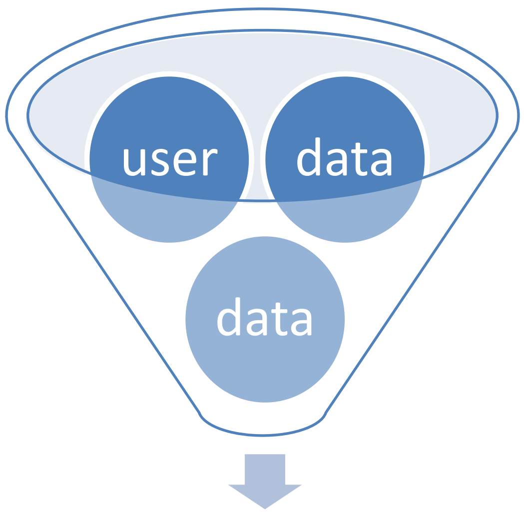
Logical Properties of Stable Conditional Independence

Mathias Niepert
Dirk Van Gucht

Department of Computer Science &
Program of Cognitive Science
Indiana University

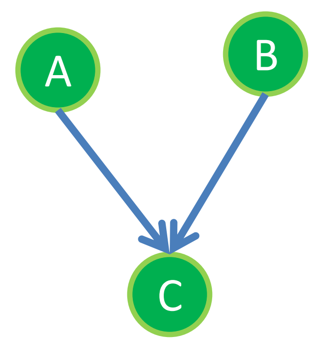
Representing and Reasoning about CI

P(A,B,C,D,E,...)



I(A, B|C), I(E, A|D), I(B, A|D), ...

Representational Choice



Learning

Yes/No

„Query“

I(E,B|D) ?

0.2346

„Query“

P(A=a,B=b) ?

Why Alternatives?

- Naive representation of CI possibly exponential in number of variables, but also
- Faithfulness problem with graphical models
- Try to find a representation which is
  - Broader (i.e, captures more probability measures) than graphical models but
  - Possibly without the exponential blow-up

What is Stable CI?

- Same as „general“ CI, except that strong union is *sound*
- I(A,B|C) → I(A,B|CD)   [strong union]
- Early work (Matus 92) called it ascending CI
- Every set of CI statements can be paritioned into stable and unstable part
- Used to reduce size of representation (de Waal & van der Gaag 04)

Logical Properties of Stable CI

- Conditional Independence Structures

Set of CI statements	C = { I(a,b|Ø), I(c,d|a), I(c,d|b) }
	$\models$
+ Implied CI statements	I(c,d|Ø)
= CI structure	C* = { I(a,b|Ø), I(c,d|a), I(c,d|b), I(c,d|Ø) }

Logical Properties of Stable CI

- Conditional Independence Structures

Set of CI statements $C = \{ I(a,b|Ø), I(c,d|a), I(c,d|b) \}$

$\models$?

+ Implied CI statements $I(c,d|Ø)$

= CI structure $C^* = \{ I(a,b|Ø), I(c,d|a), I(c,d|b), I(c,d|Ø) \}$

Logical Properties of Stable CI

Axiomatization of stable CI exists (Niepert et al. 08)

I(A,B|C) → I(B,A|C) [symmetry]

I(A,BD|C) → I(A,B|C) [decomposition]

I(A,B|CD) I(A,D|C) → I(A,BD|C) [contraction]

I(A,B|C) → I(A,B|CD) [strong union]

I(A,B|C), I(D,E|AC),I(D,E|BC) → I(D,E|C) [strong contraction]

Logical Properties of Stable CI

- For every set of stable CI statements C, there exists a discrete probability measure that satisfies the statements in C* and none other
- Not true for binary probability measures

→ Same situation as for „arbitrary“ conditional independence

Logical Properties of Stable CI

- Generalization of undirected graphical models (follows from axiomatization)
- Number of stable CI structures grows at least double exponentially with the number of variables
- For 4 variables there are **64** UG, 18478 general, and at least **4221** stable CI structures

Logical Properties of Stable CI

- Problem: How can we achieve „lossless compression“ of representation of stable CI?
- Solution in the computational complexity of the implication problem
- Implication problem for stable CI:
  - Given a set of stable CI statements C and a CI statement c. Decide if $C \models c$?
  - If so, c is a stable CI statement!

