Logical Properties of Stable Conditional Independence

Mathias Niepert and Dirk Van Gucht
Department of Computer Science
Indiana University
Bloomington, IN, USA

Abstract

We utilize recent results concerning a complete axiomatization of stable conditional independence (CI) relative to discrete probability measures to derive perfect model properties of stable CI structures. We show that stable CI can be interpreted as a generalization of undirected graphical models and establish a connection between sets of stable CI statements and propositional formulae in conjunctive normal form. Consequently, we derive that the implication problem for stable CI is coNP-complete. Finally, we show that SAT solvers can be employed to efficiently decide the implication problem and to compute non-redundant representations of stable CI, even for instances involving hundreds of variables.

1 Introduction

The importance of stable conditional independence for reducing the complexity of representation of conditional independence structures has recently been established (de Waal and van der Gaag, 2004). Stable CI is an alternative to graphical models in representing and reasoning with conditional independence. A good understanding of its logical and algorithmic properties could lead to new theoretical insights and applications in the field of uncertain reasoning and data mining. While several results regarding the characteristics of stable CI structures exist (Matúš, 1992)(de Waal and van der Gaag, 2004)(de Waal and van der Gaag, 2005), no study has investigated its logical properties as it was done for general CI and graphical models relative to the class of discrete probability measures (Geiger and Pearl, 1993). We use recent results concerning a complete axiomatization of stable CI relative to discrete probability measures (Niepert et al., 2008) to show that (1) stable CI has perfect models relative to discrete probability measures, (2) for some sets of stable CI statements there exists no perfect model relative to binary probability measures, and (3) the number of distinct stable CI structures grows at least double exponentially with the number of statistical variables. We also derive that stable CI structures can be interpreted as a generalization of undirected graphical models: for every UG model there exists a stable CI structure, and if a discrete probability measure is (perfectly) Markovian w.r.t. the UG model, then it satisfies (exactly) all the CI statements of the stable CI structure. We establish a direct connection between sets of stable CI statements and propositional formulae in conjunctive normal form and use this connection to show that the implication problem for stable conditional independence is coNP-complete. Finally, we show that existing SAT solvers can be employed to efficiently decide the implication problem and to compute non-redundant representations of stable CI, even for instances involving hundreds of variables.

2 Preliminaries

Definition 1. Throughout this paper, S will be a finite, implicit set of attributes (discrete statistical variables). The expression $I(A, B|C)$, with A, B, and C pairwise disjoint subsets of S, is called a *conditional independence (CI) statement*. If $ABC = S$, we say that $I(A, B|C)$ is *saturated*. If $A = \emptyset$ or $B = \emptyset$ or both, we say that $I(A, B|C)$ is *trivial*.

A1: $I(A, B|C) \rightarrow I(B, A|C)$
A2: $I(A, BD|C) \rightarrow I(A, D|C)$
A3: $I(A, B|CD) \wedge I(A, D|C) \rightarrow I(A, BD|C)$
A4: $I(A, B|C) \rightarrow I(A, B|CD)$
A5: $I(A, B|C) \wedge I(D, E|AC) \wedge I(D, E|BC) \rightarrow I(D, E|C)$

Figure 1: The inference rules of system $\mathcal{A}$.

The set of inference rules in Figure 1 will be denoted by $\mathcal{A}$. The *symmetry* (A1), *decomposition* (A2), and *contraction* (A3) rules are part of the semi-graphoid axioms (Pearl, 1988). *Strong union* (A4) and *strong contraction* (A5) are additional inference rules. The derivability of a CI statement c from a set of CI statements $\mathcal{C}$ under the inference rules of system $\mathcal{A}$ is denoted by $\mathcal{C} \vdash c$. The *closure* of $\mathcal{C}$ under $\mathcal{A}$, denoted $\mathcal{C}^+$, is the set $\{c \mid \mathcal{C} \vdash c\}$. Even though triviality is a sound inference rule, we will not mention it explicitly in the rest of the paper. Trivial CI statements are assumed to be implicitly present.

Definition 2. A *probability model* over $S = \{s_1, \ldots, s_n\}$ is a pair (dom, P), where dom is a domain mapping that maps each s_i to a finite domain $dom(s_i)$ and P is a probability measure having $dom(s_1) \times \cdots \times dom(s_n)$ as its sample space. For $A = \{a_1, \ldots, a_k\} \subseteq S$ we will say that $\mathbf{a}$ is a domain vector of A if $\mathbf{a} \in dom(a_1) \times \cdots \times dom(a_k)$. If $dom(s_i) = \{0, 1\}$ we say that the probability model is *binary*.

