A Logical Notion of Conditional Independence

Properties and Applications

Adnan Darwiche*
Rockwell Science Center
1049 Camino Dos Rios
Thousand Oaks, CA 91360
darwiche@risc.rockwell.com

May 15, 1997

Abstract

We propose a notion of conditional independence with respect to propositional logic and study some of its key properties. We present several equivalent formulations of the proposed notion, each oriented towards a specific application of logical reasoning such as abduction and diagnosis. We suggest a framework for utilizing logical independence computationally by structuring a propositional logic database around a directed acyclic graph. This structuring explicates many of the independences satisfied by the underlying database. Based on these structural independences, we develop an algorithm for a class of structured databases that is not necessarily Horn. The algorithm is linear in the size of a database structure and can be used for deciding entailment, computing abductions and diagnoses. The presented results are motivated by similar results in the literature on probabilistic and constraint-based reasoning.

1 Introduction

A major factor in slowing down logical computations is that reasoners tend to consider irrelevant parts of a given database when computing answers to queries. This has prompted a considerable amount of research on the notion of *irrelevance,* with the goal of identifying these irrelevant parts of a database so they can be avoided by logical reasoners [13, 15, 24, 23].

Irrelevance has also been the subject of extensive research in probabilistic reasoning, where it is typically referred to as *independence* [19]. The scope of independence in probability, however, seems to be bigger than its scope in logic for at least two reasons. First, although there is a standard definition of independence in probability, there seems to be no agreement on a definition of irrelevance in logic. This also contributes to the lack of a comprehensive theory of logical irrelevance similar to the one for probabilistic independence. Second, the computational utility that irrelevance brings to logical reasoning is perceived as a *luxury* instead of a *necessity* since influential algorithms for logical reasoning have not been based on irrelevance. This is contrary to what one finds in probabilistic reasoning where independence is the building block of almost all state-of-the-art algorithms.

Motivated by the role of independence in probability, our goal in this paper is to show that independence can play a similar role in logical reasoning, at least in the context of propositional logic. Therefore, our definition of conditional independence with respect to propositional databases resembles the definition of conditional independence with respect to probability distributions. We use Logical Conditional Independence, LCI, to refer to the proposed definition.

Our contribution here is not LCI *perse,* but rather (a) the framework we propose for exploiting it computationally and (b) the various formulations of LCI that we provide with respect to different reasoning tasks. We show that LCI introduces to propositional reasoning many of the tools and techniques that conditional

*Currently at: Department of Mathematics, American University of Beirut, PO Box 11 - 236, Beirut, Lebanon. Email: darwiche@aub.edu.lb.

independence has introduced to probabilistic reasoning. This includes a paradigm for logical reasoning in which the amount of independence information decides the computational complexity of reasoning, just as independence controls the complexity of probabilistic reasoning. The various formulations of LCI that we present make it clear how independence can be computationally valuable to the corresponding reasoning tasks. We utilize these formulations by developing an algorithm that can be used for deciding entailment, computing abductions and diagnoses. The algorithm provides a good example of how to use independence information when deriving algorithms for logical reasoning.

In the following section, we provide a more extended introduction to the proposed notion of independence where we explain the choices we had to make in developing it. We also outline the structure of the paper in light of the results to be presented.

2 Key Choices for Independence

When formulating a notion of independence, one faces a number of choice points. We devote this section to enumerating some of these points and to presenting our position on them. This helps in relating our approach to the spectrum of other existing approaches. It also provides a good opportunity for outlining the structure of the paper.

What objects can appear in an independence relation? One finds proposals in which sentences, atomic propositions, predicates, even algorithms appear as part of an independence relation [13]. In our proposal, independence is a relation between four objects: a propositional database Δ and three sets of atomic propositions X, Y and Z. Specifically, LCI decides whether the database Δ finds X independent of Y given Z. When the independence holds, we write $Ind_{\Delta}(X, Z, Y)$ and refer to it as an LCI assertion. This is similar to probabilistic independence which tells us whether a probability distribution finds X independent of Y given Z.

What decides the correctness of a definition of independence? In considering the literature on independence (both probabilistic and logical), one finds two main positions:

1. *a philosophical* position that starts with postulating some intuitive properties of independence and then proposes a definition that adheres to these properties. In most of these approaches, independence is formulated in terms of *belief change,* that is, formalizing the irrelevance of certain information to certain beliefs [11, 19]. The probabilistic notion of independence can clearly be motivated on these grounds where belief in a proposition corresponds to its probability.

2. *a pragmatic* position, where independence is not an absolute notion but rather a task specific one. That is, there is no correct or incorrect definition of independence but rather a useful or not-very-useful one. For example, in deciding an entailment test $\Delta \models \alpha$, a definition of independence may target the identification of sentences in Δ that if removed from Δ will not affect the test result.

LCI, at least as developed in this paper, is meant to be a pragmatic notion. In fact, we provide different and equivalent formulations of LCI, each explicating its usefulness to a specific reasoning task. The tasks we cover are: deciding entailment and satisfiability, which are dual tasks, and computing abductions and diagnoses, which are also dual tasks. This leads to a total of four formulations of LCI that are meant to explicate its computational role in these different reasoning tasks.

LCI, however, does have a formulation based on belief change, which happens to be the closest one to probabilistic independence. We therefore start in Section 3 with this formulation and then lead into the other four formulations in Section 4 (entailment and satisfiability) and Section 5 (abduction and diagnosis).

How should independence be used computationally? If one is adopting a pragmatic approach to independence, then the usage of independence will depend on the reasoning task of interest. A very popular use of independence though is in pruning a database before attempting certain computations. For example, in testing for logical entailment, the current practice is to use independence for reducing an entailment test

$\Delta \models \alpha$ into a simpler test $\Delta' \models \alpha$, where Δ' is obtained by removing sentences from Δ that are irrelevant to the test.

Although LCI will support such usage, its main computational role is different. In deciding entailment, for example, LCI will be used to *decompose* a global entailment test $\Delta \models \alpha$ into a number of local tests $\Delta_1 \models \alpha_1$, $\Delta_2 \models \alpha_2$, ..., $\Delta_n \models \alpha_n$, where each Δ_i is so small that its corresponding test $\Delta_i \models \alpha_i$ can be performed in constant time.

To introduce the computational role of LCI, we point out that the following decompositions are generally *not valid* in logic:

1. *Entailment:* Decomposing an entailment test $\Delta \models \alpha \vee \beta$ into two simpler tests $\Delta \models \alpha$ and $\Delta \models \beta$.

2. *Satisfiability:* Decomposing a satisfiability test with respect to $\Delta \cup \{\alpha \wedge \beta\}$ into two tests with respect to the smaller databases $\Delta \cup \{\alpha\}$ and $\Delta \cup \{\beta\}$.

3. *Abduction:* Decomposing the abductions of a finding $\alpha \vee \beta$ into the abductions of α and those of β.

4. *Diagnosis:* Decomposing the diagnoses of an observation $\alpha \wedge \beta$ into the diagnoses of α and those of β.

We shall demonstrate, however, that each one of these decompositions becomes valid when certain independences hold between the atoms appearing in α and those appearing in β. In fact, Sections 4 and 5 provide examples of how such decompositions can be used to decompose a global computation into a number of local computations.

What is the source of independence information? Most existing approaches attempt to *discover* independence information by pre-processing a given database [15, 23, 17, 16]. Although this is consistent with our utilization of independence, we advocate a different strategy that is motivated by the following result:

> If a database Δ is *graphically structured* — that is, satisfies some conditions that are dictated by a directed acyclic graph — then the topology of the structure reveals many of the independences satisfied by the database.

Therefore, instead of automatically discovering the independences satisfied by a database, we will propose *explicating* them by constructing a structured database in the first place. Such databases are defined precisely in Section 6, which also contains a key result on how to read independences from the topology of a database structure.

What is the measure of success when using independence? When using independence to prune a database, the measure of success is the degree of pruning it allows. But when using independence to decompose a global reasoning task into local tasks, the measure of success is the number of local tasks that result from the decomposition. In Section 7, we identify a class of structured databases for which we can decompose a reasoning task (entailment, abduction, diagnosis) into a number n of local tasks, where n is linear in the number of arcs and nodes of the database structure. We also discuss extensions of the algorithm to other classes of structured databases.

We start in the next three sections by providing five formulations of LCI, each oriented towards a specific reasoning task. The different formulations are then shown to be equivalent.

3 Belief Change Formulation

Common intuitions about independence suggest that it is strongly related to the notion of belief change. Therefore, many expect a formal definition of independence to be based on such a notion.

Although we shall present a few formulations of LCI in this paper, it is for these intuitions that we start with a formulation in terms of belief change. This formulation also happens to be the one most resembling probabilistic independence.

Before we present this first formulation, we need to define the notion of information.

Definition 1 *Information about a set of atomic propositions X is a propositional sentence constructed from these propositions. We use $\tilde{X}$ to denote information about propositions X.*

For example, if X contained the atoms p and q, then $p \vee q$, $\neg p$ and $p \supset q$ are three pieces of information about X.

Definition 2 *Full information about a set of atomic propositions X is a conjunction of literals, one for each atomic proposition in X. We use $\hat{X}$ to denote full information about propositions X.*

For example, there are four pieces of full information about atoms p and q: $p \wedge q$, $p \wedge \neg q$, $\neg p \wedge q$ and $\neg p \wedge \neg q$. We will also use the term *conjunctive clause* over X to mean full information about X.

Intuitively, the first formulation of LCI says that atoms X are independent of Y given Z with respect to database Δ if obtaining information about Y is irrelevant to entailing more information about X given that we have already obtained full information about Z:

Definition 3 *Let X, Y, and Z be disjoint sets of atomic propositions. Database Δ finds X independent of Y given Z, written $Ind^b_\Delta(X, Z, Y)$, iff*

$$\Delta \cup \{\hat{Z}\} \models \tilde{X} \textit{ precisely when } \Delta \cup \{\hat{Z}, \tilde{Y}\} \models \tilde{X}$$

for all $\tilde{X}, \tilde{Y}, \hat{Z}$ such that $\Delta \cup \{\hat{Z}, \tilde{Y}\}$ is consistent. If Z is empty, we simply say that X is independent of Y.

The database $\Delta \cup \{\hat{Z}\}$ results from adding the full information $\hat{Z}$ to Δ, and the database $\Delta \cup \{\hat{Z}, \tilde{Y}\}$ results from adding the extra information $\tilde{Y}$. Definition 3 says that X is independent of Y given Z if both of these databases entail the same information about X, that is, the information about Y is irrelevant.

Examples

To further illustrate Definition 3, consider the following database,

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

Let $X = \{\mathsf{sprinkler\_was\_on}\}$, $Y = \{\mathsf{it\_rained}\}$ and $Z = \emptyset$. Then $Ind^b_\Delta(X, Z, Y)$ because

$$\begin{array}{rcl} \Delta & \not\models & \mathsf{sprinkler\_was\_on} \\ \Delta \cup \{\mathsf{it\_rained}\} & \not\models & \mathsf{sprinkler\_was\_on} \\ \Delta \cup \{\neg\mathsf{it\_rained}\} & \not\models & \mathsf{sprinkler\_was\_on}, \end{array}$$

and, moreover,

$$\begin{array}{rcl} \Delta & \not\models & \neg\mathsf{sprinkler\_was\_on} \\ \Delta \cup \{\mathsf{it\_rained}\} & \not\models & \neg\mathsf{sprinkler\_was\_on} \\ \Delta \cup \{\neg\mathsf{it\_rained}\} & \not\models & \neg\mathsf{sprinkler\_was\_on}. \end{array}$$

That is, adding information about $\mathsf{it\_rained}$ to the database does not change the belief in any information about $\mathsf{sprinkler\_was\_on}$. We say in this case that database Δ finds $\mathsf{sprinkler\_was\_on}$ independent of $\mathsf{it\_rained}$.

However, if we let $Z = \{\mathsf{wet\_ground}\}$, then we no longer have $Ind^b_\Delta(X, Z, Y)$ because

$$\begin{array}{rcl} \Delta \cup \{\mathsf{wet\_ground}\} & \not\models & \mathsf{sprinkler\_was\_on} \\ \Delta \cup \{\mathsf{wet\_ground}, \neg\mathsf{it\_rained}\} & \models & \mathsf{sprinkler\_was\_on}. \end{array}$$

That is, in the presence of complete information about $\mathsf{wet\_ground}$ in the database, adding information about $\mathsf{it\_rained}$ does change the belief in some information about $\mathsf{sprinkler\_was\_on}$. We say in this case that the database Δ finds $\mathsf{sprinkler\_was\_on}$ dependent on $\mathsf{it\_rained}$ given $\mathsf{wet\_ground}$.

Consider now the database,

$$\Delta = \{\mathsf{it\_rained} \supset \mathsf{wet\_ground}, \mathsf{wet\_ground} \supset \mathsf{wet\_shoes}\}.$$

Let $X = \{\mathsf{wet\_shoes}\}$, $Y = \{\mathsf{it\_rained}\}$ and $Z = \emptyset$. Then we do not have $Ind^b_\Delta(X, Z, Y)$ because

$$\begin{array}{rcl} \Delta & \not\models & \mathsf{wet\_shoes} \\ \Delta \cup \{\mathsf{it\_rained}\} & \models & \mathsf{wet\_shoes}. \end{array}$$

That is, adding information about it_rained to the database changes the belief in some information about wet_shoes. We say in this case that database Δ finds wet_shoes dependent on it_rained.

However, if we let $Z = \{\mathsf{wet\_ground}\}$, then we have $Ind^b_\Delta(X, Z, Y)$ because, in the presence of complete information about wet_ground in the database, adding information about it_rained does not change the belief in any information about wet_shoes — we leave this for the reader to verify. We say in this case that the database Δ finds wet_shoes independent of it_rained given wet_ground.

Properties of LCI

Definition 3 of LCI satisfies the following properties.

Theorem 1 *Let X, Y, Z and L be disjoint sets of atomic propositions and let Δ be a propositional database. Then*

1. $Ind^b_\Delta(X, Z, \emptyset)$
2. $Ind^b_\Delta(X, Z, Y)$ *precisely when* $Ind^b_\Delta(Y, Z, X)$
3. $Ind^b_\Delta(X, Z, Y)$ *and* $Ind^b_\Delta(L, Z \cup X, Y)$ *precisely when* $Ind^b_\Delta(X \cup L, Z, Y)$.