Redundancy and Irredundancy

- Given a set of stable CI statements C
- Irredundant equivalent subset C‘ (Liberatore)
  - Subset of C
  - For all $c \in C$: $C' \models c$
  - For all $c \in C'$: $(C' - \{c\}) \not\models c$
- If we can decide $\models$ efficiently, we could compute more compact and lossless represenatation

Computational Complexity

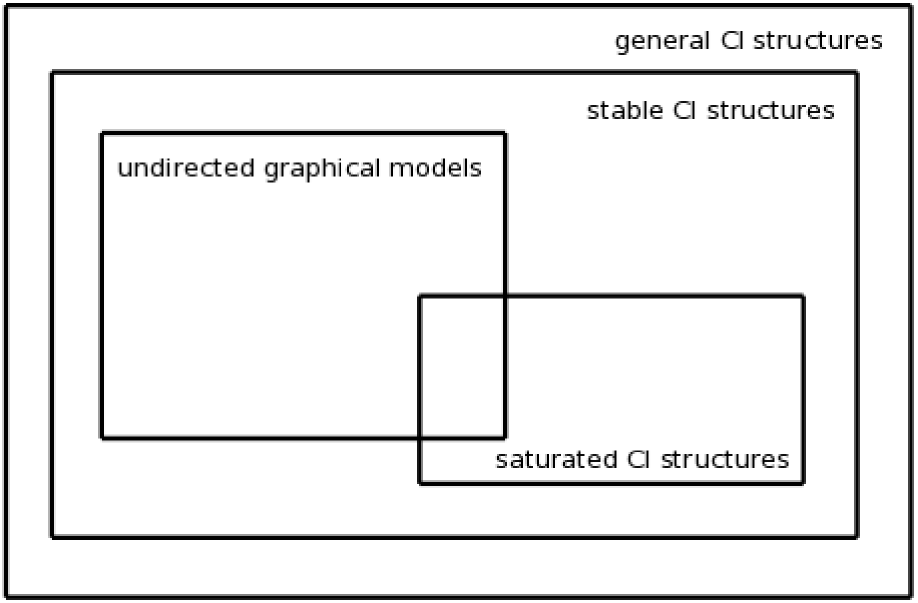
- Propositional formula ϕ in variant of 3-CNF over T
- Mapped into set C of CI statements over $T \cup \{r,s\}$
- ϕ is contradiction if and only if $C \models I(r,s|\emptyset)$
- ➔ Implication problem coNP-complete

$(a \vee c) \wedge (\neg a \vee \neg b \vee c)$ ➔ $\{I(a,c|\emptyset), I(c,r|ab), I(c,s|ab)\}$

$\{ I(a,c|\emptyset), I(c,r|ab), I(c,s|ab) \} \models I(r,s|\emptyset)$?