In what follows, we will only refer to probability measures, keeping their underlying probability models implicit. The class of discrete probability measures will be denoted by $\mathcal{P}$ and the class of binary probability measures by $\mathcal{B}$.

Definition 3. Let $I(A, B|C)$ be a CI statement and let P be a probability measure. We say that P *satisfies* $I(A, B|C)$ if for every domain vector $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ of A, B, and C, respectively, $P(\mathbf{c})P(\mathbf{a}, \mathbf{b}, \mathbf{c}) = P(\mathbf{a}, \mathbf{c})P(\mathbf{b}, \mathbf{c})$.

Relative to the notion of *satisfaction* we can now define the *conditional independence implication problem.*

Definition 4 (Probabilistic CI implication problem)**.** Let $\mathcal{C}$ be a set of CI statements, let c be a CI statement, and let $\mathcal{P}$ be the class of discrete probability measures. We say that $\mathcal{C}$ *implies* c relative to $\mathcal{P}$, and write $\mathcal{C} \models c$, if each measure $P \in \mathcal{P}$ that *satisfies* the CI statements in $\mathcal{C}$ also *satisfies* the CI statement c. The set $\{c \mid \mathcal{C} \models c\}$ will be denoted by $\mathcal{C}^*$.

A powerful tool in deriving results about the CI implication problem is the association of semi-lattices with CI statements (Niepert et al., 2008). Given subsets A and B of S we write $[A, B]$ for the lattice $\{U \mid A \subseteq U \subseteq B\}$.

Definition 5. Let $I(A, B|C)$ be a CI statement. The *semi-lattice* of $I(A, B|C)$ is defined by $\mathcal{L}(A, B|C) = [C, S] - ([A, S] \cup [B, S])$.

Example 1. Let $S = \{a, b, c, d\}$ and let $I(a, b|c)$ be a CI statement. The semi-lattice of this statement is $\{c, cd\}$.

We will often write $\mathcal{L}(c)$ to denote the semi-lattice of a CI statement c and $\mathcal{L}(\mathcal{C})$ to denote the union of semi-lattices, $\bigcup_{c' \in \mathcal{C}} \mathcal{L}(c')$, of a set of CI statements $\mathcal{C}$.

3 Stable Conditional Independence

When novel information is available to a probabilistic system, the set of associated, relevant CI statements changes dynamically. However, some of the CI statements will continue to hold. Stable CI can be thought of as a subclass of general CI: every set of stable CI statements is a set of CI statements. Some of the properties of stable CI were first investigated by Matúš (Matúš, 1992) who named it *ascending* conditional independence and later by de Waal and van der Gaag (de Waal and van der Gaag, 2004) who introduced the term *stable* conditional independence. Every set of CI statements can be partitioned into its *stable* and *unstable* part. In this section we recall an axiomatization of stable CI using inference rules and its relation to the *lattice-inclusion* property. We will use these results to show that stable CI has perfect models w.r.t. discrete probability measures, but not w.r.t. binary probability measures.

Definition 6. Let $\mathcal{C}$ be a set of CI statements, and let $\mathcal{C}^{SG+}$ be the semi-graphoid closure of $\mathcal{C}$. Then $I(A, B|C)$ is said to be *stable* in $\mathcal{C}$ if $I(A, B|C') \in \mathcal{C}^{SG+}$ for all sets C' with $C \subseteq C' \subseteq S$.

Definition 7. A stable CI structure is a set of stable conditional independence statements $\mathcal{C}$ such that $\mathcal{C} = \mathcal{C}^*$.

In the remainder of the paper, a set of stable CI statements will be *any* set of CI statements that are implicitly *known* to be stable. Hence, a set of stable CI statements $\mathcal{C}$ can be different from $\mathcal{C}^*$. We approach stable CI as a *structural representation* of conditional independence much like graphical models are possible representations of conditional independence. Now, let us turn to a crucial result for stable conditional independence. The inference system $\mathcal{A}$ was shown to be sound and complete for stable conditional independence (Niepert et al., 2008).

Theorem 1. *Let $\mathcal{C}$ be a set of stable CI statements and let c be a CI statement. Then the following statements are equivalent:*

(a) $\mathcal{C} \models c$;

(b) $\mathcal{C} \vdash c$; and

(c) $\mathcal{L}(\mathcal{C}) \supseteq \mathcal{L}(c)$.

