On the Conditional Independence Implication Problem: A Lattice-Theoretic Approach

Mathias Niepert[1] Dirk Van Gucht[1] Marc Gyssens[2]

[1]Department of Computer Science
Indiana University, USA

[2]Department WNI
Hasselt University & Transnational University of Limburg, Belgium

Outline

Outline

Conditional Independence Statements

- S implicit set of statistical variables
- We consider only *discrete* probability measures
- $A, B, C, \ldots$ subsets of S
- $I(A, B|C) \Leftrightarrow P(\mathbf{AC})P(\mathbf{BC}) = P(\mathbf{ABC})P(\mathbf{C})$
- $I(A, B|C)$ is *saturated*, if $A \cup B \cup C = S$
- $\mathcal{C}$ set of CI statements, c single CI statement
- **CI important concept in uncertain reasoning**

Definition (Probabilistic CI implication problem)

Let $\mathcal{C}$ be a set of CI statements and let c be a CI statement. We say that $\mathcal{C}$ *implies* c, and write $\mathcal{C} \models c$, if each discrete probability measure that *satisfies* the CI statements in $\mathcal{C}$ also *satisfies* the CI statement c.

The semi-graphoid axioms, Pearl 1988

$I(A, \emptyset \mid C)$	**Triviality**
$I(A, B \mid C) \rightarrow I(B, A \mid C)$	**Symmetry**
$I(A, BD \mid C) \rightarrow I(A, D \mid C)$	**Decomposition**
$I(A, B \mid CD) \wedge I(A, D \mid C) \rightarrow I(A, BD \mid C)$	**Contraction**
$I(A, BD \mid C) \rightarrow I(A, B \mid CD)$	**Weak union**

Definition (Probabilistic CI implication problem)

Let $\mathcal{C}$ be a set of CI statements and let c be a CI statement. We say that $\mathcal{C}$ *implies* c, and write $\mathcal{C} \models c$, if each discrete probability measure that *satisfies* the CI statements in $\mathcal{C}$ also *satisfies* the CI statement c.

The semi-graphoid axioms, Pearl 1988

$I(A, \emptyset \mid C)$	**Triviality**
$I(A, B \mid C) \rightarrow I(B, A \mid C)$	**Symmetry**
$I(A, BD \mid C) \rightarrow I(A, D \mid C)$	**Decomposition**
$I(A, B \mid CD) \wedge I(A, D \mid C) \rightarrow I(A, BD \mid C)$	**Contraction**
$I(A, BD \mid C) \rightarrow I(A, B \mid CD)$	**Weak union**

Semi-Lattices and CI Statements

- Semi-lattices are associated with CI statements
- $\mathcal{L}(I(A, B|C) = [C, S] - ([A, S] \cup [B, S])$
- $\mathcal{L}(\mathcal{C})$ union of semi-lattices of a set of CI statements

Example

$S = \{a, b, c, d\}$ and $\mathcal{C} = \{I(a, b|c), I(ab, d|\emptyset)\}$.

- $\mathcal{L}(I(a, b|c)) = \{c, cd\}$
- $\mathcal{L}(I(ab, d|\emptyset)) = \{\emptyset, c, ac, bc\}$
- $\mathcal{L}(\mathcal{C}) = \{\emptyset, c, ac, bc, cd\}$

Outline

Inference System $\mathcal{A}$

$I(A, \emptyset \mid C)$	**Triviality**
$I(A, B \mid C) \rightarrow I(B, A \mid C)$	**Symmetry**
$I(A, BD \mid C) \rightarrow I(A, D \mid C)$	**Decomposition**
$I(A, B \mid CD) \wedge I(A, D \mid C) \rightarrow I(A, BD \mid C)$	**Contraction**
$I(A, B \mid C) \rightarrow I(A, B \mid CD)$	**Strong union**
$I(A, B \mid C) \wedge I(D, E \mid AC) \wedge$ $I(D, E \mid BC) \rightarrow I(D, E \mid C)$	**Strong contraction**

Theorem

Let $\mathcal{C}$ be a set of CI statements, and let c be a CI statement. Then $\mathcal{C} \vdash c$ if and only if $\mathcal{L}(\mathcal{C}) \supseteq \mathcal{L}(c)$.

Inference System $\mathcal{A}$

$I(A, \emptyset \mid C)$	**Triviality**
$I(A, B \mid C) \rightarrow I(B, A \mid C)$	**Symmetry**
$I(A, BD \mid C) \rightarrow I(A, D \mid C)$	**Decomposition**
$I(A, B \mid CD) \wedge I(A, D \mid C) \rightarrow I(A, BD \mid C)$	**Contraction**
$I(A, B \mid C) \rightarrow I(A, B \mid CD)$	**Strong union**
$I(A, B \mid C) \wedge I(D, E \mid AC) \wedge I(D, E \mid BC) \rightarrow I(D, E \mid C)$	**Strong contraction**

Theorem

Let $\mathcal{C}$ be a set of CI statements, and let c be a CI statement. Then $\mathcal{C} \vdash c$ if and only if $\mathcal{L}(\mathcal{C}) \supseteq \mathcal{L}(c)$.