Properties (2) and (3) are known as the semi-graphoid axioms [19]. Property (1) was added recently to them under the name *trivial independence* [26]. Property (2) is known as symmetry. Property (3) is typically broken into three other properties:

- *Decomposition:* $Ind^b_\Delta(X \cup L, Z, Y)$ only if $Ind^b_\Delta(X, Z, Y)$,
- *Weak union:* $Ind^b_\Delta(X \cup L, Z, Y)$ only if $Ind^b_\Delta(L, Z \cup X, Y)$ and
- *Contraction:* $Ind^b_\Delta(X, Z, Y)$ and $Ind^b_\Delta(L, Z \cup X, Y)$ only if $Ind^b_\Delta(X \cup L, Z, Y)$.

The semi-graphoid axioms are important for at least two reasons. First, they are intuitive properties that independence is expected to satisfy [19]. Second, they lead to an important result about the identification of LCI assertions by examining the topology of a structured database, which is the subject of Section 6.

Structured Databases

We now introduce structured databases, which will be discussed in more detail in Section 6.

Figures 1, 2 and 4 depict a number of structured databases. In general, a structured database has two parts: (1) a directed acyclic graph over atomic propositions, and (2) a number of local databases, one for each atom n in the graph. The local database for atom n can only refer to n, its parents, and atoms that do not appear in the directed graph. The local database is typically depicted next to the atom as shown in Figures 1, 2 and 4. If a database is structured, some of the LCI assertions it satisfies can be detected by visual inspection of the database structure, using a method that we shall describe in Section 6. For example, using this method, one can detect the following LCI assertions regarding the databases in Figure 1:[1]

- $Ind^b_\Delta(\{A\}, \emptyset, \{B\})$ in Figure 1.1.
- $Ind^b_\Delta(\{A\}, \{B\}, \{C\})$ in Figure 1.2.
- $Ind^b_\Delta(\{B\}, \{A\}, \{C\})$ in Figure 1.3.

We provide the formal definition of structured databases together with some of their key properties in Section 6.

The next two sections provide four other formulations of LCI, corresponding to its use in testing logical entailment and satisfiability, and in computing abductions and diagnoses.

[1] These assertions are meant to review what will be covered in Section 6. The reader is not expected to understand how these assertions are detected at this stage.

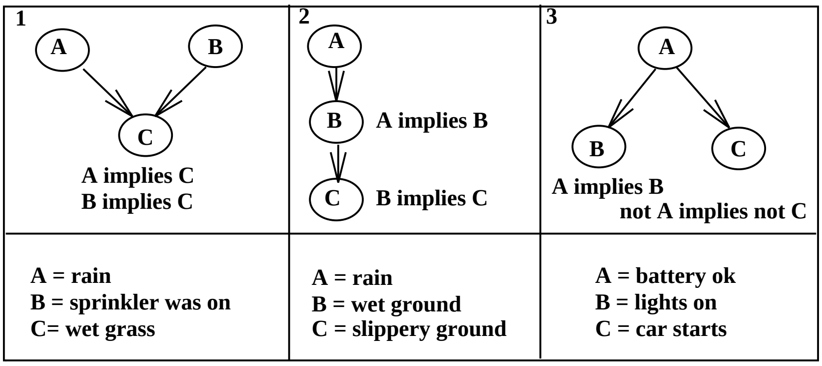

Figure 1: Three structured databases. Each graph identifies LCI assertions that are satisfied by its associated database.

4 Entailment and Satisfiability Formulations

This section discusses the computational role of LCI in deciding logical entailment and satisfiability. For this purpose, we present two formulations of LCI, each oriented towards one of these dual tasks.

We start with the formulation of LCI that is oriented towards testing logical entailment, that is, deciding whether some propositional sentence α is logically entailed by a database Δ. We first present the formulation and then show how it can be applied to logical entailment when the database is structured.

The computational difficulty in testing for logical entailment (at least in the propositional case) stems from the inability to do *decompositional* testing. That is, although one can decompose the test $\Delta \models \alpha \wedge \beta$ into testing whether $\Delta \models \alpha$ *and* $\Delta \models \beta$, one cannot (in general) decompose the test $\Delta \models \alpha \vee \beta$ into testing whether $\Delta \models \alpha$ *or* $\Delta \models \beta$. To see this, note that if $\Delta = \{p \vee q\}$, then $\Delta \models p \vee q$ holds while neither $\Delta \models p$ nor $\Delta \models q$ holds. Therefore, the decomposition

$$\Delta \models p \vee q \text{ precisely when } \Delta \models p \text{ or } \Delta \models q$$

is not valid in this case.

If such a decomposition were valid, however, testing for entailment would be very easy. We would always be able to rewrite the test $\{\alpha_1, \ldots, \alpha_n\} \models \alpha$ into the equivalent test $\neg\alpha \models \neg\alpha_1 \vee \ldots \vee \neg\alpha_n$ and then decompose the latter into n tests $\neg\alpha \models \neg\alpha_1, \ldots, \neg\alpha \models \neg\alpha_n$, which in turn are equivalent to $\alpha_1 \models \alpha, \ldots, \alpha_n \models \alpha$. This would make testing for logical entailment linear in the number n of sentences in a database.

Although the decomposition of a disjunctive test $\Delta \models \alpha \vee \beta$ is not valid in general, it can be valid when certain LCI assertions are satisfied by the database Δ. The second formulation of LCI is meant to explicate these assertions.

Entailment formulation

We first need the following supporting definition.

Definition 4 *A <u>disjunctive clause</u> over a set of atomic propositions X is a disjunction of literals, one literal for each proposition in X. We use $\check{X}$ to denote a disjunctive clause over propositions X.*

For example, there are four disjunctive clauses over p and q: $p \vee q$, $p \vee \neg q$, $\neg p \vee q$ and $\neg p \vee \neg q$.

The following formulation of LCI is in terms of decomposing disjunctive tests. Theorem 2 proves the equivalence between this formulation and the first one based on belief change.

Definition 5 *Let Δ be a propositional database and let X, Y and Z be three disjoint sets of atomic propositions. Database Δ finds X independent of Y given Z, written $Ind^e_\Delta(X, Z, Y)$, iff*

$$\Delta \models \check{X} \vee \check{Y} \vee \check{Z} \textit{ precisely when } \Delta \models \check{X} \vee \check{Z} \textit{ or } \Delta \models \check{Y} \vee \check{Z}$$

for all disjunctive clauses $\check{X}, \check{Y}, \check{Z}$.

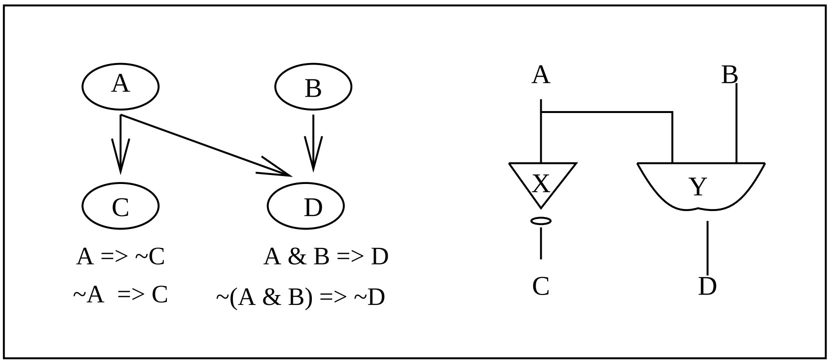

Figure 2: A structured database representing a digital circuit.

Theorem 2 *$Ind^b_\Delta(X, Z, Y)$ precisely when $Ind^e_\Delta(X, Z, Y)$.*

Theorem 2 is then showing the equivalence between

1. the validity of decomposing certain disjunctive tests, which is a computation-oriented formulation of LCI (Definition 5) and
2. the irrelevance of certain information to certain beliefs, which is an intuition-oriented formulation of LCI (Definition 3).

An important special case of Definition 5 is when $Z = \emptyset$. Here, *false* is the only disjunctive clause over Z, and we have $Ind^e_\Delta(X, \emptyset, Y)$ iff

$$\Delta \models \check{X} \vee \check{Y} \text{ precisely when } \Delta \models \check{X} \text{ or } \Delta \models \check{Y}$$

for all disjunctive clauses $\check{X}$ and $\check{Y}$.

Example

Consider the database $\Delta = \{p \supset r,\ r \supset s\}$. This database satisfies the LCI assertion $Ind^b_\Delta(\{s\}, \{r\}, \{p\})$. Therefore, the following decompositions are valid:

$$\begin{array}{rcl}
\Delta \models p \vee r \vee s & \text{precisely when} & \Delta \models p \vee r \text{ or } \Delta \models r \vee s \\
\Delta \models \neg p \vee r \vee s & \text{precisely when} & \Delta \models \neg p \vee r \text{ or } \Delta \models r \vee s \\
\Delta \models p \vee \neg r \vee s & \text{precisely when} & \Delta \models p \vee \neg r \text{ or } \Delta \models \neg r \vee s \\
& \vdots & \\
\Delta \models \neg p \vee \neg r \vee \neg s & \text{precisely when} & \Delta \models \neg p \vee \neg r \text{ or } \Delta \models \neg r \vee \neg s.
\end{array}$$

Note, however, that Δ does not satisfy the assertion $Ind^b_\Delta(\{s\}, \emptyset, \{p\})$. Therefore, it should not be surprising that the decomposition

$$\Delta \models \neg p \vee s \text{ precisely when } \Delta \models \neg p \text{ or } \Delta \models s$$

does not hold ($\Delta \models \neg p \vee s$ holds but neither $\Delta \models \neg p$ nor $\Delta \models s$ does).

The computational value of independence

We have shown so far that:

1. disjunctive tests of the form $\Delta \models \alpha \vee \beta$ cannot be decomposed in general;
2. if disjunctive tests were always decomposable, then testing for logical entailment would become linear in the number of sentences in a database;

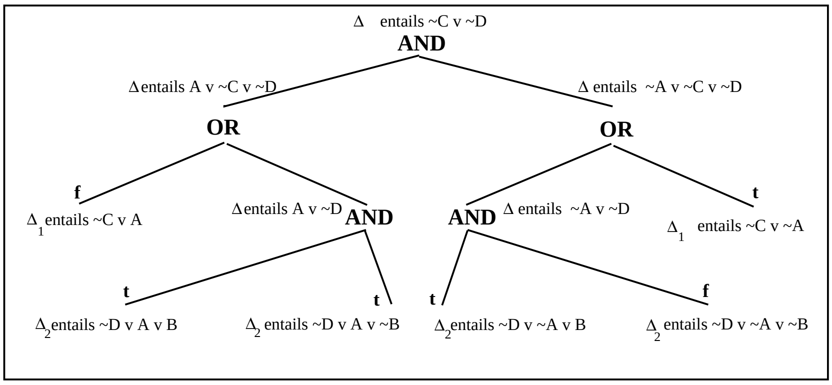

Figure 3: A rewrite process for decomposing a global test for entailment into a number of local tests. Also shown is the evaluation of local tests.

3. disjunctive tests could be decomposed when certain LCI assertions are satisfied by the database.

When all disjunctive tests are decomposable, taking advantage of this decomposability is easy as we have shown. But when only certain tests are decomposable — that is, limited independence is available — then the issue is no longer that simple.

The algorithm we present in Section 7 is a good example of how limited independence information can be utilized computationally. It decomposes a global test $\Delta \models \alpha$ into a number of local tests $\Delta_i \models \alpha_i$, which are assumed to require constant time. The decomposition is accomplished using two rewrite rules:

1. *case-analysis:* $\Delta \models \alpha$ is rewritten into $\Delta \models \alpha \vee \beta$ and $\Delta \models \alpha \vee \neg\beta$ and
2. *decomposition:* $\Delta \models \alpha \vee \beta \vee \gamma$ is rewritten into $\Delta \models \alpha \vee \gamma$ or $\Delta \models \beta \vee \gamma$.

The case-analysis rule is always valid because

1. α is equivalent to $(\alpha \vee \beta) \wedge (\alpha \vee \neg\beta)$ and
2. $\Delta \models (\alpha \vee \beta) \wedge (\alpha \vee \neg\beta)$ iff $\Delta \models \alpha \vee \beta$ and $\Delta \models \alpha \vee \neg\beta$.

However, the decomposition rule is valid only when certain LCI assertions hold according to Definition 5.

Figure 3 shows an example of this rewrite process as applied to the test $\Delta \models \neg C \vee \neg D$ with respect to the database in Figure 2. Specifically, the figure depicts a trace of a rewrite process that decomposes the global test $\Delta \models \neg C \vee \neg D$ into six local tests: $\Delta_1 \models \neg C \vee A$, $\Delta_1 \models \neg C \vee \neg A$, $\Delta_2 \models \neg D \vee A \vee B$, $\Delta_2 \models \neg D \vee A \vee \neg B$, $\Delta_2 \models \neg D \vee \neg A \vee B$, and $\Delta_2 \models \neg D \vee \neg A \vee \neg B$. Here Δ_1 and Δ_2 are the local databases:

$$\Delta_1 = \{A \supset \neg C,\ \neg A \supset C\}$$

and

$$\Delta_2 = \{A \wedge B \supset D,\ \neg(A \wedge B) \supset \neg D\}.$$

The process proceeds top-down, rewriting each test into a logical combination of the tests below it. The case-analysis rule rewrites a test into a conjunction of other tests. The decomposition rule rewrites a test into a disjunction of other tests. Therefore, the rewrite process constructs an and-or tree, the root of which is the global test, and the leaves of which are local tests. Note here that the two applications of the decomposition rule are validated by the LCI assertion $Ind^e_\Delta(\{C\}, \{A\}, \{D\})$, which follows from the structure in Figure 2.

Figure 3 also depicts the result of evaluating local tests. It is clear that this tree evaluates to *true*, which is the answer computed for the global test $\Delta \models \neg C \vee \neg D$.