Example 2. Let $S = \{a, b, d, e\}$, let $\mathcal{C} = \{I(a, b|\emptyset), I(d, e|a), I(d, e|b)\}$ be a set of stable CI statements, and let $c = I(d, e|\emptyset)$. We know by *strong contraction* that $\mathcal{C} \vdash c$ and, therefore, $\mathcal{C} \models c$ by Theorem 1. Now, $\mathcal{L}(\mathcal{C}) = \{\emptyset, d, e, de\} \cup \{a, ab\} \cup \{b, ab\} = \{\emptyset, a, b, d, e, ab, de\} \supseteq \{\emptyset, a, b, ab\} = \mathcal{L}(c)$.

Definition 8. Let $\mathcal{C}$ be a set of CI statements. A probability measure is a *perfect model* for $\mathcal{C}$ if it satisfies precisely the statements $\mathcal{C}^*$, that is, all the statements that are implied by $\mathcal{C}$ and none other.

The next result follows from the existence of discrete perfect models with respect to CI statements (Geiger and Pearl, 1993), a result which was later strengthened by (Peña et al., 2006).

Proposition 1. *For every set of stable CI statements $\mathcal{C}$ there exists a discrete probability measure P such that P satisfies exactly the statements in $\mathcal{C}^*$ and none other, that is, P is a perfect model for $\mathcal{C}$.*

The previous result does not hold for the class of binary probability measures and it follows that stable CI shares the perfect model properties with general CI.

Proposition 2. *There exists a set of stable CI statements $\mathcal{C}$ for which no binary probability model is perfect.*

Proof. Let $S = \{a, b, c\}$ and let $\mathcal{C} = \{I(a, b|\emptyset), I(a, b|c)\}$. Clearly, $\mathcal{C}$ is a set of stable CI statements. By Theorem 1(c) neither $I(a, c|\emptyset)$ nor $I(b, c|\emptyset)$ are implied by $\mathcal{C}$. From (Geiger and Pearl, 1993) we know that every binary probability measure that satisfies the elements in $\mathcal{C}$ also satisfies either $I(a, c|\emptyset)$ or $I(b, c|\emptyset)$. Thus, no binary probability measure is perfect for $\mathcal{C}$. □

4 Graphical Models and Stable CI

Our goal is to relate stable CI to graphical models and more specifically undirected graphical models. Ultimately, we will show that stable CI can be seen as a generalization of undirected graphical models. The following theorem establishes that the CI statements present in a Markov network form a stable CI structure.

Theorem 2. *Let G be a Markov network (i.e., an undirected graphical model) and let $\mathcal{C}(G)$ be the set of all CI statements* encoded *in G. Then $\mathcal{C}(G)$ is a stable CI structure.*

Proof. It is well-known that *strong union* is a sound inference rule for separation in undirected graphs (Pearl, 1988). In addition, it can be verified that the inference rule *strong contraction* is sound for undirected graph separation. Thus, inference system $\mathcal{A}$ is sound for separation in undirected graphs and the statement of the theorem follows. □

Corollary 1. *For every Markov network G there exists a stable CI structure $\mathcal{C}$ and every discrete probability measure that is (perfectly) Markovian w.r.t. G satisfies the elements in $\mathcal{C}$ (and none other).*

This shows that stable conditional independence can be interpreted as a generalization of Markov networks. In what follows, we investigate how much broader this representation is compared to graphical models in general. First,

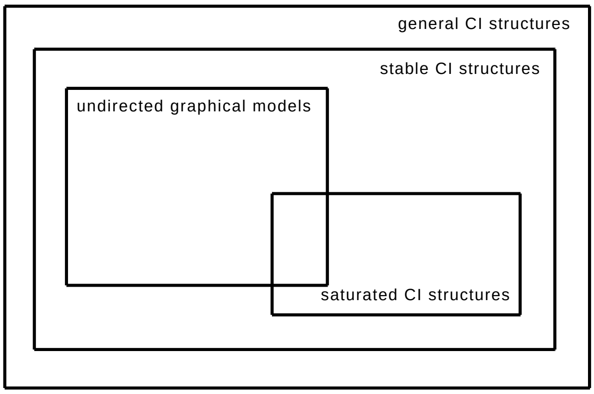

Figure 2: Inclusion relationships between different representations of conditional independence. Every undirected graphical model is a stable CI structure. Every saturated CI statements is trivially a stable CI statement.

we provide an example which demonstrates that there exists a stable CI structure that cannot be represented with a Markov network.