The Additive Implication Problem

We can link the multiplicative notion of conditional independence to an additive one (Studený 2005):

Definition

The *multi-information function* $M_P : 2^S \rightarrow [0, \infty]$ induced by P is defined as

$$M_P(A) := H(P^A | \prod_{a \in A} P^{\{a\}})$$

$$P(\mathbf{AC})P(\mathbf{BC}) = P(\mathbf{ABC})P(\mathbf{C})$$
$$\Leftrightarrow$$
$$M_P(AC) + M_P(BC) = M_P(ABC) + M_P(C)$$

The Additive Implication Problem

Question

Given some class of real-valued functions $\mathcal{F}$. What are neccessary and/or sufficient conditions under which inference system $\mathcal{A}$ is sound and/or complete for the additive implication problem?

(Some) Answers

- $\mathcal{A}$ is sound for *saturated* statements if and only if *weak union* is a sound inference rule
- $\mathcal{A}$ is sound if and only if *strong union* and *decomposition* are sound inference rules
- $\mathcal{A}$ is complete if $\mathcal{A}$ is sound and complete for *saturated* statements

The Additive Implication Problem

Question

Given some class of real-valued functions $\mathcal{F}$. What are neccessary and/or sufficient conditions under which inference system $\mathcal{A}$ is sound and/or complete for the additive implication problem?

(Some) Answers

- $\mathcal{A}$ is sound for *saturated* statements if and only if *weak union* is a sound inference rule
- $\mathcal{A}$ is sound if and only if *strong union* and *decomposition* are sound inference rules
- $\mathcal{A}$ is complete if $\mathcal{A}$ is sound and complete for *saturated* statements

The Probabilistic CI Implication Problem

These results follow directly for the **probabilistic CI implication problem**:

Theorem

$\mathcal{A}$ is sound and complete for saturated CI statements.

Theorem

$\mathcal{A}$ is sound and complete for stable CI statements.

Theorem

$\mathcal{A}$ is complete for CI statements.

The Probabilistic CI Implication Problem

These results follow directly for the **probabilistic CI implication problem**:

Theorem

$\mathcal{A}$ is sound and complete for saturated CI statements.

Theorem

$\mathcal{A}$ is sound and complete for stable CI statements.

Theorem

$\mathcal{A}$ is complete for CI statements.

The Probabilistic CI Implication Problem

These results follow directly for the **probabilistic CI implication problem**:

Theorem

$\mathcal{A}$ is sound and complete for saturated CI statements.

Theorem

$\mathcal{A}$ is sound and complete for stable CI statements.

Theorem

$\mathcal{A}$ is complete for CI statements.

The Probabilistic CI Implication Problem

These results follow directly for the **probabilistic CI implication problem**:

Theorem

$\mathcal{A}$ is sound and complete for saturated CI statements.

Theorem

$\mathcal{A}$ is sound and complete for stable CI statements.

Theorem

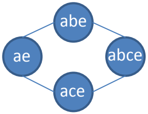
$\mathcal{A}$ is complete for CI statements.

Outline

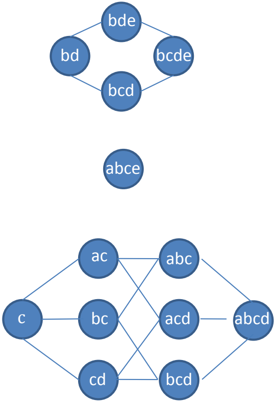
Falsification Algorithm

Corollary

Let $\mathcal{C}$ be a set of CI statements. If $\mathcal{L}(\mathcal{C}) \not\supseteq \mathcal{L}(c)$, then $\mathcal{C} \not\models c$.

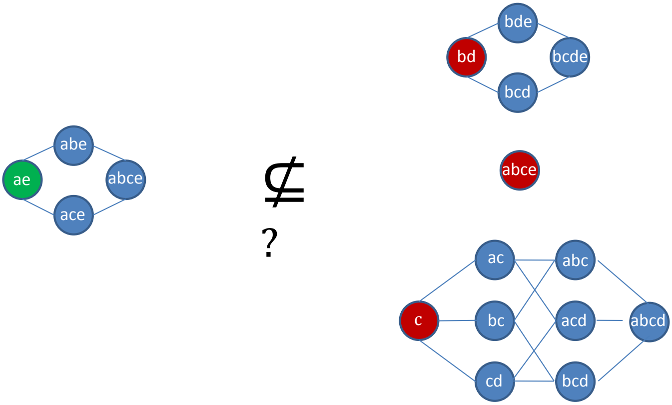


$\not\subseteq$
?



