# **Logic and Probability**

**The Computational Connection** 

Adnan Darwiche UCLA

**Deduction at Scale Seminar, March, 2011** 

## Inference

Probabilistic Graphical Models:

Marginal and conditional probabilities Most likely instantiations...

Propositional Knowledge Bases:

Logical entailment Existential quantification Model counting...

## **Two Main Themes**

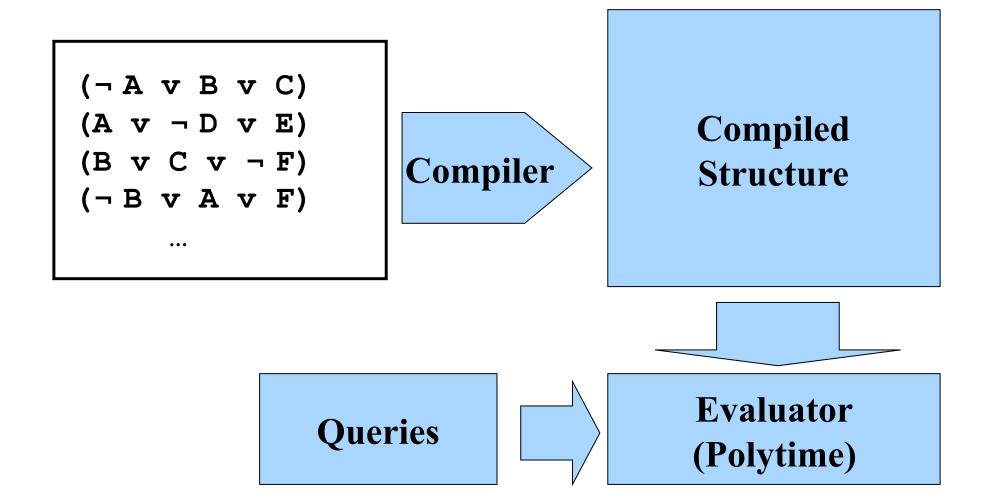
#### Exact inference as:

Enforcing decomposability and determinism on propositional knowledge bases

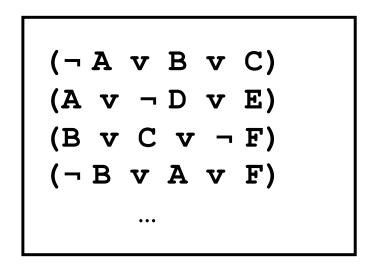
#### Approximate inference as:

Relaxing, compensating for, and recovering equivalence constraints (equalities)