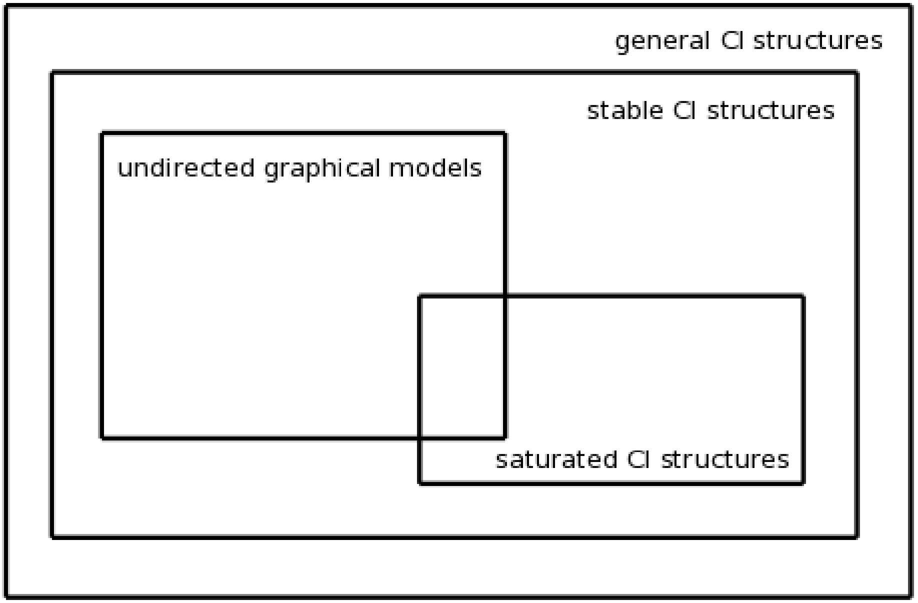
Logical Properties of Stable CI

Property	Stable CI
Complete finite axiomatization	Yes
Implication Problem	coNP-complete
Perfect models	Yes (but not for binary measures)



Logical Properties of Stable CI

Property	Stable CI
Complete finite axiomatization	Yes
Implication Problem	coNP-complete
Perfect models	Yes (but not for binary measures)



Logical Properties of Stable CI

- Mapping g from set of CI statements to propositional formula exists (linear time)
- SAT solvers can be used to decide the implication problem