We have a few points to make with respect to this example. First, it exemplifies the computational paradigm that we are advocating in this paper: decompose a global computation into a number of local, constant time computations. Second, for each of the computational tasks we shall consider, there will be a corresponding case-analysis and decomposition rules. The case-analysis rule will always be valid, but the decomposition rule will be valid only when certain LCI assertions hold. Therefore, to develop an independence-based algorithm, one must

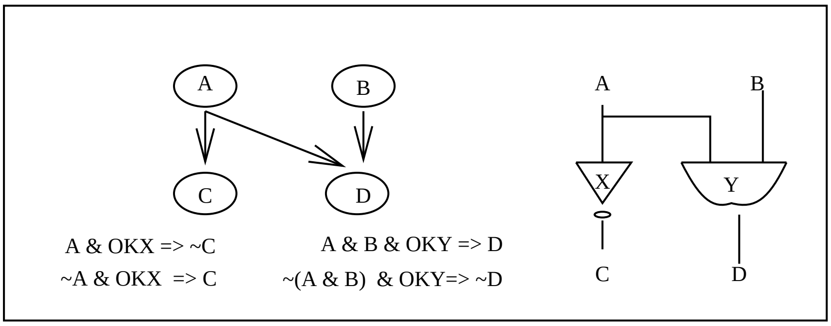

Figure 4: A structured database representing a digital circuit.

1. know when the decomposition rule is valid *(soundness)*
2. control the rewrite process so that:
   (a) it terminates *(completeness)*
   (b) it invokes the least number of rewrites possible *(efficiency)*

Structured databases that we shall introduce in Section 6 are central for achieving these goals: the topology of a database tells us which decompositions are valid and can be used to control the rewrite process so that it terminates.

The number of invoked rewrites decides the complexity of reasoning and is the parameter for measuring the superiority of one algorithm versus another. As we shall see, for certain database structures, we can decompose a task using only a linear number of rewrites. These computational issues will be discussed in more detail later.

Satisfiability formulation

We now provide the third formulation of LCI, which is oriented towards testing satisfiability:

Definition 6 *Let Δ be a propositional database and let X, Y and Z be three disjoint sets of atomic propositions. Database Δ finds X independent of Y given Z, written $Ind^s_\Delta(X,Z,Y)$, iff $\Delta \cup \{\hat{Z},\hat{X}\}$ and $\Delta \cup \{\hat{Z},\hat{Y}\}$ are both satisfiable precisely when $\Delta \cup \{\hat{Z},\hat{Y},\hat{X}\}$ is satisfiable for all conjunctive clauses $\hat{Z}$, $\hat{Y}$ and $\hat{X}$.*

This formulation is dual to the second formulation for testing entailment. The equivalence is established below:

Theorem 3 *$Ind^s_\Delta(X,Z,Y)$ precisely when $Ind^e_\Delta(X,Z,Y)$.*

Now that we have established the equivalence between Ind^b_Δ, Ind^e_Δ and Ind^s_Δ, we will sometimes drop the superscripts and simply write Ind_Δ. The specific interpretation of Ind_Δ will then be chosen depending on the context.

5 Diagnosis and Abduction Formulations

This section presents two more formulations of LCI oriented towards the dual tasks of diagnosis and abduction. We define the notion of a *consequence* for diagnosis tasks, which is a semantical characterization of all diagnoses. We also define the notion of an *argument* for abduction tasks, which is a semantical characterization of all abductions. We show that the difficulties in computing diagnoses and abduction are rooted in the inability to decompose the computations of consequences and arguments in general. We also show that independence assertions can validate this decomposition in certain cases. We present two more formulations of LCI to spell out these cases. The two formulations are dual since consequences are dual to arguments.

We start with the fourth formulation of LCI that is oriented towards diagnostic applications. Before we present it though, we need to review some basics of diagnosis [7].

Basics of diagnosis

In the diagnostic literature [7], a *system* is typically characterized by a tuple (Δ, P, W) where Δ is a database constructed from atomic propositions $P \cup W$. Moreover, a system *observation* is typically characterized by a sentence γ constructed from atoms P. The atoms in W are called assumables and those in P are called non-assumables. The intention is that the database Δ describes the system behavior, the assumables W represent the modes of system components and the sentence γ represents the observed system behavior. For example, in Figure 4, the assumables are okX and okY, the non-assumables are A, B, C and D and a possible system observation γ is $C \wedge D$.

A *diagnosis* is defined as a conjunctive clause over the set of assumables W that is consistent with $\Delta \cup \{\gamma\}$. Therefore, a diagnosis is an assignment of modes to components that is consistent with the system description and its observed behavior. In Figure 4, okX and okY are the assumables and $okX \wedge \neg okY$ is a potential diagnosis.

One goal of diagnostic reasoning is to characterize all diagnoses compactly so they can be presented to a user. Another goal is to extract a subset of these diagnoses according to some preference criterion. We have shown elsewhere that these objectives can be achieved in two steps [4]. First, we compute the *consequence* of observation γ, which is a Boolean expression that characterizes all the diagnoses of γ. Second, we extract the most preferred diagnoses from the computed consequence.

The consequence of an observation is defined formally below:

Definition 7 *Let Δ be a propositional database, α be a propositional sentence and W be a set of atomic propositions that do not appear in α. The* <u>*consequence*</u> *of sentence α with respect to database Δ and atoms W, written $Cons_W^{\Delta}(\alpha)$, is a logically strongest sentence $\tilde{W}$ such that $\Delta \cup \{\alpha\} \models \tilde{W}$.*[2]

When Δ and W are clear from the context, we will write $Cons(\alpha)$ instead of $Cons_W^{\Delta}(\alpha)$ for simplicity. In Figure 4, for example, the consequence of observation $C \wedge D$ is $\neg okX \vee \neg okY$ because it is a logically strongest sentence (constructed from assumables) that can be concluded from the given observation and system description.

A consequence characterizes all diagnoses in the following way:

Theorem 4 *$\hat{W}$ is a diagnosis for system (Δ, P, W) and observation γ iff $\hat{W} \models Cons_W^{\Delta}(\gamma)$.*

The consequence $\neg okX \vee \neg okY$ for example characterizes three diagnoses: $\neg okX \wedge \neg okY$, $okX \wedge \neg okY$ and $\neg okX \wedge okY$.

Diagnosis formulation

Similar to testing for logical entailment, the difficulty with computing diagnoses is that it cannot be done compositionally. In particular, although $Cons(\alpha \vee \beta)$ is equivalent to $Cons(\alpha) \vee Cons(\beta)$, $Cons(\alpha \wedge \beta)$ is not equivalent to $Cons(\alpha) \wedge Cons(\beta)$ in general.

The following formulation of LCI is in terms of decomposing consequences. It is followed by a theorem that shows the equivalence between this formulation and the one based on belief change, therefore, identifying conditions under which decomposing consequences is valid.

Definition 8 *Let Δ be a database and let X, Y, Z and W be disjoint sets of atomic propositions. The pair (Δ, W) finds X independent of Y given Z, written $Ind^c_{(\Delta,W)}(X, Z, Y)$, iff*

$$\models Cons(\hat{X} \wedge \hat{Y} \wedge \hat{Z}) \equiv Cons(\hat{X} \wedge \hat{Z}) \wedge Cons(\hat{Y} \wedge \hat{Z})$$

for all conjunctive clauses $\hat{X}, \hat{Y}$ and $\hat{Z}$.

Theorem 5 *$Ind^c_{(\Delta,W)}(X, Z, Y)$ precisely when $Ind^b_{\Delta}(X, Z \cup W, Y)$.*

Given Theorem 5, we now have an equivalence between

1. the irrelevance of information to beliefs (Definition 3),

[2] A sentence α is stronger than sentence β iff $\alpha \models \beta$. The sentence β is said to be weaker than α in such a case.

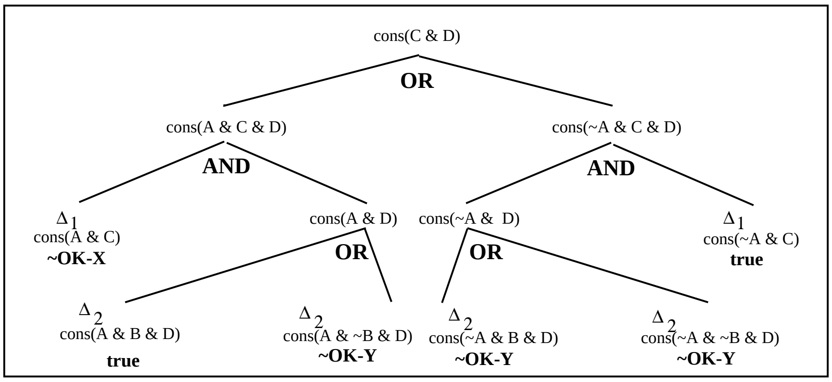

Figure 5: A rewrite process for decomposing a global consequence into a number of local consequences. Also shown are the values of local consequences.

2. the validity of decomposing entailment tests (Definition 5),

3. the validity of decomposing satisfiability tests (Definition 6), and

4. the validity of decomposing consequences (Definition 8).

Before we present an example on the computational use of this LCI formulation, we note the following special case of Definition 8. When Z is empty, *true* is the only conjunctive clause over Z, and we have $Ind^c_{(\Delta,W)}(X,\emptyset,Y)$ iff

$$\models Cons(\hat{X} \wedge \hat{Y}) \equiv Cons(\hat{X}) \wedge Cons(\hat{Y})$$

for all conjunctive clauses $\hat{X}$ and $\hat{Y}$.

Example

We have presented elsewhere an algorithm for computing consequences with respect to a structured database [4]. The algorithm works by rewriting a global consequence of the form $Cons^{\Delta}(\alpha)$ into a Boolean expression that involves logical connectives and local consequences of the form $Cons^{\Delta_i}(\alpha_i)$, where Δ_i is a local database. The algorithm is based on the following two rewrites:

1. *case-analysis:* $Cons(\alpha)$ is rewritten into $Cons(\alpha \wedge \beta) \vee Cons(\alpha \wedge \neg\beta)$;[3] and

2. *decomposition:* $Cons(\alpha \wedge \beta \wedge \gamma)$ is rewritten into $Cons(\alpha \wedge \gamma) \wedge Cons(\beta \wedge \gamma)$ when the corresponding LCI assertion holds.

We will now consider an example using these rewrites with respect to Figure 4. Our goal is to compute the consequence $Cons(C \wedge D)$, where the set W contains the atoms okX and okY.

Figure 5 depicts a trace of the rewrite process that decomposes the global consequence $Cons(C \wedge D)$ into six local consequences: $Cons^{\Delta_1}(A \wedge C)$, $Cons^{\Delta_1}(\neg A \wedge C)$, $Cons^{\Delta_2}(D \wedge A \wedge B)$, $Cons^{\Delta_2}(D \wedge A \wedge \neg B)$, $Cons^{\Delta_2}(D \wedge \neg A \wedge B)$ and $Cons^{\Delta_2}(D \wedge \neg A \wedge \neg B)$. Here, Δ_1 and Δ_2 are the local databases:

$$\Delta_1 = \{A \wedge okX \supset \neg C,\ \neg A \wedge okX \supset C\}$$

and

$$\Delta_2 = \{A \wedge B \wedge okY \supset D,\ \neg(A \wedge B) \wedge okY \supset \neg D\}.$$

[3] The case-analysis rule follows because:

(a) α is equivalent to $(\alpha \wedge \beta) \vee (\alpha \wedge \neg\beta)$ and

(b) $Cons((\alpha \wedge \beta) \vee (\alpha \wedge \neg\beta))$ is equivalent to $Cons(\alpha \wedge \beta) \vee Cons(\alpha \wedge \neg\beta)$.

The process proceeds top-down, rewriting each consequence into the nodes below it. The case-analysis rule rewrites a consequence into a disjunction of other consequences, while the decomposition rule rewrites a consequence into a conjunction of other consequences. Therefore, the rewrite process constructs an and-or tree, the root of which is the global consequence, and the leaves of which are local consequences. Given the values of local consequences, the tree simplifies to $\neg okX \vee \neg okY$, which is the value of the global consequence $Cons(C \wedge D)$. Note that local consequences can be computed by operating on local databases, which is assumed to require constant time. Note also that the two applications of the decomposition rule are validated by the LCI assertion $Ind^c_{(\Delta,W)}(\{C\},\{A\},\{D\})$, which can be inferred by examining the database structure.

Abduction formulation

The fifth formulation of LCI is oriented towards the computation of *abductions.*

There is no standard definition of an abduction, although existing definitions seem to agree on the following two properties:

1. *Generality:* An abduction should be as general as possible while still explaining the finding.
2. *Scope:* The language in which an abduction is phrased must be restricted in order to avoid trivial abductions (α being an abduction of itself).

The following definition gives a generic notion of an abduction, called an *argument.* The ATMS label of a proposition, for example, is only a syntactic variation on the argument for that proposition [6, 21].

Definition 9 *Let Δ be a propositional database, α be a propositional sentence, and W be a set of atomic propositions that do not appear in α. The argument for sentence α with respect to database Δ and atoms W, written $Arg^{\Delta}_{W}(\alpha)$, is the logically weakest sentence $\tilde{W}$ such that $\Delta \cup \{\tilde{W}\} \models \alpha$.*[4]

The intention here is that Δ represents background information, α represents a finding, and W restricts the language used in phrasing an abduction. For example, if we choose W to contain okX and okY in Figure 4, the argument for finding $\neg C \vee \neg D$ would then be $okX \wedge okY$.

Arguments are dual to consequences as the following theorem shows:

Theorem 6 $Arg^{\Delta}_{W}(\alpha) \equiv \neg Cons^{\Delta}_{W}(\neg\alpha)$.

This also explains, indirectly, why ATMS engines are typically used for computing diagnoses.

Similar to logical entailment, the difficulty in computing abductions/arguments is related to the inability to decompose arguments. That is, although $Arg(\alpha \wedge \beta)$ is equivalent to $Arg(\alpha) \wedge Arg(\beta)$, $Arg(\alpha \vee \beta)$ is not equivalent to $Arg(\alpha) \vee Arg(\beta)$ in general. For example, if $\Delta = \{p \supset (r \vee s)\}$ and $W = \{p\}$, then $Arg(r \vee s) = p$ while $Arg(r) = false$ and $Arg(s) = false$.