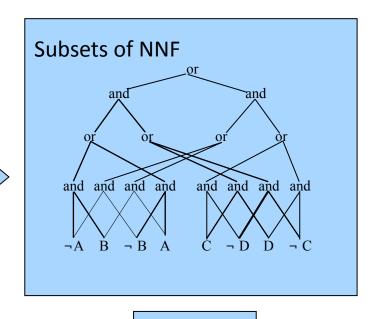
# **Knowledge Compilation**



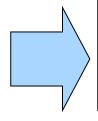
# **Knowledge Compilation**



Compiler

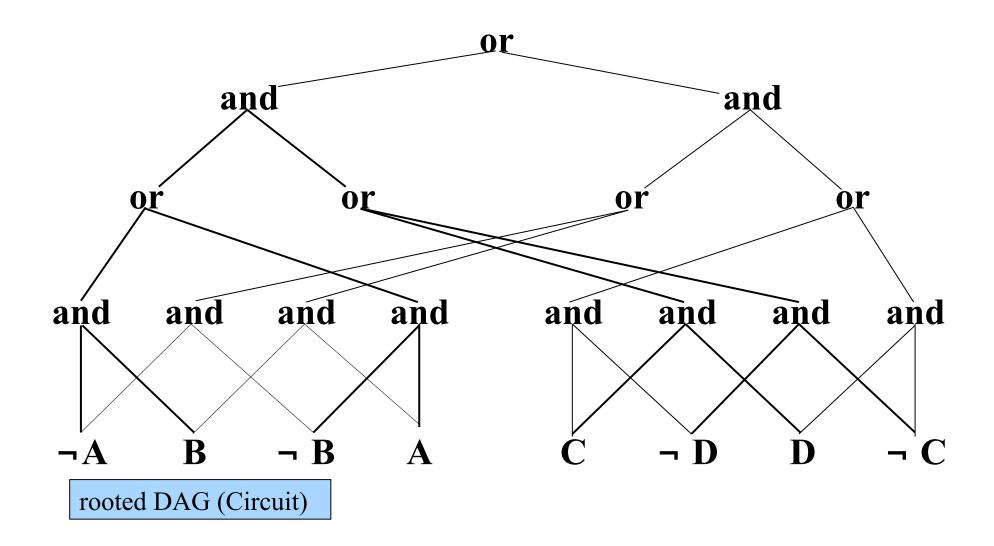


Queries

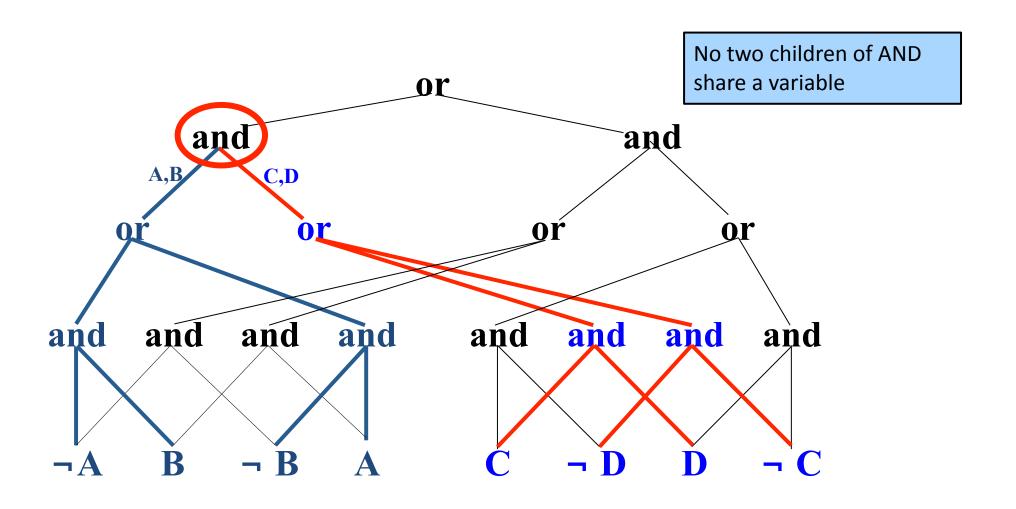


**Evaluator** (Polytime)

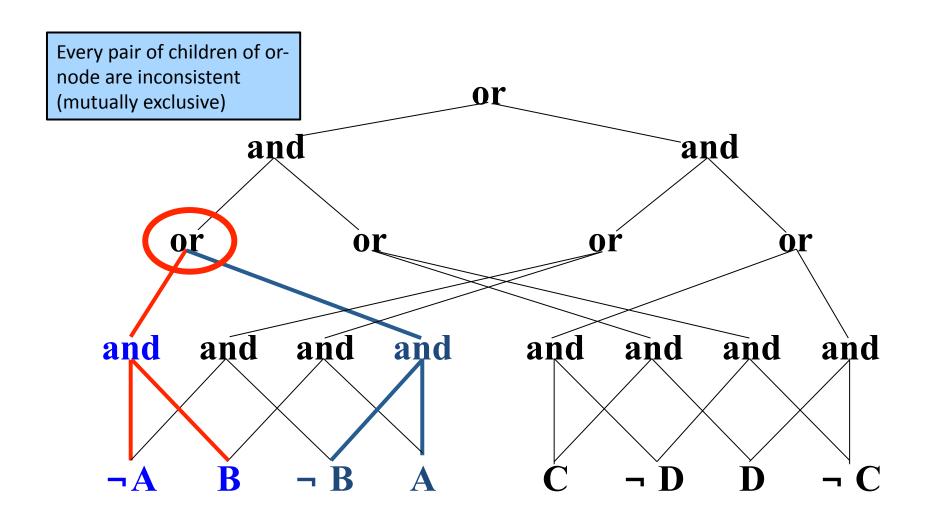
