus, in addition to encoding a set of current beliefs, and belief changes due to hypothetical observations, a causal database constrains how future beliefs would change in response to every hypothetical action or actions combination (Pearl 1993). These latter constraints can in fact be viewed as the defining characteristic of causal relationships, of which the Markovian condition is a byproduct (Pearl & Verma 1991).

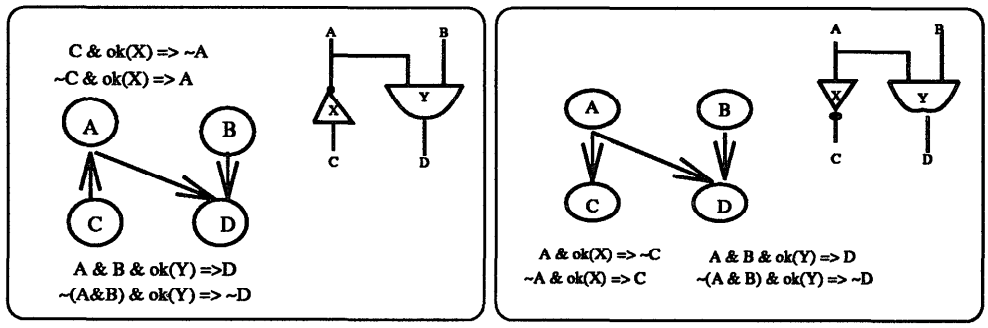

Figure 5: Different symbolic causal networks leading to logically equivalent databases.

Conclusion

If a classical logic database is to faithfully represent our beliefs about the world, the database must be consistent with our perceptions of causal influences. In this paper, we proposed a language for representing such perceptions and then formalized the consistency of a logical database relative to a given causal structure. We also introduced symbolic causal networks as tools for constructing databases that are guaranteed to be causally consistent. Finally, we discussed other applications of symbolic causal networks, including the maintenance of logical consistency, nonmonotonic reasoning, Dempster–Shafer reasoning, truth maintenance, and reasoning about actions and change.

Acknowledgments

The research was partially supported by ARPA contract #F30602-91-C-0031, Air Force grant #AFOSR 90 0136, NSF grant #IRI-9200918, and Northrop–Rockwell Micro grant #93-124. This work benefitted from discussions with Moises Goldszmidt and Sek–Wah Tan.

References

Darwiche, A., and Pearl, J. 1994. Symbolic causal networks for reasoning about actions and plans. Working notes: AAAI Spring Symposium on Decision–Theoretic Planning.

Darwiche, A. 1992. *A Symbolic Generalization of Probability Theory.* Ph.D. Dissertation, Stanford University.

Darwiche, A. 1993. Argument calculus and networks. In *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence (UAI)*, 420–427.

de Kleer, J. 1986. An assumption-based TMS. *Artificial Intelligence* 28:127–162.

Dechter, R., and Pearl, J. 1991. Directed constraint networks: A relational framework for causal modeling. In *Proceedings, 12th International Joint Conference of Artificial Intelligence (IJCAI-91)*, 1164–1170.

Goldszmidt, M., and Pearl, J. 1992. Rank–based systems: A simple approach to belief revision, belief update and reasoning about evidence and actions. In *Proceedings of the Third Conference on Principles of Knowledge Representation and Reasoning*, 661–672. Morgan Kaufmann Publishers, Inc., San Mateo, California.

Katsuno, H., and Mendelzon, A. 1991. On the difference between updating a knowledge base and revising it. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference*, 387–394.

Laskey, K. B., and Lehner, P. E. 1989. Assumptions, beliefs, and probabilities. *Artificial Intelligence* 41(1):65–77.

Pearl, J., and Verma, T. 1991. A theory of inferred causation. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference*, 441–452.

Pearl, J. 1988a. Embracing causality in default reasoning. *Artificial Intelligence* 35:259–271.

Pearl, J. 1988b. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann Publishers, Inc., San Mateo, California.

Pearl, J. 1990. Which is more believable, the probably provable or the provably probable. In *Proceedings, CSCSI-90, 8th Canadian Conference on AI*, 1–7.

Pearl, J. 1993. From bayesian networks to causal networks. Technical Report R-195-LLL, Cognetive Systems Laboratory, UCLA. (Short version in *Statistical Science*, Vol. 8, No. 3 (1993), pp. 266-269.).

Reiter, R., and de Kleer, J. 1987. Foundations of assumption-based truth maintenance systems: Preliminary report. In *Proceedings of AAAI*, 18? -188. AAAI.

Reiter, R. 1987. Nonmonotonic reasoning. *Ann. Rev. Comput. Sci* 2:147–186.

Spirtes, P.; Glymour, C.; and Schienes, R. 1993. *Causation, Prediction, and Search.* New York: Springer–Verlag.