Example 3. Let $S = \{a, b, c, d\}$ and let $\mathcal{C} = \{I(a, b|cd), I(a, d|bc)\}$ be a set of stable CI statements. Note that by Theorem 1(c) no other CI statements are implied by $\mathcal{C}$ and hence, $\mathcal{C}$ is a stable CI structure. However, every Markov network that represents these two CI statements also represents the CI statement $I(a, bd|c)$ by the inference rule *intersection* which is sound for separation in undirected graphs (Pearl, 1988). Thus, the class of all CI structures induced by the class of Markov networks is a strict subclass of the class of stable CI structures.

Figure 2 depicts some relationships between different representations of conditional independence.

Proposition 3. *Let S be a finite set and let $x_i = \binom{|S|}{i}\binom{i}{2}$. The number of distinct stable CI structures over S is at least*

$$d_S =_{def} \sum_{i=2}^{|S|} (2^{x_i} - 1).$$

Proof. We sketch the proof. Let S be a finite set, let $V \subseteq S$ with $|V| = |S| - 2$, and let $U \subseteq V$. For every lattice $[U, V]$ there exists a stable CI structure $\mathcal{C}$ such that $\mathcal{L}(\mathcal{C}) = [U, V]$. Let $\ell = |S| - i$ for $2 \leq i \leq |S|$. Now, we have $2^{\binom{|S|}{i}\binom{i}{2}} - 1$ distinct combinations of lattices of the form $[U, V]$ with $|U| = \ell$ and each of these combinations represents a distinct stable CI structure by Theorem 1. □

Example 4. For $|S| = 3$ there are 8 UG, 22 discrete (Studený, 2005), and 14 stable discrete CI structures. For $|S| = 4$ there are 64 UG (Studený, 2005), 18478 discrete (Šimeček, 2006), and at least 4221 distinct stable CI structures. For $|S| = 5$ there are at least 2147485692 distinct stable CI structures, which is also a lower bound for the number of discrete CI structures.

As a consequence of Proposition 3 the number of stable CI structures grows double exponentially with the size of S.

5 Complexity of the Stable CI Implication Problem

In this section we will investigate the computational complexity of an important decision problems related to stable CI. Given a set of stable CI statements $\mathcal{C}$ and a CI statement c. Decide whether c is implied by $\mathcal{C}$. We will prove this decision problem to be `coNP`-complete. However, we will later show that a simple reduction to `UNSAT` exists. This allows one to make use of the many available `SAT` solvers and we will show experimentally that the problem can be decided very efficiently, even for instances involving hundreds of variables. We start with the formal definition of the decision problem.

Definition 9. Let $\mathcal{C}$ be a set of stable CI statements and let c be a CI statement. `STABLE-IMPLICATION` is the problem of deciding whether c is implied by $\mathcal{C}$, or, equivalently, whether the statement $\mathcal{C} \models c$ holds.

Lemma 1. `STABLE-IMPLICATION` *is in* `coNP`.

Proof. We show that the complement is in `NP`. Since $\mathcal{C} \not\models c$ if and only if $\mathcal{L}(\mathcal{C}) \not\supseteq \mathcal{L}(c)$ it is sufficient to find a $U \in \mathcal{L}(c)$ with $U \notin \mathcal{L}(\mathcal{C})$. This set can be guessed and then verified in polynomial time by checking for all $I(A, B|C) \in \mathcal{C}$ if $(U \not\supseteq C) \vee (U \supseteq A) \vee (U \supseteq B)$. □

We will now establish the correspondence between sets of stable CI statements and propositional formulae in conjunctive normal form, where a set of stable CI statement corresponds to a clause in the `CNF` formula and vice versa.

Definition 10. `3-CNFV` is the set of all propositional formulae in conjunctive normal form with clauses of the form $x \vee y$, $\neg x \vee y \vee z$, $\neg x \vee \neg y \vee z$, and $\neg x \vee \neg y \vee \neg z$.