# **Negation Normal Form**



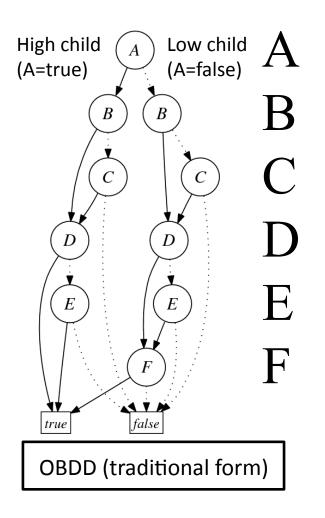
# **Decomposability (DNNF)**

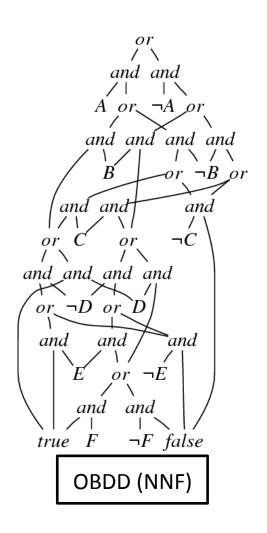


# **Determinism (d-DNNF)**



# **OBDD:**d-DNNF + Additional Properties





## **Queries and Transformations**

#### Queries

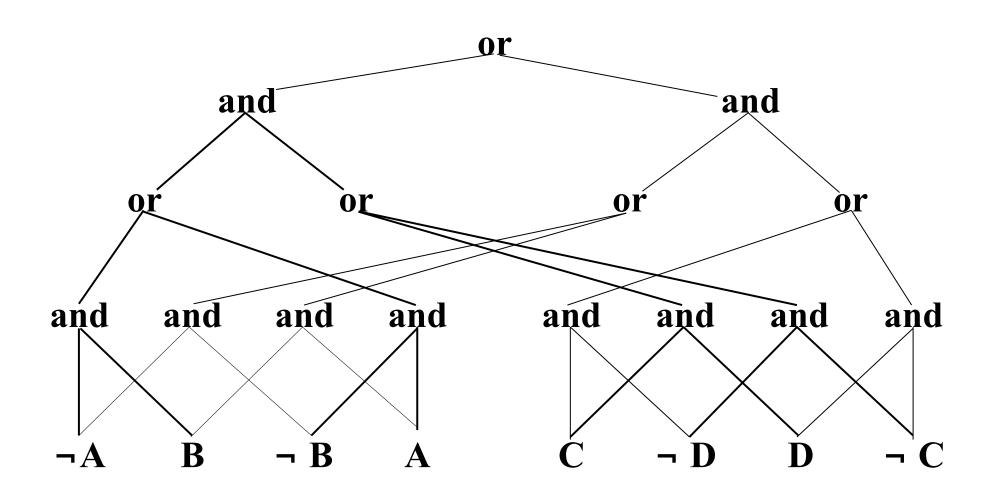
SAT, MAXSAT, logical entailment, equivalence testing, model counting,...