The following formulation of LCI is in terms of decomposing arguments. It is followed by a theorem that shows the equivalence between this formulation and previous ones, thus establishing conditions under which it is valid to decompose arguments.

Definition 10 *Let Δ be a database and let X, Y, Z and W be disjoint sets of atomic propositions. The pair (Δ, W) finds X independent of Y given Z, written $Ind^a_{(\Delta,W)}(X,Z,Y)$, iff*

$$\models Arg(\check{X} \vee \check{Y} \vee \check{Z}) \equiv Arg(\check{X} \vee \check{Z}) \vee Arg(\check{Y} \vee \check{Z}),$$

for all disjunctive clauses $\check{X}, \check{Y}, \check{Z}$.

Theorem 7 $Ind^a_{(\Delta,W)}(X,Z,Y)$ *precisely when* $Ind^c_{(\Delta,W)}(X,Z,Y)$.

[4] We will typically drop Δ and W, thus writing $Arg(\alpha)$, when no confusion is expected.

Now that we have established the equivalence between $Ind^a_{(\Delta,W)}$ and $Ind^c_{(\Delta,W)}$, we will sometimes drop the superscripts and simply write $Ind_{(\Delta,W)}$. The specific interpretation of Ind_Δ will then be chosen depending on the context.

We close this section with an important special case of Definition 10. When Z is empty, *false* is the only disjunctive clause over Z, and we have $Ind^a_{(\Delta,W)}(X, \emptyset, Y)$ iff

$$Arg(\check{X} \vee \check{Y}) \equiv Arg(\check{X}) \vee Arg(\check{Y}),$$

for all disjunctive clauses $\check{X}$ and $\check{Y}$.

6 Structured Databases

The different formulations of LCI show how independence information can validate the decomposition of a computation into smaller computations that can be performed in parallel, a decomposition that is not sound in general. Although this ability to decompose a computation could be valuable from a complexity viewpoint, exploiting such decompositions is not always straightforward.

Structured databases make this utilization of independence more feasible. In particular, the structure of a database plays two important roles in designing independence-based algorithms:

1. it graphically explicates valid applications of the decomposition rule, and

2. it specifies a control flow that guarantees the termination of the decomposition process.

This is also the role that directed acyclic graphs have been playing in probabilistic reasoning, and our goal here is to extend the role to propositional logic.[5]

We will now provide the formal definition of structured databases, and then present two important operations on their structures: (1) reading independences off the database structure and (2) pruning irrelevant parts of a database structure before performing certain computations.

6.1 The Syntax of a Structured Database

A structured database is a pair $(\mathcal{G}, \Delta)$ where

- $\mathcal{G}$ is a *directed acyclic graph* with nodes representing atomic propositions, and

- Δ is the union of *local databases,* one database for each node in $\mathcal{G}$.

The atoms that appear in the database Δ but do not appear in the structure $\mathcal{G}$ are called *exogenous* atoms. For example, in Figure 4, the exogenous atoms are okX and okY. The local database corresponding to proposition n in $\mathcal{G}$, denoted by Δ_n, must satisfy the following conditions:

1. *locality:* the only atoms in $\mathcal{G}$ that Δ_n can mention are the family of atom n;[6] and

2. *modularity:* if Δ_n entails a disjunctive clause that does not mention atom n, then the clause must be valid.

The first condition makes sure that each local database is restricted to specifying the relationship between an atom and its parents. The second condition ensures that a local database for atom n be only concerned with specifying how the parents of n determine the truth value of n. That is, the database should not be concerned with specifying a direct relationship between the parents of n. Note also that the modularity condition ensures the consistency of a local database since the empty clause (falsehood) does not mention atom n.

If the arcs of a structured database represent causal influences, then the conditions above are typically self–imposed. For example, suppose that a structured database is used to describe the functionality of a digital circuit as shown in Figures 2 and 4. Each atom in the structure represents the state of a wire in the

[5] However, we later discuss a key difference between the propositional and probabilistic framework.

[6] Recall that the family of n consists of n and its parents in the database structure.

circuit, and the arcs point from the inputs of a gate into its output. The local database associated with atom n specifies the behavior of the gate having output n. For this class of structured databases, the locality and modularity conditions are typically self–imposed since

1. the locality condition means that one should not mention any wire that is not an input or an output of the gate for which we are specifying a behavior, and
2. the modularity condition means that one should not specify a relationship between the inputs of a gate in the process of specifying its behavior.

We have focused elsewhere on a causal interpretation of structured databases, referred to as *symbolic causal networks*, where we discussed non-computational applications of LCI that include reasoning about actions and identifying anomalous extensions of nonmonotonic theories [5].

6.2 Structure-Based Independence

By construction, a structured database satisfies some LCI assertions that can be easily detected from the database structure:

Theorem 8 *Let $(\mathcal{G}, \Delta)$ be a structured database, W be its exogenous atoms, n be a node in $\mathcal{G}$, U be its parents, and N be its non-descendants. Then $Ind_{\Delta}(\{n\}, U \cup W, N)$.*

That is, each atom in the database structure is independent of its non-descendants given its parents and exogenous atoms. With respect to the database in Figure 4, this theorem states that

1. $\{D\}$ is independent of $\{C\}$ given $\{A, B, okX, okY\}$,
2. $\{B\}$ is independent of $\{A\}$ given $\{okX, okY\}$, and
3. $\{C\}$ is independent of $\{D\}$ given $\{A, okX, okY\}$.

Theorem 8 brings up a key difference between structured databases and Bayesian networks: the independences characterized by this theorem are *properties* of a structured database, but their corresponding probabilistic independences are *part of the definition* of a Bayesian network. That is, the restrictions on a structured database are sufficient to imply these independences, but similar restrictions on a Bayesian network are not enough to imply the corresponding probabilistic independences. There are two implications of this:

- By constructing a structured database, one is making no explicit commitment to independences; one is committing only to the logical content of local databases. By constructing a Bayesian network, however, one is explicitly committing to independences which correspond to Theorem 8 [19].
- One can recover the independences satisfied by a structured database from only its local databases, that is, without having to consider the database structure.[7]

Note that having an explicit structure is very useful in deducing further independences. Specifically, since LCI is a symmetric relation, the LCI assertion in (3) above implies

4. $\{D\}$ is independent of $\{C\}$ given $\{A, okX, okY\}$,

which *cannot* be detected by applying Theorem 8 directly. As this example illustrates, a structured database satisfies more LCI assertions than those characterized by Theorem 8. We now present a topological test, called d-separation, for identifying some of these additional assertions.

d-separation is a test on directed acyclic graphs that tells us whether two sets of nodes are d-separated by a third set [19]. According to this test, for example, the nodes $\{B, D\}$ are d-separated from the node $\{C\}$ by the node $\{A\}$ in Figure 4. Therefore, d-separation could be viewed as a relation $Sep_{\mathcal{G}}$ where $Sep_{\mathcal{G}}(X, Z, Y)$ holds precisely when Z d-separates X from Y in the directed acyclic graph $\mathcal{G}$. The d-separation test has

[7] We do not investigate this direction in this paper however.

been instrumental in reasoning about independence in Bayesian networks. We shall use it analogously in structured databases.

We will define d-separation later, but we first go over how it can be used to identify LCI assertions. The following result, which is basically a corollary of a theorem in [25], reveals the importance of d-separation:

Theorem 9 *Let $(\mathcal{G}, \Delta)$ be a structured database and let W be its exogenous atoms. If Z d-separates X from Y in $\mathcal{G}$, then Δ finds X and Y independent given $Z \cup W$. That is, $Sep_{\mathcal{G}}(X, Z, Y)$ implies $Ind_{\Delta}(X, Z \cup W, Y)$.*

Therefore, d-separation allows one to infer LCI assertions satisfied by a given structured database by simply examining the topology of its structure.

Following are the key computational implications of Theorem 9. If atoms X and Y are d-separated by Z in the structure of database $(\mathcal{G}, \Delta)$, then the following decompositions are valid with respect to database Δ:

1. $\Delta \models \check{X} \vee \check{Z} \vee \check{W} \vee \check{Y}$ iff $\Delta \models \check{X} \vee \check{Z} \vee \check{W}$ or $\Delta \models \check{Z} \vee \check{W} \vee \check{Y}$,
2. $\Delta \cup \{\hat{X} \wedge \hat{Z} \wedge \hat{W} \wedge \hat{Y}\}$ is satisfiable iff $\Delta \cup \{\hat{X} \wedge \hat{Z} \wedge \hat{W}\}$ and $\Delta \cup \{\hat{Z} \wedge \hat{W} \wedge \hat{Y}\}$ are satisfiable,
3. $Arg(\check{X} \vee \check{Z} \vee \check{Y})$ is equivalent to $Arg(\check{X} \vee \check{Z}) \vee Arg(\check{Z} \vee \check{Y})$, and
4. $Cons(\hat{X} \wedge \hat{Z} \wedge \hat{Y})$ is equivalent to $Cons(\hat{X} \wedge \hat{Z}) \vee Cons(\hat{Z} \wedge \hat{Y})$,

for all disjunctive clauses $\check{X}, \check{Y}, \check{Z}$ and all conjunctive clauses $\hat{X}, \hat{Y}, \hat{Z}$. These implications show why the complexity of reasoning in the framework we are proposing is decided by the topology of a structured database.

The d-separation test will be defined next. We first need the following supporting definition:

Definition 11 *Let $\mathcal{G}$ be a directed acyclic graph and let X, Y, and Z be three disjoint sets of nodes in $\mathcal{G}$. A path between X and Y is <u>Z-active</u> precisely when its nodes satisfy the following conditions:*

1. *each converging node belongs to Z or has a descendent in Z, and*
2. *each diverging or linear node is outside Z.*

When Z is empty, we say the path is <u>active.</u> Figure 6 gives the definition of converging, diverging, and linear nodes.

In Figure 4, for example,

- the path $A \to D \leftarrow B$ is $\{D\}$-active,
- the path $C \leftarrow A \to D$ is active, and
- the path $A \to D \leftarrow B$ is not active.

Following is the definition of d-separation:

Definition 12 ([19]) *In a directed acyclic graph $\mathcal{G}$, nodes Z d-separate X from Y, written $Sep_{\mathcal{G}}(X, Z, Y)$, precisely when there is no Z-active path between X and Y in $\mathcal{G}$. When Z is empty, we say that X and Y are d-separated.*

The d-separation test is not complete in the sense that it cannot identify all LCI assertions satisfied by a database. To illustrate this, consider the structured databases in Figure 7. In the first one, $\{A\}$ and $\{C\}$ are independent although they are not d-separated. And in the second, $\{B\}$ and $\{C\}$ are also independent although not d-separated. Intuitively, in the first database, there is no information about A that could lead us to prove anything about C. Similarly in the second database, neither B nor $\neg B$ will help in proving anything about C.

Existing structure-based algorithms use only the database structure for identifying LCI assertions. It should be clear then that such algorithms cannot be optimal because they are bound to miss some independences that could be useful computationally. This also seems to be the practice in the probabilistic literature, except possibly for some recent work on utilizing non-structural independence [1, 18].

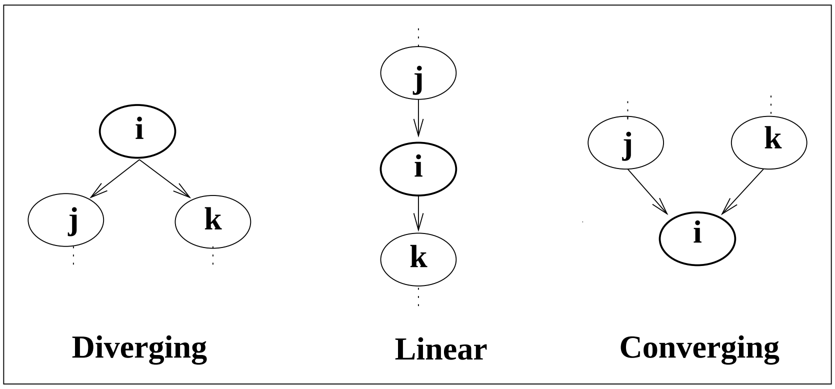

Figure 6: There are three types of intermediate nodes on a given path. The type of a node is determined by its relation to its neighbors. A node is diverging if both neighbors are children. A node is linear if one neighbor is a parent and the other is a child. A node is converging if both neighbors are parents.

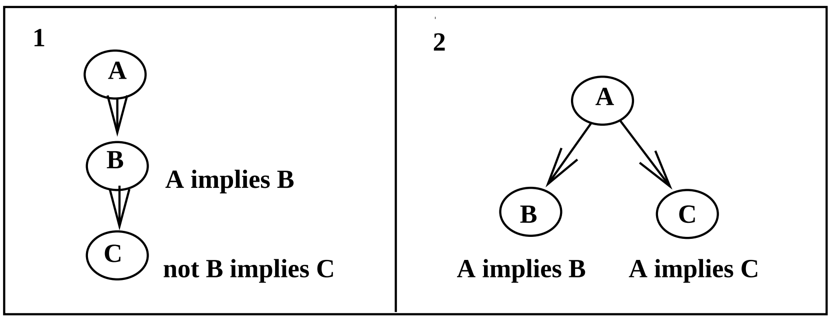

Figure 7: Structured databases satisfying more independences than is revealed by their graphical structures.

6.3 Structure-Based Pruning

The modularity condition on local databases has strong implications that make structured databases attractive from computational complexity and knowledge acquisition viewpoints. To illustrate this, suppose that we are constructing a structured database incrementally by adding leaf nodes together with their local databases. When adding node n, the modularity condition ensures that the added database Δ_n does not contradict the structured database Δ constructed so far. For the database Δ_n to contradict Δ, it must entail some clause that is inconsistent with Δ. But this is impossible. Every non-valid clause that is entailed by Δ_n must mention the atom n. And any clause that mentions n cannot be inconsistent with Δ, since Δ does not mention n to start with.