Heuristics

Heuristic 1 tests if the minimal element of $\mathcal{L}(c)$ is a superset of one of the minimal elements of the semi-lattices of $\mathcal{C}$.



Heuristics (cont.)

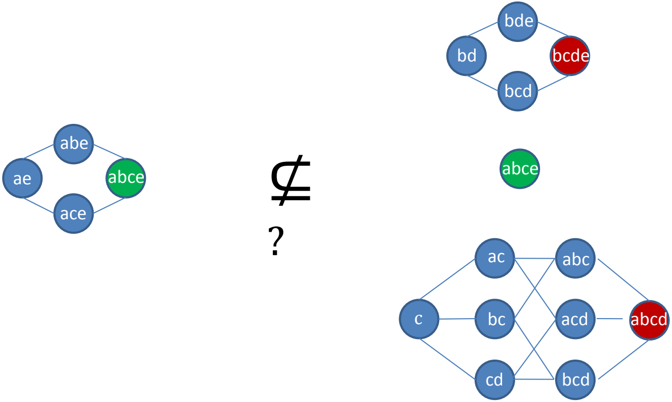
Example (Intersection)

$I(A, B|DC) \wedge I(A, D|BC) \rightarrow I(A, BD|C)$

Intersection is *not* a sound inference rule relative to the class of discrete probability measures. Heuristic 1 can reject this instance in polynomial time.

Heuristics (cont.)

Heuristic 2 tests if one of the maximal elements of the semi-lattice of c is a subset of one of the maximal elements of the semi-lattices of $\mathcal{C}$.



Experiments

- 77000 implication problems for sets with 3 antecedents, 76000 for sets with 4 antecedents, down to 70000 for sets with 10 antecedents
- Compared the falsification procedure to the racing algorithm (Bouckaert and Studený)
- Compared the heuristics to the falsification procedure

Results

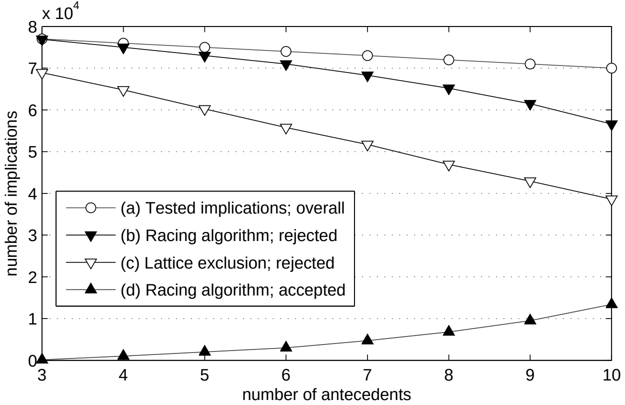


Figure: Rejection and acceptance curves of the racing and falsification algorithms, respectively, for five attributes.

Results

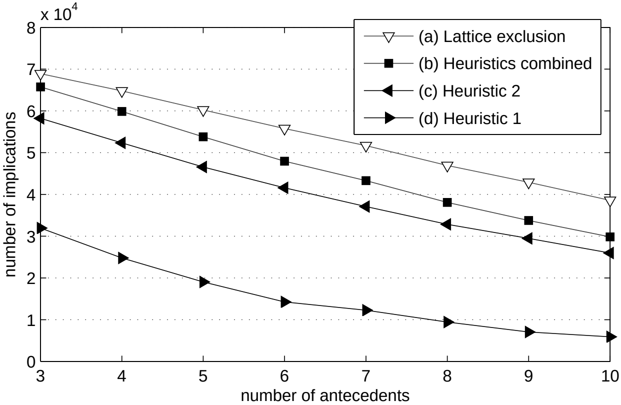


Figure: Falsifications based on the lattice-exclusion criterion and the heuristics, for five attributes.

Applications

- Can be used as part of a CI inference algorithm (e.g., the racing algorithm)
- Structure learning for Markov networks
  - Constrained-based learning (Margaritis & Thrun; Ghandi et al.)
  - Guided by “faithfulness" of the model
- Stable Independence as new structural representation (Niepert & Van Gucht PGM’08)
  - Generalization of Markov networks
  - Implication problem coNP-complete but “easy" for SAT solvers for thousands of variables
  - Compact representation of CI that is *more faithful* to the data

References I

- Dan Geiger and Judea Pearl.
  Logical and Algorithmic Properties of Conditional Independence and Graphical Models.
  The Annals of Statistics, 21(4):2001–2021, 1993.
- Remco R. Bouckaert and Milan Studený.
  Racing Algorithms for Conditional Independence Inference.
  Int. J. Approx. Reasoning, 45(2):386–401, 2007.
- Peter de Waal and Linda C. van der Gaag.
  Stable Independence and Complexity of Representation.
  Proceedings UAI, 112–119, 2004.