#### Transformations:

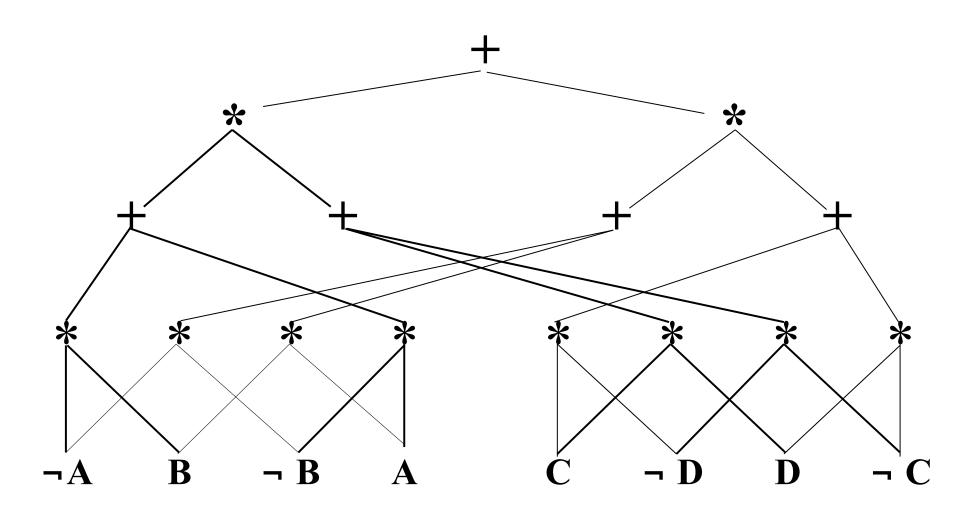
Existential quantification, conjunction, disjunction, negation...

 More properties imply more polytime queries and transformations, but less succinctness

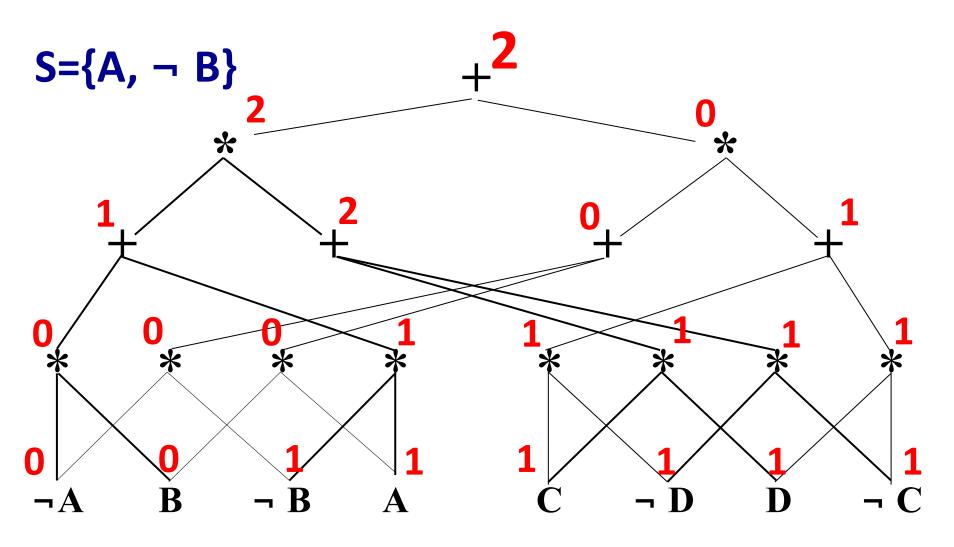
# Counting Models (d-DNNF)