Therefore, structured databases are guaranteed to be consistent by construction, which is a very attractive property from a knowledge acquisition viewpoint.[8]

Another important implication of the modularity condition on local databases is the ability to prune certain (irrelevant) parts of a structured database before computing answers to certain queries. The following theorem identifies cases where this pruning is possible.

Theorem 10 *Let $(\mathcal{G}, \Delta)$ be a structured database with exogenous variables W, let N be some atoms in structure $\mathcal{G}$ and n be a leaf node in $\mathcal{G}$ that does not belong to N. If $(\mathcal{G}', \Delta')$ is a structured database (with exogenous variables W') that results from removing node n and its local database Δ_n from $(\mathcal{G}, \Delta)$, then*

1. *Entailment:* $\Delta \models \check{N}$ *iff* $\Delta' \models \check{N}$,
2. *Satisfiability:* $\Delta \cup \{\hat{N}\}$ *is satisfiable iff* $\Delta' \cup \{\hat{N}\}$ *is satisfiable,*

[8] Modularity is also a key property responsible for the desirable properties of directed constraint networks [10].

3. *Abduction:* $Arg^{\Delta}_{W}(\check{N})$ *is equivalent to* $Arg^{\Delta'}_{W'}(\check{N})$ *and*

4. *Diagnosis:* $Cons^{\Delta}_{W}(\hat{N})$ *is equivalent to* $Cons^{\Delta'}_{W'}(\hat{N})$

for all disjunctive clauses $\check{N}$ *and conjunctive clauses* $\hat{N}$.

For example, when attempting to test for the logical entailment of a clause $\check{N}$ that does not involve a leaf atom n, we can drop atom n and its local database Δ_n without affecting the result of the test. Applying this theorem recursively may prune a significant part of a structured database in certain cases. The amount of pruning, however, depends on how the atoms of clause $\check{N}$ are spread topologically in the structure. For example, at one extreme, clause $\check{N}$ may refer to all leaf nodes in the database structure. In this case, no pruning is possible. At another extreme, the clause may only mention root nodes in the structure, in which case all nodes except for those mentioned in the clause will be pruned.

From a computational complexity viewpoint, the significance of Theorem 10 is in showing how the complexity of a reasoning task is affected by the specific query.

Theorem 10 presents a result which is close in spirit to the traditional usage of irrelevance information in the literature on logical reasoning [15, 24, 23]. That is, before applying some algorithm for testing entailment, we prune some parts of the given database knowing that we will not compromise the soundness of the test. We have to mention, however, that although such pruning can lead to great computational savings in practice, it is only a secondary usage of independence information in the framework that we are proposing. This is also consistent with the way independence is used in probabilistic and constraint-based reasoning. The key usage of independence information is in decomposing a computation with respect to a global database into a number of independent computations with respect to local databases. This usage will be illustrated in the following section, which presents an independence-based algorithm for a class of structured databases.

We close this section with the following corollary which says that we can prune a leaf node from a structured database, without affecting the independences that hold between any of the remaining nodes.

Corollary 1 *Let* $(\mathcal{G}, \Delta)$ *be a structured database with exogenous variables* W. *Let* X, Y *and* Z *be atoms in structure* $\mathcal{G}$ *and let* $(\mathcal{G}', \Delta')$ *be a structured database (with exogenous variables* W'*) that results from removing a leaf node from* $(\mathcal{G}, \Delta)$ *that does not belong to* $X \cup Y \cup Z$. *Then*

$$Ind_{\Delta}(X, Z \cup W, Y) \text{ precisely when } Ind_{\Delta'}(X, Z \cup W', Y).$$

7 Structure-Based Reasoning

This section presents an algorithm for computing arguments. The algorithm applies only to structured databases that are singly-connected, which are introduced in the following section. The algorithm can be generalized straightforwardly to arbitrary structured databases. The computational complexity of the algorithm and its extensions depend on the topology of the structured database to which it applies. This is why this class of algorithms is referred to as *structure-based.*

Using this algorithm, one can also decide entailment and satisfiability in addition to computing consequences. We provide more details on this later.

Singly-connected structures

A singly-connected structure is one in which there is only one undirected path between any two nodes — the structure has no undirected cycles. The database in Figure 4 has a singly-connected structure, but the one in Figure 8 has a multiply-connected structure.

Singly-connected databases are important for at least two reasons:

1. Computation on a singly-connected structure is linear in the number of nodes and arcs of that structure, exponential only in the size of a family.[9]

[9] In the literature on structure-based reasoning, the size of a family is typically assumed to be small enough to justify the assumption that any computation involving only a family will require constant time.

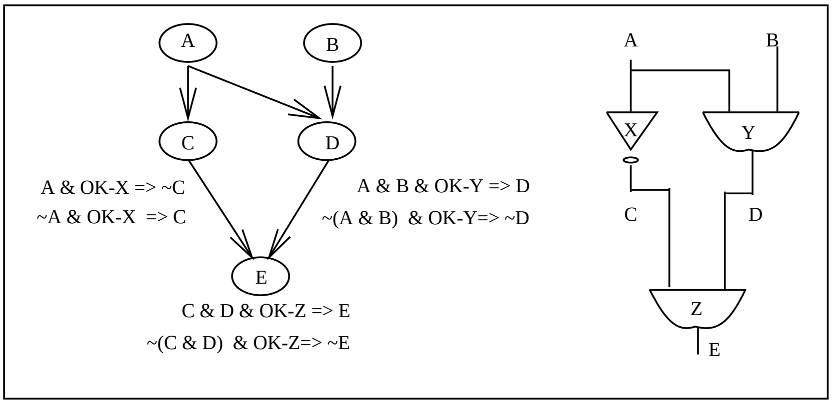

Figure 8: A structured database representing a digital circuit. The structure of this database is multiply-connected because there is more than one undirected path between nodes A and E.

2. There is a simple algorithm for reducing a computation on a multiply-connected structure into a number n of computations on singly-connected structures, where n is exponential in the size of a structure cutset [19, 2].[10]

A singly-connected database is not necessarily Horn. For example, the databases in Figures 2 and 4 are singly-connected and yet contain non-Horn clauses.

The rewrite paradigm

Given a singly-connected database $(\mathcal{G}, \Delta)$, and a disjunctive clause $\check{O}$ over some atoms in $\mathcal{G}$, the algorithm we shall present computes the argument for $\check{O}$ with respect to database Δ. The algorithm is symmetric to a well-known one in the probabilistic literature, known as the polytree algorithm [19], and has similar computational complexity: linear in the number of nodes and arcs in $\mathcal{G}$ and exponential in the size of each family in $\mathcal{G}$. The algorithm can also be used to compute the argument of an arbitrary sentence α by first converting the sentence into conjunctive normal form $\check{O}_1 \wedge \ldots \wedge \check{O}_n$, and then applying the algorithm to each of the conjuncts:

$$Arg(\alpha) \equiv Arg(\check{O}_1 \wedge \ldots \wedge \check{O}_n) \equiv \bigwedge_i Arg(\check{O}_i).$$

The algorithm works by rewriting a global argument $Arg^{\Delta}(\check{O})$ into an expression that includes logical connectives and local arguments of the form $Arg^{\Delta_i}(\alpha_i)$, where Δ_i is a local database. A local argument can be evaluated by operating on a local database, which is assumed to require constant time operation. The algorithm alternates between applying the case-analysis and decomposition rewrite rules:

1. *case-analysis:* $Arg(\alpha)$ is rewritten into $Arg(\alpha \vee \beta) \wedge Arg(\alpha \vee \neg\beta)$, and
2. *decomposition:* $Arg(\alpha \vee \beta \vee \gamma)$ is rewritten into $Arg(\alpha \vee \gamma) \vee Arg(\beta \vee \gamma)$ when the corresponding LCI assertion holds.

The alternation between the two rules takes place because of the following. To decompose an argument $Arg(\check{X} \vee \check{Y})$, atoms X and Y must be independent. If they are not, the algorithm performs a case analysis on atoms Z (that make X and Y independent) in order to apply the decomposition rule. That is,

1. the algorithm rewrites $Arg(\check{X} \vee \check{Y})$ into
$$\bigwedge_{\check{Z}} Arg(\check{X} \vee \check{Y} \vee \check{Z})$$
using the case-analysis rule, and then

[10] A cutset of a directed acyclic graph is a set of nodes in the graph that satisfies the following property: if we remove the arcs outgoing from every node in the cutset, the resulting graph becomes singly-connected.

2. rewrites the result into
$$\bigwedge_{\check{Z}} Arg(\check{X} \vee \check{Z}) \vee Arg(\check{Y} \vee \check{Z})$$
using the decomposition rule.

Therefore, when an argument cannot be decomposed, it is first expanded using case-analysis and then decomposed. This process continues until we have a Boolean expression in which all arguments are local.

The algorithm

We now introduce some notation that is needed to state the algorithm as a set of rewrite rules. Let

- x be an arbitrary node in the database structure,
- u and u' be distinct parents of x,
- y and y' be distinct children of x, and
- U be all parents of x.

The algorithm can be described as a recursive and deterministic[11] application of the following rewrite rules:

$$Arg(\check{O}) \rightarrow (\pi(x) \vee \lambda(x)) \bigwedge (\pi(\neg x) \vee \lambda(\neg x)), \text{ for some node } x \text{ in the structure } \mathcal{G} \quad (1)$$

$$\pi(\check{x}) \rightarrow \bigwedge_{\check{U}} Arg^{\Delta_x}(\check{x} \vee \check{U}) \vee \bigvee_{\check{u} \models \check{U}} \pi_x(\check{u}) \quad (2)$$

$$\lambda(\check{x}) \rightarrow \Lambda(\check{x}) \vee \bigvee_{y} \lambda_y(\check{x}) \quad (3)$$

$$\pi_y(\check{x}) \rightarrow \pi(\check{x}) \vee \bigvee_{y'} \lambda_{y'}(\check{x}) \quad (4)$$

$$\lambda_x(\check{u}) \rightarrow \bigwedge_{\check{x}} \lambda(\check{x}) \vee \bigwedge_{\check{u} \models \check{U}} Arg^{\Delta_x}(\check{x} \vee \check{U}) \vee \bigvee_{\check{u'} \models \check{U}} \pi_x(\check{u'}) \quad (5)$$

$$\Lambda(\check{x}) \rightarrow \begin{cases} \mathit{true} & \text{if } \neg\check{x} \text{ appears in } \check{O}; \\ \mathit{false} & \text{otherwise.} \end{cases} \quad (6)$$

The algorithm starts with the argument $Arg(\check{O})$. It then keeps on applying the rewrites given above until it reaches an expression that contains only the connectives $\wedge$ and $\vee$ in addition to *true*, *false* and local arguments of the form $Arg^{\Delta_x}(\check{x} \vee \check{U})$. In its intermediate stages, the constructed expression will also include the following intermediate terms π, λ, π_y, λ_x and Λ, which will be rewritten into other expressions using the above rewrites. The algorithm is guaranteed to construct an expression that is free of these intermediate terms. Given that the database structure is singly connected, it is easy to verify that the rewrite process will terminate and

- Rewrite 1 will be applied only once,
- Rewrites 2, 3 and 6 will be applied at most once per node in the database structure, and
- Rewrites 4 and 5 will be applied at most once for each arc in the database structure.

Therefore, the complexity of the algorithm and the size of the constructed Boolean expression are linear in the number of nodes and arcs, and exponential only in the size of families of the database structure.

[11] Except for the first rewrite rule, which applies only once.

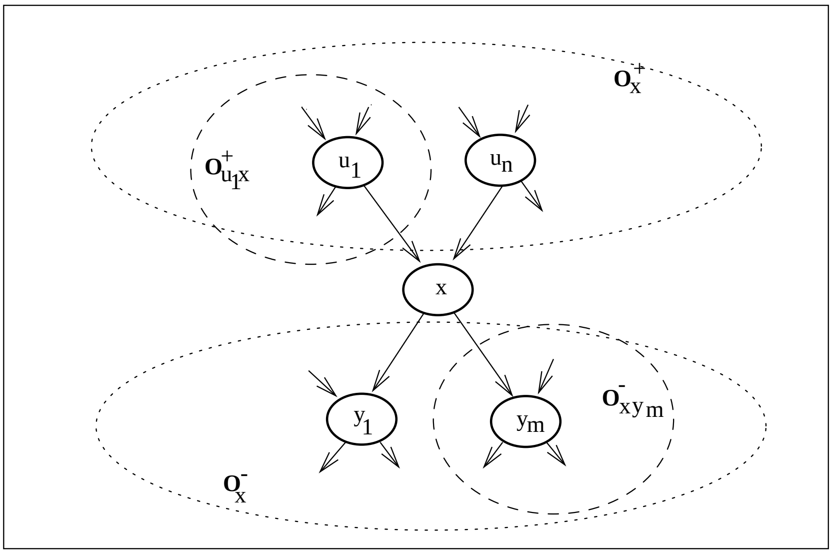

Figure 9: Decomposing a disjunctive clause in a singly-connected structure.

Deriving the rewrites

We will now go over the derivation of this algorithm. This discussion would typically be part of Appendix A, but we include it here to provide an example of the techniques involved in developing a structure-based algorithm. This is important because in certain applications, one may arrive at database structures that imply strong independences and therefore may be easy to solve using a specialized structure-based algorithm. To develop such an algorithm, one needs to go into an exercise similar to the following.

The algorithm starts by rewriting $Arg(\check{O})$ using the case-analysis rule into

$$Arg(x \vee \check{O}) \bigwedge Arg(\neg x \vee \check{O})$$

for some atom x in the structure (we will assume that x does not appear in $\check{O}$ without loss of generality). To decompose the argument $Arg(\check{x} \vee \check{O})$, the algorithm partitions the disjunctive clause $\check{O}$ into two parts: a clause $\check{\mathbf{O}}_x^-$ mentioning only descendants of x and a clause $\check{\mathbf{O}}_x^+$ mentioning only non-descendants of x; see Figure 9. This validates the decomposition:

$$\begin{aligned} Arg(\check{x} \vee \check{O}) &\longmapsto Arg(\check{x} \vee \check{\mathbf{O}}_x^+ \vee \check{\mathbf{O}}_x^-) \\ &\longmapsto \underbrace{Arg(\check{x} \vee \check{\mathbf{O}}_x^+)}_{\pi(\check{x})} \bigvee \underbrace{Arg(\check{x} \vee \check{\mathbf{O}}_x^-)}_{\lambda(\check{x})} \end{aligned}$$

because x d-separates any of its descendants from any of its non-descendants.