Proposition 4. *Let T be a set of propositional variables and let $\Phi \in$ `3-CNFV`(T). Deciding whether Φ is satisfiable is* `NP`*-complete.*

Proof. This can be verified by a reduction from standard 3-`CNF-SAT`: the set of clauses we use in our construction are the clauses that occur in standard 3-`CNF` formulae except that every clause $x \vee y \vee z$ will be replaced by $(x \vee y \vee \neg w) \wedge (z \vee w)$, where w is a new variable. This reduction is possible in polynomial time and preserves satisfiability. □

Corollary 2. *Let T be a set of propositional variables and let $\Phi \in$ `3-CNFV`(T). Deciding whether Φ is a contradiction is* `coNP`*-complete.*

Definition 11. Let T be a set of propositional variables and let X be a subset of T. The *minterm* associated with X, denoted $\mathbf{X}$, is the formula $\bigwedge_{a \in X} a \wedge \bigwedge_{b \in \overline{X}} \neg b$. Let Φ be a propositional formula over T. The *minset* of Φ, denoted *minset*(Φ), is the set $\{X \mid \mathbf{X} \models_{prop} \Phi\}$ where $\models_{prop}$ is the logical implication relation for propositional logic. The negative minset of Φ, denoted *negminset*(Φ), is the set *minset*$(\neg\Phi)$.

Definition 12. Let $T = \{t_1, ..., t_n\}$ be a set of propositional variables, let $\Phi \in$ `3-CNFV`(T), let $\mathcal{C}(\Phi)$ be the set of clauses in Φ, let $S = T \cup \{r, s\}$ with $r \notin T$ and $s \notin T$, and let $\mathcal{T}(S)$ be the set of all non-trivial CI statements over S. Then $f :$ `3-CNFV`$(T) \rightarrow 2^{\mathcal{T}(S)}$ is defined as follows:

- $f(\Phi) = \bigcup_{c \in \mathcal{C}(\Phi)} f(c)$; with
- $f(t_i) = \{I(t_i, x|\emptyset) \mid x \in S - \{t_i\}\}$
- $f(\neg t_i) = \{I(x, y|t_i) \mid x, y \in S - \{t_i\}, x \neq y\}$
- $f(t_i \vee t_j) = \{I(t_i, t_j|\emptyset)\}$
- $f(\neg t_i \vee t_j) = \{I(t_j, x|t_i) \mid x \in S - \{t_i, t_j\}\}$
- $f(\neg t_i \vee \neg t_j) = \{I(x, y|\{t_i, t_j\}) \mid x, y \in S - \{t_i, t_j\}, x \neq y\}$
- $f(\neg t_i \vee t_j \vee t_k) = \{I(t_j, t_k|t_i)\}$
- $f(\neg t_i \vee \neg t_j \vee t_k) = \{I(t_k, x|\{t_i, t_j\}) \mid x \in S - \{t_i, t_j, t_k\}\}$
- $f(\neg t_i \vee \neg t_j \vee \neg t_k) = \{I(x, y|\{t_i, t_j, t_k\}) \mid x, y \in S - \{t_i, t_j, t_k\}, x \neq y\}$

Notice that the mapping f can be computed in polynomial time in the size of Φ and the number of variables involved. Furthermore, note that for any clause $c \in \mathcal{C}(\Phi)$ and for any $U \subseteq T$ we have $U \in \mathcal{L}(f(c))$ if and only if $\mathbf{U} \models_{prop} \neg c$.

Example 5. Let $T = \{a, b, c\}$, let $S = T \cup \{d, e\}$, and let $\Phi = (a \vee c) \wedge (\neg a \vee \neg b \vee c)$. Then $f(\Phi) = f(a \vee c) \cup f(\neg a \vee \neg b \vee c) = \{I(a, c|\emptyset)\} \cup \{I(c, d|ab), I(c, e|ab)\} = \{I(a, c|\emptyset), I(c, d|ab), I(c, e|ab)\}$ with $\mathcal{L}(f(\Phi)) = \{\emptyset, b, d, e, bd, be, bde, ab, abd, abe\}$ and *negminset*$(\Phi) = \{\emptyset, b, ab\}$.

Lemma 2. *Let T be a set of propositional variables, let $S = T \cup \{r, s\}$ with $r \notin T$, $s \notin T$, let f be the function from Definition 12, and let $\Phi \in$ 3-CNFV(T). Then we have the following:*

(1) negminset$(\Phi) \subseteq \mathcal{L}(f(\Phi))$; and

(2) Φ is a contradiction if and only if $\mathcal{L}(I(r, s|\emptyset)) \subseteq \mathcal{L}(f(\Phi))$.

Proof. To show (1) let $U \in$ *negminset*(Φ). Then there exists a clause c in $\mathcal{C}(\Phi)$ such that $\mathbf{U} \models_{prop} \neg c$. But then for $I(x, y|U') \in f(c)$ it must be $U \supseteq U'$, $x \notin U$ and $y \notin U$ since otherwise $\mathbf{U} \models_{prop} c$. It follows that $U \in \mathcal{L}(f(c))$ and therefore $U \in \mathcal{L}(f(\Phi))$.