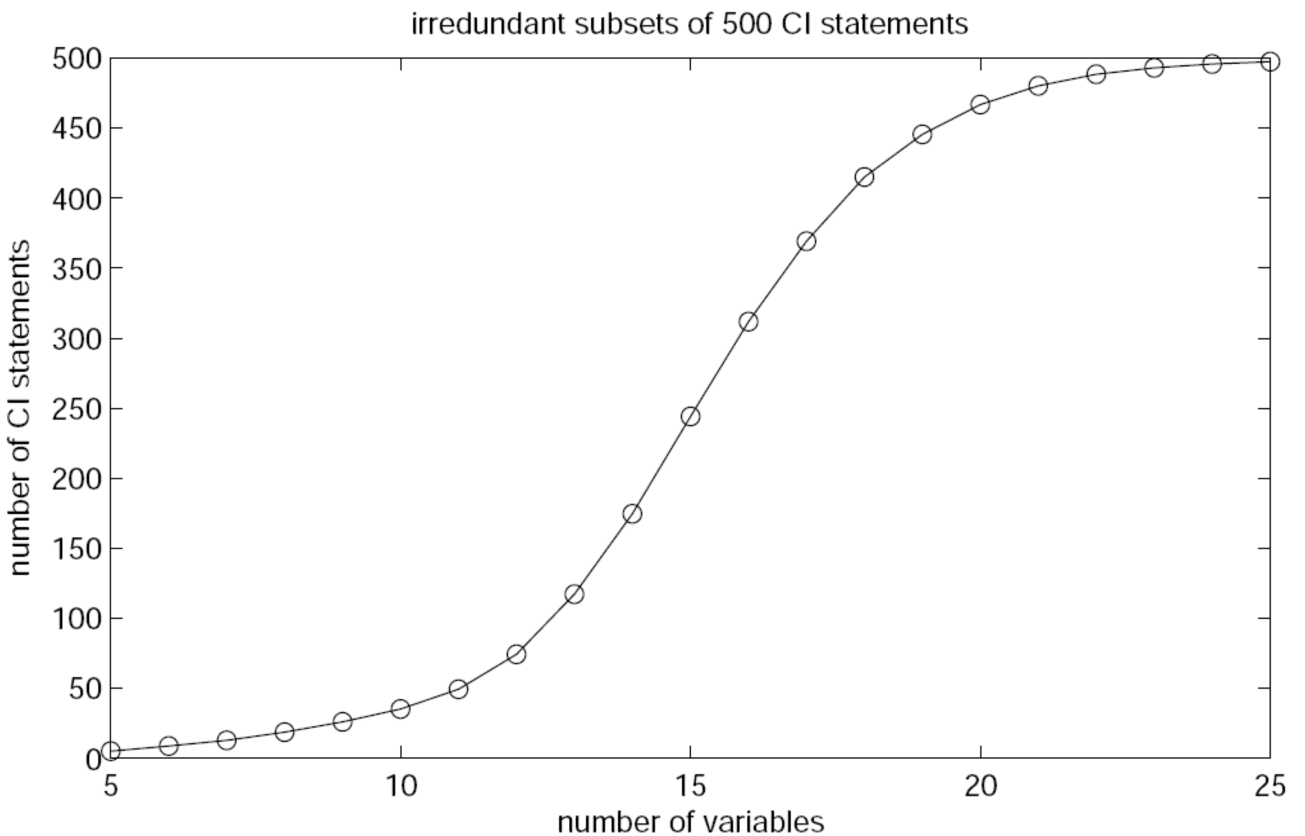
---

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

$\mathcal{C}' := \mathcal{C}$
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 $\mathcal{C}'$

Irredundant Representation

- Experiments
  - Initially 500 randomly generated CI statements



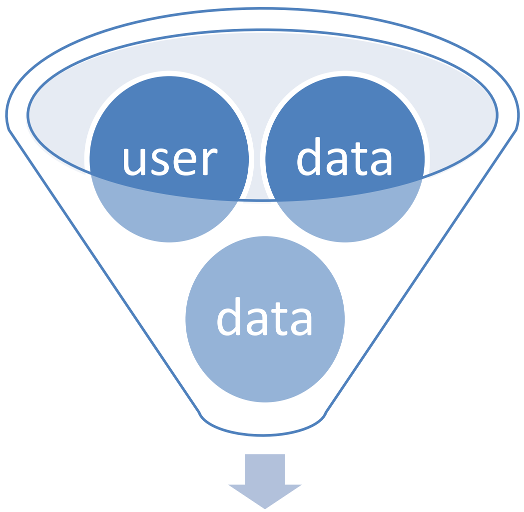
Irredundant Representation

- Experiments (cont.)
  - 100,000 antecedents
  - Randomly generated
  - Time (in ms) to decide the implication problem

Variables	50	100	200	300	400
Time [ms]	740	1523	3362	5627	7076

Representing and Reasoning about CI

P(A,B,C,D,E,...)



I(A, B|C), I(E, A|D), I(B, A|D), ...

SAT solver
>>100 variables

Create Stable CI representation using SAT solvers

Yes/No

„Query“

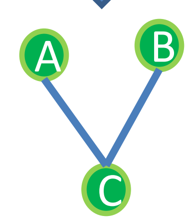
I(E,B|D) ?

0.2346

„Query“

P(A=**a**,B=**b**) ?

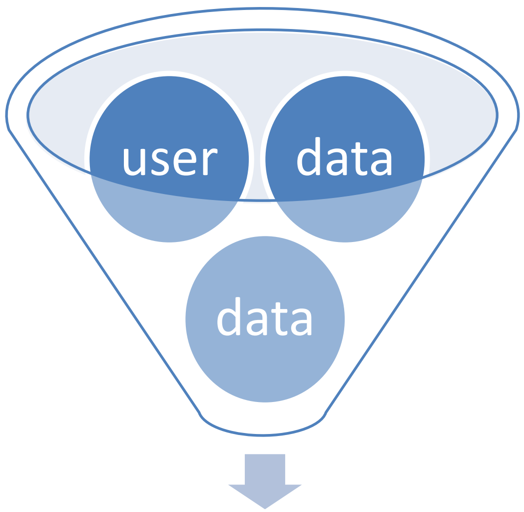
Irredundant Stable CI



„Unique“
UG Model

Representing and Reasoning about CI

P(A,B,C,D,E,...)



I(A, B|C), I(E, A|D), I(B, A|D), ...

Create Stable CI representation using SAT solvers

Irredundant Stable CI

Yes/No

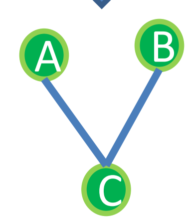
„Query“

I(E,B|D) ?

???? 0.2346

„Query“

P(A=**a**,B=**b**) ?



„Unique“ UG Model

Stable Conditional Independence

- Alternative for representing and reasoning about conditional independence
- Complete, finite axiomatization exists
- Perfect models for discrete prob. measures
- Generalization of UG models
- Correspondence to propositional logic
- SAT solver can decide the implication problem

Thank you!