To decompose the argument $Arg(\check{x} \vee \check{\mathbf{O}}_x^-)$, the algorithm partitions the clause $\check{\mathbf{O}}_x^-$ into a number of clauses $\check{\mathbf{O}}_{xy}^-$, each about the nodes connected to one child y of x; see Figure 9. This validates the decomposition:

$$\begin{aligned} Arg(\check{x} \vee \check{\mathbf{O}}_x^-) &\longmapsto Arg(\check{x} \vee \bigvee_y \check{\mathbf{O}}_{xy}^-) \\ &\longmapsto \bigvee_y \underbrace{Arg(\check{x} \vee \check{\mathbf{O}}_{xy}^-)}_{\lambda_y(\check{x})} \end{aligned}$$

because x d-separates any nodes connected to one of its children from any nodes connected to other children.

To decompose the argument $Arg(\check{x} \vee \check{\mathbf{O}}_x^+)$, the algorithm applies the case-analysis rule:

$$Arg(\check{x} \vee \check{\mathbf{O}}_x^+) \longmapsto \bigwedge_{\check{U}} Arg(\check{x} \vee \check{U} \vee \check{\mathbf{O}}_x^+),$$

where $\check{U}$ is a conjunctive clause over U, the parents of x.

Using Theorem 11 in the Appendix, we can decompose the argument

$$Arg(\check{x} \vee \check{U} \vee \check{\mathbf{O}}_x^+)$$

into

$$Arg^{\Delta_x}(\check{x} \vee \check{U}) \bigvee Arg(\check{U} \vee \check{\mathbf{O}}_x^+).$$

Note here that the first argument is local, that is, can be computed using only the local database for x.

To decompose the argument $Arg(\check{U} \vee \check{\mathbf{O}}_x^+)$, the algorithm partitions the clause $\check{\mathbf{O}}_x^+$ into a number of clauses $\check{\mathbf{O}}_{ux}^+$, each about the nodes connected to x through its parent u; see Figure 9. This allows the application of the decomposition rule:

$$\begin{aligned} Arg(\check{U} \vee \check{\mathbf{O}}_x^+) &\longmapsto Arg(\bigvee_{\check{u} \models \check{U}} \check{u} \vee \check{\mathbf{O}}_{ux}^+) \\ &\longmapsto \bigvee_{\check{u} \models \check{U}} \underbrace{Arg(\check{u} \vee \check{\mathbf{O}}_{ux}^+)}_{\pi_x(\check{u})}, \end{aligned}$$

which is validated by d-separation: each parent u of x and the nodes in clause $\check{\mathbf{O}}_{ux}^+$ are d-separated from every other parent u' of x and the nodes in clause $\check{\mathbf{O}}_{u'x}^+$. The rewrite for $\lambda_y(\check{x})$ and that for $\pi_x(\check{u})$ can be derived similarly. Moreover, given that x may appear in $\check{O}$, the term Λ must be inserted in the Rewrites 1–6 as shown.

The algorithm presented in this section can be extended to multiply-connected databases straightforwardly, leading to an algorithm called *cutset conditioning* that is exponential in the size of a structure cutset [20, 19, 2]. There are more sophisticated extensions to multiply-connected structures, however, but they are outside the scope of this paper [14, 22, 3]. Note that multiply-connected structures are not necessarily harder than singly-connected ones in general. For example, n-bit adders lead to multiply-connected structures that behave computationally like singly-connected ones.

Other reasoning tasks

The algorithm we just presented computes the argument for a given disjunctive clause $\check{O}$. Given the duality between arguments and consequences, the algorithm can also be used to compute the consequence of a conjunctive clause $\hat{O}$ using

$$Cons(\hat{O}) \equiv \neg Arg(\neg\hat{O}).$$

The algorithm can also be used to decide entailment since

$$\Delta \models \alpha \text{ precisely when } Arg_{\emptyset}^{\Delta}(\alpha) \equiv true.$$

Therefore, if the structured database $(\mathcal{G}, \Delta)$ has no exogenous variables — that is, all atoms that appear in Δ also appear in $\mathcal{G}$ — then one can compute arguments and declare a clause $\check{O}$ as entailed by the database Δ iff the computed argument for $\check{O}$ is equivalent to *true*.

Given the duality between satisfiability and entailment in propositional logic, one can also use the algorithm for testing satisfiability.

8 Conclusion

We have presented a notion of conditional independence with respect to propositional databases that resembles its probabilistic counterpart. We have also presented several formulations of logical independence, together with their applications to deciding logical entailment, computing abductions and diagnoses. We have demonstrated how structuring a database around a directed acyclic graph can lead to explicating the independences it satisfies and how independence-based algorithms can then take advantage of this structure.

Our proposed approach for utilizing independence ties the computational complexity of logical reasoning to the availability of independence information and therefore to database structure. This leads to a computational paradigm in which the complexity of reasoning is parameterized by the topology of a database structure. This structure-based approach is at the heart of causal-networks in the probabilistic literature and constraint-networks in the constraint satisfaction literature [9, 12, 8]. In both cases, a graphical structure is the key aspect deciding the difficulty of a reasoning problem. This structure is what users need to control in order to ensure an appropriate response time for their applications. The probabilistic literature, for example, contains techniques for tweaking this structure to ensure certain response times, most of which can be adopted by our proposed framework.

A Proofs

The proofs will not have the same order as the theorems. We will first establish the equivalence between all LCI formulations, then prove the semi-graphoid axioms using the abduction formulation in Definition 10.

Proving equivalence between LCI definitions

We first prove the equivalence between dualities: entailment/satisfiability and diagnosis/abduction. We then prove the equivalence between belief-change and abduction formulations followed by the equivalence between belief-change and entailment formulations. We will then have covered Theorems 2, 3, 5 and 7.

Equivalence between entailment and consistency formulations

Suppose that

$$\Delta \models \check{X} \vee \check{Y} \vee \check{Z} \text{ precisely when } \Delta \models \check{X} \vee \check{Z} \text{ or } \Delta \models \check{Y} \vee \check{Z}$$

for all disjunctive clauses $\check{X}, \check{Y}, \check{Z}$. Then $\Delta \cup \{\neg\check{Z}, \neg\check{Y}, \neg\check{X}\}$ is inconsistent precisely when either $\Delta \cup \{\neg\check{Z}, \neg\check{X}\}$ or $\Delta \cup \{\neg\check{Z}, \neg\check{Y}\}$ is inconsistent for all disjunctive clauses $\check{X}, \check{Y}, \check{Z}$. This means that $\Delta \cup \{\hat{Z}, \hat{Y}, \hat{X}\}$ is satisfiable precisely when both $\Delta \cup \{\hat{Z}, \hat{X}\}$ and $\Delta \cup \{\hat{Z}, \hat{Y}\}$ are satisfiable for all conjunctive clauses $\hat{X}, \hat{Y}$ and $\hat{Z}$.

The other direction follows similarly. ∎

Equivalence between diagnosis and abduction formulations

We now prove the equivalence between the diagnosis and abduction formulations of LCI, therefore, proving the equivalence between Definitions 8 and 10 if independence.

By Theorem 6, we have that $Arg_W^{\Delta}(\alpha) \equiv \neg Cons_W^{\Delta}(\neg\alpha)$. Therefore,

$$\models Arg(\check{X} \vee \check{Y} \vee \check{Z}) \equiv Arg(\check{X} \vee \check{Z}) \vee Arg(\check{Y} \vee \check{Z})$$

is equivalent to

$$\models \neg Cons(\neg\check{X} \wedge \neg\check{Y} \wedge \neg\check{Z}) \equiv \neg Cons(\neg\check{X} \wedge \neg\check{Z}) \vee \neg Cons(\neg\check{Y} \wedge \neg\check{Z}),$$

which is equivalent to

$$\models \neg Cons(\neg\check{X} \wedge \neg\check{Y} \wedge \neg\check{Z}) \equiv \neg(Cons(\neg\check{X} \wedge \neg\check{Z}) \wedge Cons(\neg\check{Y} \wedge \neg\check{Z})),$$

and

$$\models Cons(\neg\check{X} \wedge \neg\check{Y} \wedge \neg\check{Z}) \equiv Cons(\neg\check{X} \wedge \neg\check{Z}) \wedge Cons(\neg\check{Y} \wedge \neg\check{Z}).$$

Therefore,

$$\models Arg(\check{X} \vee \check{Y} \vee \check{Z}) \equiv Arg(\check{X} \vee \check{Z}) \vee Arg(\check{Y} \vee \check{Z})$$

for all $\check{X}$, $\check{Y}$ and $\check{Z}$ iff

$$\models Cons(\hat{X} \wedge \hat{Y} \wedge \hat{Z}) \equiv Cons(\hat{X} \wedge \hat{Z}) \wedge Cons(\hat{Y} \wedge \hat{Z})$$

for all $\hat{X}$, $\hat{Y}$ and $\hat{Z}$. ∎

Equivalence between belief change and abduction formulations

We now prove that $Ind_{\Delta}^{b}(X, Z \cup W, Y)$ (Definition 3) is equivalent to $Ind_{(\Delta,W)}^{a}(X, Z, Y)$ (Definition 10).

The following observations (lemmas) are useful in understanding the proof of this theorem:

- Each sentence $\tilde{X}$ is equivalent to the conjunction of some $\check{X}$'s. Example: $a \wedge b$ is equivalent to $(a \vee b) \wedge (a \vee \neg b) \wedge (\neg a \vee b)$.
- Each sentence $\tilde{X}$ is equivalent to the disjunction of some $\hat{X}$'s. Example: $a \supset b$ is equivalent to $(a \wedge b) \vee (\neg a \wedge b) \vee (\neg a \wedge \neg b)$.
- Suppose that there is no $\check{X}$ and $\hat{Y}$ such that

1. $\Delta \cup \{\hat{Y}\}$ is consistent,
2. $\Delta \not\models \check{X}$,
3. $\Delta \cup \{\hat{Y}\} \models \check{X}$.

Then there is no $\check{X}$ and $\tilde{Y}$ such that

1. $\Delta \cup \{\tilde{Y}\}$ is consistent,
2. $\Delta \not\models \check{X}$,
3. $\Delta \cup \{\tilde{Y}\} \models \check{X}$.

Because if $\tilde{Y}$ is consistent with Δ and $\Delta \cup \{\tilde{Y}\} \models \check{X}$, then there must exist some $\hat{Y}$ such that $\hat{Y}$ is consistent with Δ and $\Delta \cup \{\hat{Y}\} \models \check{X}$.

- Suppose that there is no $\tilde{Y}$ and $\check{X}$ such that

  1. $\Delta \cup \{\tilde{Y}\}$ is consistent,
  2. $\Delta \not\models \check{X}$, and
  3. $\Delta \cup \{\tilde{Y}\} \models \check{X}$.

  Then there is no $\tilde{Y}$ and $\tilde{X}$ such that

  1. $\Delta \cup \{\tilde{Y}\}$ is consistent,
  2. $\Delta \not\models \tilde{X}$, and
  3. $\Delta \cup \{\tilde{Y}\} \models \tilde{X}$.

  Because if $\Delta \not\models \tilde{X}$ and $\Delta \cup \{\tilde{Y}\} \models \tilde{X}$, then there must exist some $\check{X}$ such that $\Delta \not\models \check{X}$ and $\Delta \cup \{\tilde{Y}\} \models \check{X}$.

- The following is always true:

  $$Arg(\check{Z} \vee \check{X}) \vee Arg(\check{Z} \vee \check{Y}) \models Arg(\check{Z} \vee \check{X} \vee \check{Y}).$$

  Because if

  $$\Delta \cup \{\tilde{W}\} \models \check{Z} \vee \check{X} \text{ or } \Delta \cup \{\tilde{W}\} \models \check{Z} \vee \check{Y},$$

  then also

  $$\Delta \cup \{\tilde{W}\} \models \check{Z} \vee \check{X} \vee \check{Y}.$$

The proof of the theorem is given below.

$\Longrightarrow$ Suppose that $Ind^a_{(\Delta,W)}(X, Z, Y)$ does not hold. That is, we have

$$Arg(\check{Z} \vee \check{X} \vee \check{Y}) \not\models Arg(\check{Z} \vee \check{X}) \vee Arg(\check{Z} \vee \check{Y})$$

for some $\check{X}, \check{Y}$ and $\check{Z}$. We want to show that $Ind^b_\Delta(X, Z \cup W, Y)$ does not hold either.

By supposition, there must exist some $\hat{W}$ such that

1. $\Delta \cup \{\hat{W}\} \models \check{Z} \vee \check{X} \vee \check{Y}$,
2. $\Delta \cup \{\hat{W}\} \not\models \check{Z} \vee \check{X}$,
3. $\Delta \cup \{\hat{W}\} \not\models \check{Z} \vee \check{Y}$.

This means that

1. $\Delta \cup \{\hat{W}, \neg\check{Z}, \neg\check{Y}\} \models \check{X}$,
2. $\Delta \cup \{\hat{W}, \neg\check{Z}\} \not\models \check{X}$,
3. $\Delta \cup \{\hat{W}, \neg\check{Z}, \neg\check{Y}\}$ is consistent.

Therefore, $Ind^b_\Delta(X, Z \cup W, Y)$ does not hold. We have then proved that $Ind^b_\Delta(X, Z \cup W, Y)$ of implies $Ind^a_{(\Delta,W)}(X, Z, Y)$.

$\Longleftarrow$ Suppose that $Ind^b_\Delta(X, Z \cup W, Y)$ does not hold. Then there must exist $\hat{W}$, $\hat{Z}$, $\tilde{Y}$, and $\tilde{X}$ such that

1. $\Delta \cup \{\hat{W}, \hat{Z}\} \not\models \tilde{X}$,
2. $\Delta \cup \{\hat{W}, \hat{Z}, \tilde{Y}\} \models \tilde{X}$,
3. $\Delta \cup \{\hat{W}, \hat{Z}, \tilde{Y}\}$ is consistent.