To show (2) let Φ be a contradiction. Notice that Φ is a contradiction if and only if *negminset*$(\Phi) = 2^T$. Now, $\mathcal{L}(I(r, s|\emptyset)) = 2^T =$ *negminset*$(\Phi) \subseteq \mathcal{L}(f(\Phi))$, where the last inclusion follows from (1).

To show the other direction of (2) let $\mathcal{L}(I(r, s|\emptyset)) = 2^T \subseteq \mathcal{L}(f(\Phi))$. Assume that Φ is not a contradiction. Then there exists a set $U \subseteq T$ with $U \notin$ *negminset*(Φ). Now, since $2^T \subseteq \mathcal{L}(f(\Phi))$ there must be a clause $c \in \mathcal{C}(\Phi)$ such that $U \in \mathcal{L}(f(c))$. Hence, $\mathbf{U} \models_{prop} \neg c$ and

Property	Stable CI
Complete finite axiomatization	Yes
Implication algorithm	coNP-complete
Perfect models [$\mathcal{P}$]	Yes
Perfect models [$\mathcal{B}$]	No

Figure 3: Summary of properties of stable CI.

thus $U \in negminset(\Phi)$, a contradiction to our assumption that $U \notin negminset(\Phi)$. □

Theorem 3. `STABLE-IMPLICATION` *is* `coNP`-*complete.*

Proof. Let T be a set of propositional variables, let $r \notin T$, $s \notin T$, and let $\Phi \in$ `3-CNFV`(T). Then, by Lemma 2 and Theorem 1, Φ is a contradiction if and only if $\mathcal{L}(f(\Phi)) \supseteq \mathcal{L}(I(r, s|\emptyset))$ if and only if $f(\Phi) \vdash I(r, s|\emptyset)$ if and only if $f(\Phi) \models I(r, s|\emptyset)$, where f is computable in polynomial time. Hence, `STABLE-IMPLICATION` is `coNP`-hard. The statement now follows from Lemma 1. □

The logical and algorithmic properties of stable CI are summarized in Figure 3.

6 Implication Testing and Redundancy Elimination Using SAT Solvers

In this section we will show that every set of stable CI statements can be reduced to a propositional formula. This allows us to employ SAT solvers to decide the implication problem and to compute irredundant equivalent subsets of stable CI structures. Stable CI can considerably reduce the size of representation of CI structures (de Waal and van der Gaag, 2004). First, we will define the notion of *irredundancy* and *redundancy* of representation for sets of stable CI statements. We will use terminology that was previously introduced in the context of propositional formulae in conjunctive normal form (Liberatore, 2005).

Definition 13. A set of stable CI statements $\mathcal{C}$ is *irredundant* if and only if $\mathcal{C} - \{c\} \not\models c$ for all $c \in \mathcal{C}$. Otherwise it is *redundant.*

A related definition is that of an irredundant equivalent subset. Note that a set of stable CI statements may have several different irredundant equivalent subsets and that the cardinality of these sets can differ.

Definition 14. Let $\mathcal{C}$ be a set of stable CI statements. A set of stable CI statements $\mathcal{C}'$ is an *irredundant equivalent subset* of $\mathcal{C}$ if and only if:

1. $\mathcal{C}' \subseteq \mathcal{C}$;
2. $\mathcal{C}' \models c$ for all $c \in \mathcal{C}$; and
3. $\mathcal{C}'$ is irredundant.

Example 6. Let $S = \{a, b, c\}$ and let $\mathcal{C} = \{I(a, b|\emptyset), I(a, b|c)\}$. Then, $\mathcal{C}' = \{I(a, b|\emptyset)\}$ is an irredundant equivalent subset of $\mathcal{C}$.

By Theorem 1 a stable CI structure can be derived from each of its irredundant equivalent subsets using the inference rules of system $\mathcal{A}$.

Definition 15. Let S be a finite set, let $\mathcal{C}$ be a set of CI statements, and let $I(A, B|C)$ be a CI statement. The mapping $g : 2^{\mathcal{T}(S)} \rightarrow$ `CNF`(S) is defined as

- $g(\mathcal{C}) = \bigwedge_{c \in \mathcal{C}}(g(c))$; with
- $g(I(A, B|C)) = \bigwedge_{a \in A} a \vee \bigwedge_{b \in B} b \vee \bigvee_{c \in C} \neg c$

The mapping g can be computed in linear time in the size of $\mathcal{C}$. Now, based on this mapping we can state the following theorem.