By the above lemmas, there must exist $\hat{Y}$ and $\check{X}$ such that

1. $\Delta \cup \{\hat{W}, \hat{Z}\} \not\models \check{X}$,
2. $\Delta \cup \{\hat{W}, \hat{Z}, \hat{Y}\} \models \check{X}$,
3. $\Delta \cup \{\hat{W}, \hat{Z}, \hat{Y}\}$ is consistent.

That is, we have

1. $\Delta \cup \{\hat{W}\} \not\models \neg\hat{Z} \vee \check{X}$,
2. $\Delta \cup \{\hat{W}\} \models \neg\hat{Z} \vee \check{X} \vee \neg\hat{Y}$,
3. $\Delta \cup \{\hat{W}\} \not\models \neg\hat{Z} \vee \neg\hat{Y}$.

Therefore,

$$Arg(\neg\hat{Z} \vee \check{X} \vee \neg\hat{Y}) \not\models Arg(\neg\hat{Z} \vee \check{X}) \vee Arg(\neg\hat{Z} \vee \neg\hat{Y}),$$

which means that $Ind^a_{(\Delta,W)}(X, Z, Y)$ does not hold. We have then proved that $Ind^a_{(\Delta,W)}(X, Z, Y)$ implies $Ind^b_\Delta(X, Z \cup W, Y)$. ∎

Equivalence between belief change and entailment formulations

Suppose that $Ind^b_\Delta(X, Z, Y)$. Then $Ind^a_{(\Delta,W=\emptyset)}(X, Z, Y)$ and, hence,

$$\models Arg(\check{X} \vee \check{Y} \vee \check{Z}) \equiv Arg(\check{X} \vee \check{Z}) \vee Arg(\check{Y} \vee \check{Z}).$$

Since W is empty, arguments with respect to (Δ, W) can only be *true* or *false*. Moreover,

- $Arg(\alpha) = true$ iff $\Delta \models \alpha$.
- $Arg(\alpha) = false$ iff $\Delta \not\models \alpha$.

Therefore,

$$\models Arg(\check{X} \vee \check{Y} \vee \check{Z}) \equiv Arg(\check{X} \vee \check{Z}) \vee Arg(\check{Y} \vee \check{Z})$$

is equivalent to

$$\Delta \models \check{X} \vee \check{Y} \vee \check{Z} \text{ precisely when } \Delta \models \check{X} \vee \check{Z} \text{ or } \Delta \models \check{Y} \vee \check{Z}$$

and we have $Ind^e_\Delta(X, Z, Y)$. The other direction follows similarly. ∎

Proof of Theorem 1

The proof uses Definitions 8 of independence. If we choose $W = \emptyset$, we obtain a proof with respect to Definitions 3 of independence.

1. Trivial Independence: $Ind_\Delta(X, Z, \emptyset)$.

   Note that $\check{\emptyset}$ is *false* and that $Arg(\check{Z}) \models Arg(\check{X} \vee \check{Z})$ since $\check{Z} \models \check{X} \vee \check{Z}$. We have:

$$\begin{aligned} \models Arg(\check{X} \vee \check{\emptyset} \vee \check{Z}) &\equiv Arg(\check{X} \vee \mathit{false} \vee \check{Z}) \\ &\equiv Arg(\check{X} \vee \check{Z}) \\ &\equiv Arg(\check{X} \vee \check{Z}) \vee Arg(\check{Z}) \\ &\equiv Arg(\check{X} \vee \check{Z}) \vee Arg(\check{\emptyset} \vee \check{Z}). \end{aligned}$$

   Therefore, $Ind_\Delta(X, Z, \emptyset)$.

2. Symmetry: $Ind_{(\Delta,W)}(X, Z, Y)$ precisely when $Ind_{(\Delta,W)}(Y, Z, X)$. Follows directly from definition.

3. Decomposition: If $Ind_{(\Delta,W)}(X \cup L, Z, Y)$, then $Ind_{(\Delta,W)}(X, Z, Y)$.

   This proof uses the following facts:

   (a) If $\models \phi_i \equiv \psi_i$, then $\models \bigwedge_i \phi_i \equiv \bigwedge_i \psi_i$.
   (b) $\models \bigwedge_i (\phi \vee \psi_i) \equiv \phi \vee \bigwedge_i \psi_i$.
   (c) $\models \bigwedge_{\check{X}} Arg(\check{X} \vee \check{Y}) \equiv Arg(\check{Y})$.

   Suppose that $Ind_{(\Delta,W)}(X \cup L, Z, Y)$:

$$\models Arg(\check{Z} \vee \check{X} \vee \check{L} \vee \check{Y}) \equiv Arg(\check{Z} \vee \check{X} \vee \check{L}) \vee Arg(\check{Z} \vee \check{Y}).$$

   We have

$$\models \bigwedge_{\check{L}} Arg(\check{Z} \vee \check{X} \vee \check{L} \vee \check{Y}) \equiv \bigwedge_{\check{L}} Arg(\check{Z} \vee \check{X} \vee \check{L}) \vee Arg(\check{Z} \vee \check{Y}),$$

$$\begin{aligned} \models Arg(\check{Z} \vee \check{X} \vee \check{Y}) &\equiv Arg(\check{Z} \vee \check{Y}) \vee \bigwedge_{\check{L}} Arg(\check{Z} \vee \check{X} \vee \check{L}) \\ &\equiv Arg(\check{Z} \vee \check{Y}) \vee Arg(\check{Z} \vee \check{X}). \end{aligned}$$

   Hence, $Ind_{(\Delta,W)}(X, Z, Y)$.

4. Weak Union: If $Ind_{(\Delta,W)}(X \cup L, Z, Y)$, then $Ind_{(\Delta,W)}(L, Z \cup X, Y)$.

   Suppose that $Ind_{(\Delta,W)}(X \cup L, Z, Y)$:

$$\models Arg(\check{Z} \vee \check{X} \vee \check{L} \vee \check{Y}) \equiv Arg(\check{Z} \vee \check{X} \vee \check{L}) \vee Arg(\check{Z} \vee \check{Y}).$$

   Since $Arg(\check{Z} \vee \check{X}) \models Arg(\check{Z} \vee \check{X} \vee \check{L})$, we have

$$\models Arg(\check{Z} \vee \check{X} \vee \check{L}) \equiv Arg(\check{Z} \vee \check{X} \vee \check{L}) \vee Arg(\check{Z} \vee \check{X}),$$

   and

$$\models Arg(\check{Z} \vee \check{X} \vee \check{L} \vee \check{Y}) \equiv [Arg(\check{Z} \vee \check{X} \vee \check{L}) \vee Arg(\check{Z} \vee \check{X})] \vee Arg(\check{Z} \vee \check{Y}).$$

   Moreover, by decomposition, we have $Ind_{(\Delta,W)}(X, Z, Y)$:

$$\models Arg(\check{Z} \vee \check{X} \vee \check{Y}) \equiv Arg(\check{Z} \vee \check{X}) \vee Arg(\check{Z} \vee \check{Y}).$$

   Hence,

$$\models Arg(\check{Z} \vee \check{X} \vee \check{L} \vee \check{Y}) \equiv Arg(\check{Z} \vee \check{X} \vee \check{L}) \vee Arg(\check{Z} \vee \check{X} \vee \check{Y})$$

   and $Ind_{(\Delta,W)}(L, Z \cup X, Y)$.

5. <u>Contraction: If $Ind_{(\Delta,W)}(X,Z,Y)$ and $Ind_{(\Delta,W)}(L,Z\cup X,Y)$, then $Ind_{(\Delta,W)}(X\cup L,Z,Y)$.</u>

   Suppose that $Ind_{(\Delta,W)}(X,Z,Y)$ and $Ind_{(\Delta,W)}(L,Z\cup X,Y)$:

$$\models Arg(\check{Z}\vee\check{X}\vee\check{Y})\equiv Arg(\check{Z}\vee\check{Y})\vee Arg(\check{Z}\vee\check{X}) \text{ and}$$

$$\models Arg(\check{Z}\vee\check{X}\vee\check{L}\vee\check{Y})\equiv Arg(\check{Z}\vee\check{X}\vee\check{L})\vee Arg(\check{Z}\vee\check{X}\vee\check{Y}).$$

   Expanding the second equation using the first:

$$\models Arg(\check{Z}\vee\check{X}\vee\check{L}\vee\check{Y})\equiv Arg(\check{Z}\vee\check{X}\vee\check{L})\vee Arg(\check{Z}\vee\check{Y})\vee Arg(\check{Z}\vee\check{X}),$$

   which reduces to

$$\models Arg(\check{Z}\vee\check{X}\vee\check{L}\vee\check{Y})\equiv Arg(\check{Z}\vee\check{X}\vee\check{L})\vee Arg(\check{Z}\vee\check{Y}),$$

   because

$$\models Arg(\check{Z}\vee\check{X})\models Arg(\check{Z}\vee\check{X}\vee\check{L}).$$

   Hence, $Ind_{(\Delta,W)}(X\cup L,Z,Y)$. ∎

Proof of Theorem 4

Suppose that $Cons^{\Delta}_{W}(\gamma)=\hat{W}_1\vee\ldots\vee\hat{W}_n$. Then $\Delta\cup\{\gamma\}\models\hat{W}_1\vee\ldots\vee\hat{W}_n$. We need to prove that

1. each $\hat{W}_i$ is a diagnosis, and
2. if $\hat{W}$ is a diagnosis, then $\hat{W}\models\hat{W}_1\vee\ldots\vee\hat{W}_n$.

To show that $\hat{W}_i$ is a diagnosis, we need to show that $\Delta\cup\{\gamma,\hat{W}_i\}$ is consistent. Suppose that $\Delta\cup\{\gamma,\hat{W}_i\}$ is not consistent. Then $\Delta\cup\{\gamma\}\models\neg\hat{W}_i$ and

$$\Delta\cup\{\gamma\}\models\hat{W}_1\vee\ldots\vee\hat{W}_{i-1}\vee\hat{W}_{i+1}\vee\ldots\vee\hat{W}_n.$$

But this contradicts with $\hat{W}_1\vee\ldots\vee\hat{W}_n$ being a logically strongest sentence entailed by $\Delta\cup\{\gamma\}$.

Suppose that $\hat{W}$ is a diagnosis. Then $\Delta\cup\{\gamma,\hat{W}\}$ is consistent and $\Delta\cup\{\gamma,\hat{W}\}\models\hat{W}_1\vee\ldots\vee\hat{W}_n$. This means that $\hat{W}$ must belong to $\hat{W}_1\vee\ldots\vee\hat{W}_n$, otherwise $\Delta\cup\{\gamma,\hat{W}\}\models\neg\hat{W}_1,\ldots,\Delta\cup\{\gamma,\hat{W}\}\models\neg\hat{W}_n$. This means that $\Delta\cup\{\gamma,\hat{W}\}$ is inconsistent, which we know it is not. ∎

Proof of Theorem 6

Suppose that $Arg^{\Delta}_{W}(\alpha)\equiv\beta$. Then β is a logically weakest sentence that is constructed from atoms W such that

$$\Delta\cup\{\beta\}\models\alpha.$$

It then follows that $\neg\beta$ is a logically strongest sentence constructed from W such that

$$\Delta\cup\{\neg\alpha\}\models\neg\beta.$$

Therefore, $\neg\beta$ is the consequence of $\neg\alpha$, that is, $Cons^{\Delta}_{W}(\neg\alpha)\equiv\neg\beta$. This implies that $Cons^{\Delta}_{W}(\neg\alpha)\equiv\neg Arg^{\Delta}_{W}(\alpha)$ and $Arg^{\Delta}_{W}(\alpha)\equiv\neg Cons^{\Delta}_{W}(\neg\alpha)$. ∎

Proof of Theorem 8

We will prove $Ind_{\Delta}(\{n\},U\cup W,N)$ by assuming that n is a leaf node in $\mathcal{G}$. The proof will also work when n is not a leaf node since, by Corollary 1 of Theorem 10, we can prune the descendants of n to make it a leaf node without invalidating the proof.

To prove $Ind_{\Delta}(\{n\},U\cup W,N)$, we need to prove

$$\Delta\cup\{\hat{W},\hat{U}\}\models\tilde{n} \text{ precisely when } \Delta\cup\{\hat{W},\hat{U},\tilde{N}\}\models\tilde{n},$$

for all $\hat{W}, \hat{U}, \tilde{N}, \tilde{n}$ where $\Delta \cup \{\hat{W}, \hat{U}, \tilde{N}\}$ is consistent. The direction

$$\Delta \cup \{\hat{W}, \hat{U}\} \models \tilde{n} \text{ only if } \Delta \cup \{\hat{W}, \hat{U}, \tilde{N}\} \models \tilde{n}$$

is trivial. The direction

$$\Delta \cup \{\hat{W}, \hat{U}, \tilde{N}\} \models \tilde{n} \text{ only if } \Delta \cup \{\hat{W}, \hat{U}\} \models \tilde{n}$$

is equivalent to

$$\Delta \cup \{\hat{W}, \hat{U}\} \not\models \tilde{n} \text{ only if } \Delta \cup \{\hat{W}, \hat{U}, \tilde{N}\} \not\models \tilde{n}$$

and

$$\Delta' \cup \Delta_n \cup \{\hat{W}, \hat{U}\} \not\models \tilde{n} \text{ only if } \Delta' \cup \Delta_n \cup \{\hat{W}, \hat{U}, \tilde{N}\} \not\models \tilde{n}$$

where $\Delta' = \Delta \setminus \Delta_n$.

Suppose that

$$\Delta' \cup \Delta_n \cup \{\hat{W}, \hat{U}\} \not\models \tilde{n}.$$

One of the following must be true:

1. $\Delta_n \cup \{\hat{W}, \hat{U}\}$ is equivalent to $\{\hat{W}, \hat{U}\}$ or
2. $\Delta_n \cup \{\hat{W}, \hat{U}\}$ is equivalent to $\{\neg\tilde{n}, \hat{W}, \hat{U}\}$.