Theorem 4. *Let $\mathcal{C}$ be a set of stable CI statements and let c be a CI statement. Then the following statements are equivalent:*

- *$\mathcal{C} \models c$; and*
- *$g(\mathcal{C}) \models_{prop} g(c)$.*

Proof. We will again use the concepts *minset* and *negminset* introduced in Definition 11. Let $\mathcal{C}$ be a set of CI statements and let c be a CI statement. One can verify that $\mathcal{L}(\mathcal{C}) = negminset(g(\mathcal{C}))$ and $\mathcal{L}(c) = negminset(g(c))$.

irredundant-subset ($\mathcal{C}$: set) $\mathcal{C}'$: set

$\mathcal{C}' := \mathcal{C}$
for each $c \in \mathcal{C}'$
  begin
    if $g(\mathcal{C}' - \{c\}) \wedge \neg g(c)$ *not* satisfiable
      then $\mathcal{C}' := \mathcal{C}' - \{c\}$
  end
return $\mathcal{C}'$

Figure 4: A function to compute an irredundant equivalent subset.

By Theorem 1 we have that if $\mathcal{C}$ is a set of stable CI statements, then $\mathcal{C} \models c$ if and only if $\mathcal{L}(\mathcal{C}) \supseteq \mathcal{L}(c)$. Now, $\mathcal{L}(\mathcal{C}) \supseteq \mathcal{L}(c)$ if and only if $negminset(g(\mathcal{C})) \supseteq negminset(g(c))$ if and only if $g(\mathcal{C}) \models_{prop} g(c)$. □

Example 7. Let $S = \{a, b, d, e\}$, let $\mathcal{C} = \{I(a, b|\emptyset), I(d, e|a), I(d, e|b)\}$, and let $c = I(d, e|\emptyset)$. We have $g(\mathcal{C}) = (a \vee b) \wedge (d \vee e \vee \neg a) \wedge (d \vee e \vee \neg b)$ and $g(c) = d \vee e$. We also have $g(\mathcal{C}) \models_{prop} g(c)$ if and only if $g(\mathcal{C}) \wedge \neg g(c)$ is not satisfiable. Now, $g(\mathcal{C}) \wedge \neg g(c) = (a \vee b) \wedge (d \vee e \vee \neg a) \wedge (d \vee e \vee \neg b) \wedge \neg d \wedge \neg e$. This formula is not satisfiable. Hence, $\mathcal{C} \models c$ by Theorem 4.

Corollary 3. *Let $\mathcal{C}$ be a set of stable CI statements. Then $\mathcal{C}$ is irredundant if and only if for all c in $\mathcal{C}$ we have that $g(\mathcal{C} - \{c\}) \wedge \neg g(c)$ is satisfiable.*

The algorithm in Figure 4 is based on Corollary 3. It takes as input a set of stable CI statements $\mathcal{C}$ and returns an irredundant equivalent subset of $\mathcal{C}$ based on several satisfiability tests. For each number of attributes from 5 to 25 we randomly created sets of 500 CI statements and determined the size of the irredundant equivalent subsets using the algorithm. Figure 5 shows the average size of 1000 different runs. As one can expect, the fewer attributes there are the smaller is the irredundant equivalent subset.

The performance of the SAT solvers applied to instances of the implication problem was quite remarkable. We used MiniSat[1] by Niklas Eén and Niklas Sörensson on a Pentium4 dual-core Linux system for the experiments. For the 500 satisfiability tests made to compute an irredundant equivalent subset, the algorithm took at most 1100 ms, where the majority of the time was spent on unsatisfiable instances of the problem. This amounts on average to 2ms per satisfiability test for sets of 500 CI statements.

[1]http://minisat.se

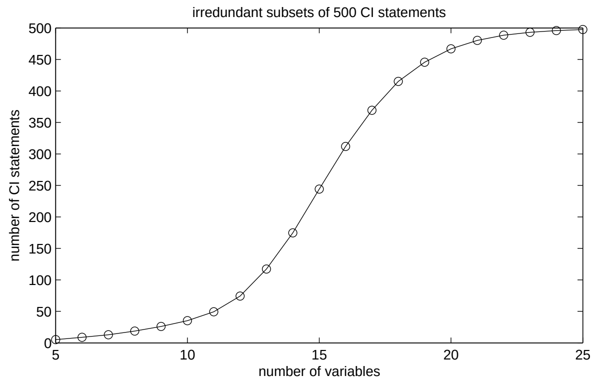

Figure 5: Size of irredundant equivalent subset of a set of initially 500 CI statements for different numbers of attributes.

variables	50	100	200	300	400
time [ms]	740	1523	3362	5627	7076

Figure 6: Average time needed (in milliseconds) to decide the implication problem for different numbers of variables and 100,000 antecedents.

In a second experiment we applied the SAT solver to larger, randomly generated instances of the stable CI implication problem with up to 400 variables. Figure 6 shows the average time (out of 10 tests) needed to decide the implication problem $\mathcal{C} \models c$ for $|\mathcal{C}| = 100,000$ and different numbers of variables.

7 Discussion and Future Work

We used a finite complete axiomatization of stable conditional independence to show that stable CI has the same perfect model properties as general conditional independence. In addition, we proved that stable conditional independence can be interpreted as a generalization or extension of undirected graphical models in that the class of stable CI structures is a strict superset of the class of CI structures induced by undirected graphical models. Many procedures that learn graphical models are based on the *data faithfulness assumption*, see for example (Studený, 2005). The data faithfulness assumption states that data are "generated" by a probability measure P which is perfectly Markovian with respect to an instance of the class of graphical model under consideration. Now, learning methods based on these procedures are only safely applicable if the data faithfulness assumption is guaranteed.

While the data faithfulness assumption is also *not* guaranteed for the class of stable CI structures, we have as a consequence of Proposition 3 that the number of stable CI structures grows double exponentially with the size of S and, therefore, *more* probability measures are perfect with respect to a stable CI structure. On one hand, this implies that a reasonable graphical representation of stable CI is unlikely, using arguments similar to those made in (Studený, 2005) on page 63. On the other hand, it shows that the class of stable CI structures is the broadest and only double exponentially growing class of CI structures for which a complete finite axiomatization using inference rules and an implication algorithm are known. We also know that this class of CI structures includes the class of all CI structures induced by undirected graphical models and that there exists an interesting, direct connection to propositional logic. Furthermore, we have demonstrated that SAT solvers can be used to efficiently decide the implication problem for stable conditional independence, even for large numbers of variables. Future research should be concerned with the development of algorithms that can learn stable CI models from data and for *probabilistic* inference in the context of stable CI.

In addition to the aforementioned possible applications, stable CI can also be used as part of a probabilistic system to store information about conditional independencies more efficiently, using irredundant equivalent subsets computed by the algorithm in Figure 4.

Acknowledgments

We thank Marc Gyssens and Paul Purdom for helpful discussions and the anonymous reviewers for their valuable suggestions.

References

Peter R. de Waal and Linda C. van der Gaag. 2004. Stable independence and complexity of representation. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, pages 112–119.

Peter R. de Waal and Linda C. van der Gaag. 2005. Stable independence in perfect maps. In *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence*, pages 161–168.

Dan Geiger and Judea Pearl. 1993. Logical and algorithmic properties of conditional independence and graphical models. *The Annals of Statistics*, 21(4):2001–2021.

Paolo Liberatore. 2005. Redundancy in logic i: Cnf propositional formulae. *Artif. Intell.*, 163(2):203–232.

František Matúš. 1992. Ascending and descending conditional independence relations. In *Transactions of the 11th Prague Conference on Information Theory*, pages 189–200.

Mathias Niepert, Dirk Van Gucht, and Marc Gyssens. 2008. On the conditional independence implication problem: A lattice-theoretic approach. In *Proceedings of the 24th Conference on Uncertainty in Artificial Intelligence*, pages 435–443.

Judea Pearl. 1988. *Probabilistic reasoning in intelligent systems: networks of plausible inference.* Morgan Kaufmann Publishers Inc.

Jose Peña, Roland Nilsson, Johan Björkegren, and Jesper Tegnér. 2006. Reading dependencies from the minimal undirected independence map of a graphoid that satisfies weak transitivity. In *Proccedings of the European Workshop on Probabilistic Graphical Models*, pages 247–254.

Milan Studený. 2005. *Probabilistic Conditional Independence Structures.* Springer-Verlag.

Petr Šimeček. 2006. A short note on discrete representability of independence models. In *Procceedings of the European Workshop on Probabilistic Graphical Models*, pages 287–292.