This holds because by definition of a local database, Δ_n must be equivalent to a set of disjunctive clauses, each referring only to W, U and n and containing either $\tilde{n}$ or $\neg\tilde{n}$. Moreover, each one of these clauses must either

1. be entailed by $\{\hat{W}, \hat{U}\}$ in which case it can be removed from Δ_n or
2. resolve with $\{\hat{W}, \hat{U}\}$ to yield $\neg\tilde{n}$ in which case it can be replaced by $\neg\tilde{n}$.[12]

This means that we either have

1. $\Delta' \cup \Delta_n \cup \{\hat{W}, \hat{U}\}$ is equivalent to $\Delta' \cup \{\hat{W}, \hat{U}\}$ and $\Delta' \cup \{\hat{W}, \hat{U}\} \not\models \tilde{n}$ or
2. $\Delta' \cup \Delta_n \cup \{\hat{W}, \hat{U}\}$ is equivalent to $\Delta' \cup \{\neg\tilde{n}, \hat{W}, \hat{U}\}$ and $\Delta' \cup \{\neg\tilde{n}, \hat{W}, \hat{U}\} \not\models \tilde{n}$,

and therefore either

1. $\Delta' \cup \{\hat{W}, \hat{U}, \tilde{N}\} \not\models \tilde{n}$ since $\Delta' \cup \{\hat{W}, \hat{U}, \tilde{N}\}$ is consistent and does not refer to n, or
2. $\Delta' \cup \{\neg\tilde{n}, \hat{W}, \hat{U}, \tilde{N}\} \not\models \tilde{n}$ since $\Delta' \cup \{\neg\tilde{n}, \hat{W}, \hat{U}, \tilde{N}\}$ is consistent.

In either case, it follows that

$$\Delta \cup \{\hat{W}, \hat{U}, \tilde{N}\} \not\models \tilde{n}. \quad \blacksquare$$

Proof of Theorem 9

Verma has shown in [25] that if

1. N is a set of atomic propositions,
2. $\mathcal{I}$ is an independence relation over N, a subset of $2^N \times 2^N \times 2^N$, that satisfies the semi-graphoid axioms,
3. $\mathcal{G}$ is a DAG over nodes N,
4. each node n in $\mathcal{G}$ is independent of its non-descendants given its parents according to $\mathcal{I}$,

then whenever X and Y are d-separated by Z in $\mathcal{G}$, we must have $I(X, Z, Y)$.

Given a structured database $(\mathcal{G}, \Delta)$, we know that the independence relation Ind_Δ satisfies the semi-graphoid axioms. And by Theorem 8, we know that each node in $\mathcal{G}$ is independent of its non-descendants given its parents, according to Ind_Δ. Therefore, whenever X and Y are d-separated by Z in $\mathcal{G}$, we must have $Ind_\Delta(X, Z, Y)$. $\blacksquare$

[12] The result of the resolution cannot be $\tilde{n}$ since $\Delta' \cup \Delta_n \cup \{\hat{W}, \hat{U}\} \not\models \tilde{n}$.

Proof of Theorem 10

1. The equivalence between $\Delta \models \check{N}$ and $\Delta' \models \check{N}$ follows immediately from the proof in (2) below.

2. The consistency of $\Delta \cup \{\hat{N}\}$ is equivalent to the consistency of

$$\Delta \cup \{ \bigvee_{\hat{W},\hat{A}} \hat{N} \wedge \hat{W} \wedge \hat{A} \},$$

   where A are the other atoms of $\mathcal{G}$ (different from n and N). Similarly, the consistency of $\Delta' \cup \{\hat{N}\}$ is equivalent to the consistency of

$$\Delta' \cup \{ \bigvee_{\hat{W},\hat{A}} \hat{N} \wedge \hat{W} \wedge \hat{A} \}.$$

   Therefore, it suffices to show that the consistency of each $\Delta \cup \{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$ is equivalent to the consistency of its corresponding $\Delta' \cup \{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$. Suppose that $\Delta \cup \{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$ is consistent. Then $\Delta' \cup \{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$ must also be consistent since $\Delta' \subseteq \Delta$. Now suppose that $\Delta' \cup \{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$ is consistent. Then it must be equivalent to $\{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$ since $N \cup W \cup A$ are all the atoms appearing in Δ. Therefore, $\Delta \cup \{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$ must be equivalent to $\Delta_n \cup \{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$ since $\Delta = \Delta' \cup \Delta_n$. By the definition of local database Δ_n, it follows that $\Delta_n \cup \{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$ must be consistent.[13]

   We have therefore established the equivalence between the consistency of $\Delta \cup \{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$ and that of $\Delta' \cup \{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$. By that we have also established that $\Delta \cup \{\hat{N}\}$ is consistent iff $\Delta' \cup \{\hat{N}\}$ is.

3. Note that $Arg^{\Delta}_{W}(\check{N})$ is equivalent to $Arg^{\Delta}_{W \cup Z}(\check{N})$ if atoms Z do not appear in Δ. Therefore, $Arg^{\Delta'}_{W'}(\check{N})$ is equivalent to $Arg^{\Delta'}_{W}(\check{N})$ since $W \setminus W'$ do not appear in Δ'. All we need to show then is the equivalence between $Arg^{\Delta}_{W}(\check{N})$ and $Arg^{\Delta'}_{W}(\check{N})$, which can be established by proving that $\Delta \cup \{\hat{W}\} \models \check{N}$ iff $\Delta' \cup \{\hat{W}\} \models \check{N}$. This can be proved as shown in (2) above.

4. The equivalence between $Cons^{\Delta}_{W}(\hat{N})$ and $Cons^{\Delta'}_{W'}(\hat{N})$ follows immediately from (3) and Theorem 6. ∎

Proof of Corollary 1

What we need to show is

$$Arg^{\Delta}_{W}(\check{X} \vee \check{Z} \vee \check{Y}) \equiv Arg^{\Delta}_{W}(\check{X} \vee \check{Z}) \vee Arg^{\Delta}_{W}(\check{Y} \vee \check{Z})$$

precisely when

$$Arg^{\Delta'}_{W'}(\check{X} \vee \check{Z} \vee \check{Y}) \equiv Arg^{\Delta'}_{W'}(\check{X} \vee \check{Z}) \vee Arg^{\Delta'}_{W'}(\check{Y} \vee \check{Z}).$$

This follows from Theorem 10, according to which

$\models Arg^{\Delta}_{W}(\check{X} \vee \check{Z} \vee \check{Y}) \equiv Arg^{\Delta'}_{W'}(\check{X} \vee \check{Z} \vee \check{Y})$,

$\models Arg^{\Delta}_{W}(\check{X} \vee \check{Z}) \equiv Arg^{\Delta'}_{W'}(\check{X} \vee \check{Z})$ and

$\models Arg^{\Delta}_{W}(\check{Y} \vee \check{Z}) \equiv Arg^{\Delta'}_{W'}(\check{Y} \vee \check{Z})$. ∎

Theorem 11 *Let $(\mathcal{G}, \Delta)$ be a structured database, n be a node in $\mathcal{G}$, U be its parents and N be some of its other ancestors. We then have*

$$Arg^{\Delta}(\check{n} \vee \check{U} \vee \check{N}) \equiv Arg^{\Delta_n}(\check{n} \vee \check{U}) \vee Arg^{\Delta}(\check{U} \vee \check{N}).$$

[13] Recall that any disjunctive clause entailed by database Δ_n must include atom n, which does not appear in $\{\hat{N} \wedge \hat{W} \wedge \hat{A}\}$.

Proof of Theorem 11

By Theorem 8, we have

$$\models Arg^{\Delta}_{W}(\check{n} \vee \check{U} \vee \check{N}) \equiv Arg^{\Delta}_{W}(\check{n} \vee \check{U}) \vee Arg^{\Delta}_{W}(\check{U} \vee \check{N})$$

since $Ind_{\Delta}(\{n\}, U \cup W, N)$. Therefore, it suffices to show

$$\models Arg^{\Delta}_{W}(\check{n} \vee \check{U}) \equiv Arg^{\Delta_n}_{W}(\check{n} \vee \check{U}) \vee Arg^{\Delta}_{W}(\check{U})$$

since $Arg^{\Delta}_{W}(\check{U})$ entails $Arg^{\Delta}_{W}(\check{U} \vee \check{N})$. The direction

$$Arg^{\Delta_n}_{W}(\check{n} \vee \check{U}) \vee Arg^{\Delta}_{W}(\check{U}) \models Arg^{\Delta}_{W}(\check{n} \vee \check{U})$$

is trivial. To show the direction

$$Arg^{\Delta}_{W}(\check{n} \vee \check{U}) \models Arg^{\Delta_n}_{W}(\check{n} \vee \check{U}) \vee Arg^{\Delta}_{W}(\check{U})$$

we need to prove that $\Delta \cup \{\hat{W}\} \models \check{n} \vee \check{U}$ implies either $\Delta_n \cup \{\hat{W}\} \models \check{n} \vee \check{U}$ or $\Delta \cup \{\hat{W}\} \models \check{U}$. If we take $\hat{n}$ to be $\neg\check{n}$ and $\hat{U}$ to be $\neg\check{U}$, then we need to show that $\Delta \cup \{\hat{W}, \hat{n}, \hat{U}\}$ is inconsistent only if $\Delta_n \cup \{\hat{W}, \hat{n}, \hat{U}\}$ or $\Delta \cup \{\hat{W}, \hat{U}\}$ is inconsistent.

Suppose that $\Delta \cup \{\hat{W}, \hat{n}, \hat{U}\}$ is inconsistent and $\Delta_n \cup \{\hat{W}, \hat{n}, \hat{U}\}$ is consistent. Then $\Delta_n \cup \{\hat{W}, \hat{n}, \hat{U}\}$ is equivalent to $\{\hat{W}, \hat{n}, \hat{U}\}$ and therefore $\Delta \cup \{\hat{W}, \hat{n}, \hat{U}\}$ is equivalent to $\Delta/\Delta_n \cup \{\hat{W}, \hat{n}, \hat{U}\}$. This means that $\Delta/\Delta_n \cup \{\hat{W}, \hat{n}, \hat{U}\}$ is inconsistent, which implies that $\Delta/\Delta_n \cup \{\hat{W}, \hat{U}\}$ is also inconsistent since atom n does not appear in $\Delta/\Delta_n \cup \{\hat{W}, \hat{U}\}$. Therefore, $\Delta \cup \{\hat{W}, \hat{U}\}$ is inconsistent. ∎

References

[1] B.D. D'Ambrosio. Local Expression Languages for Probabilistic Dependence. *International Journal of Approximate Reasoning*, 11:1–15, 1994.

[2] Adnan Darwiche. *A Symbolic Generalization of Probability Theory*. PhD thesis, Stanford University, 1992.

[3] Adnan Darwiche. Conditioning algorithms for exact and approximate inference in causal networks. In *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 99–107, 1995.

[4] Adnan Darwiche. Model-based diagnosis using causal networks. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, pages 211–217, 1995.

[5] Adnan Darwiche and Judea Pearl. Symbolic causal networks. In *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI)*, pages 238–244, 1994.

[6] Johan de Kleer. An assumption-based TMS. *Artificial Intelligence*, 28:127–162, 1986.

[7] Johan de Kleer, Alan K. Mackworth, and Raymond Reiter. Characterizing diagnoses and systems. *Artificial Intelligence*, 56(2-3):197–222, 1992.

[8] Rina Dechter and Avi Dechter. Belief maintenance in dynamic constraint networks. In *Proceedings of AAAI*, pages 37–42. AAAI, 1988.

[9] Rina Dechter and Avi Dechter. Structure-driven algorithms for truth maintenance. *Artificial Intelligence*, 1994. To appear.

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

[11] Peter Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. The MIT press, 1988.

[12] Hector Geffner and Judea Pearl. An improved constraint-propagation algorithm for diagnosis. In *Proceedings of IJCAI*, pages 1105–1111, Milan, Italy, 1987.

[13] Russell Greiner and Devika Subramanian. Relevance. In *Working Notes, AAAI-94 Fall Symposium Series*, The Monteleone Hotel, New Orleans, Louisiana, 1994.

[14] F. V. Jensen, S.L. Lauritzen, and K.G. Olesen. Bayesian updating in recursive graphical models by local computation. *Computational Statistics Quarterly*, 4:269–282, 1990.

[15] Alon Levy. *Irrelevance Reasoning in Knowledge Based Systems*. PhD thesis, Stanford University, Stanford, CA, 1993.

[16] Alon Levy, Inderpal Singh Mumick, and Yehoshua Sagiv. Query optimization by predicate movearound. In *Proceedings of VLDB-94*, Santiago, Chile, 1994.

[17] Alon Levy and Yehoshua Sagiv. Queries independent of updates. In *Proceedings of the 19th VLDB Conference*, pages 171–181, Dublin, Ireland, 1993.

[18] Z. Li and B.D. D'Ambrosio. Efficient Inference in Bayes Networks as a Combinatorial Optimization Problem. *International Journal of Approximate Reasoning*, 11:55–81, 1994.

[19] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, Inc., San Mateo, California, 1988.

[20] Mark A. Peot and Ross D. Shachter. Fusion and propagation with multiple observations in belief networks. *Artificial Intelligence*, 48(3):299–318, 1991.

[21] Ray Reiter and Johan de Kleer. Foundations of assumption-based truth maintenance systems: Preliminary report. In *Proceedings of AAAI*, pages 183–188. AAAI, 1987.

[22] R. Shachter, S.K. Andersen, and P. Szolovits. Global Conditioning for Probabilistic Inference in Belief Networks. In *Proc. Tenth Conference on Uncertainty in AI*, pages 514–522, Seattle WA, 1994.

[23] Devika Subramanian. *A Theory of Justified Reformulations*. PhD thesis, Stanford University, Stanford, CA, 1989.

[24] Devika Subramanian and Michael R. Genesereth. The relevance of irrelevance. In *Proc. International Joint Conference on Artifical Intelligence (IJCAI)*, pages 416–422, 1987.

[25] Tom Verma and Judea Pearl. Causal networks: Semantics and expressiveness. In *Proceedings of the* 4^{th} *Workshop on Uncertainty in AI*, pages 352–359, Minneapolis, MN, 1988.

[26] Nic Wilson. Generating graphoids from generalised conditional probability. In *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 583–590, 1994.