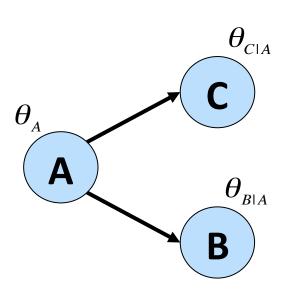
# **Counting Graph**



# **Counting Graph**



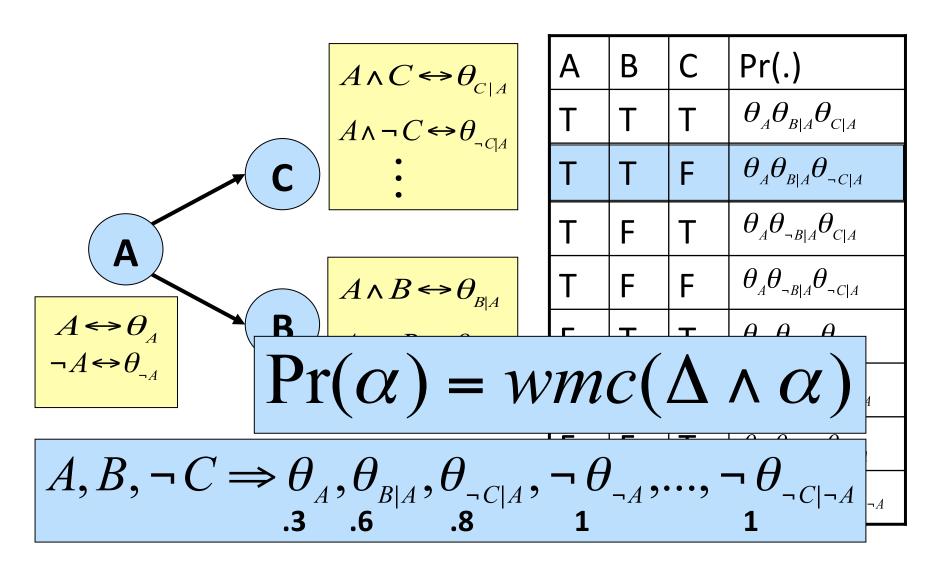
# Probabilistic Inference by Weighted Model Counting



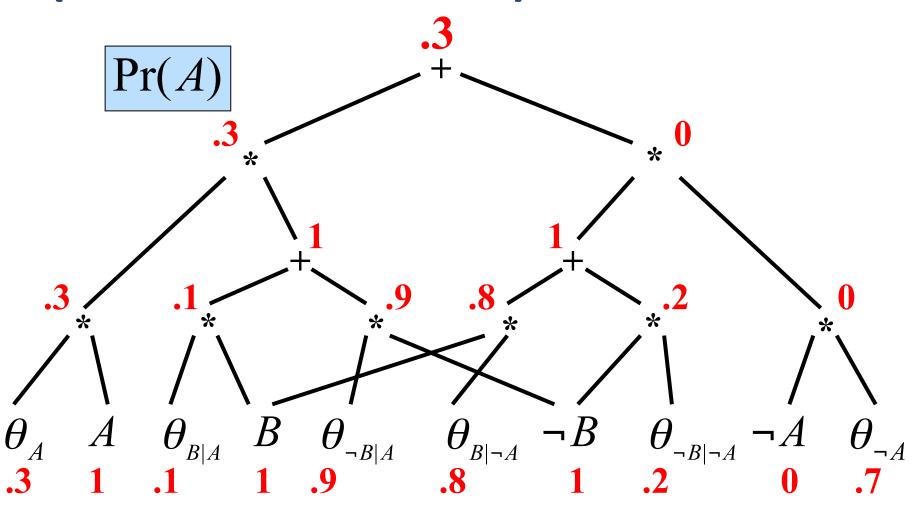


Α	В	С	Pr(.)
Т	Т	Т	$oldsymbol{ heta}_{\scriptscriptstyle A} oldsymbol{ heta}_{\scriptscriptstyle B A} oldsymbol{ heta}_{\scriptscriptstyle C A}$
Т	Т	F	$ heta_{\scriptscriptstyle A} heta_{\scriptscriptstyle B A} heta_{\scriptscriptstyle  eg C A}$
Т	F	Т	$oldsymbol{ heta}_{\scriptscriptstyle A} oldsymbol{ heta}_{\scriptscriptstyle \cap B A} oldsymbol{ heta}_{\scriptscriptstyle C A}$
Т	F	F	$ heta_{\scriptscriptstyle A} heta_{\scriptscriptstyle  eg B A} heta_{\scriptscriptstyle  eg C A}$
F	Т	Т	$oldsymbol{ heta}_{ eg A} oldsymbol{ heta}_{B  eg A} oldsymbol{ heta}_{C  eg A}$
F	Т	F	$oxed{ heta_{ abla_A} heta_{ abla_{  abla_A}} heta_{ abla_{  abla_A}}}$
F	F	Т	$ heta_{ eg A} heta_{ eg B  eg A} heta_{C  eg A}$
F	F	F	$oldsymbol{ heta}_{ eg_A} oldsymbol{ heta}_{ eg_{B  eg_A}} oldsymbol{ heta}_{ eg_{C  eg_A}}$

# Probabilistic Inference by Weighted Model Counting



# Weighted Model Counting (Arithmetic Circuits)



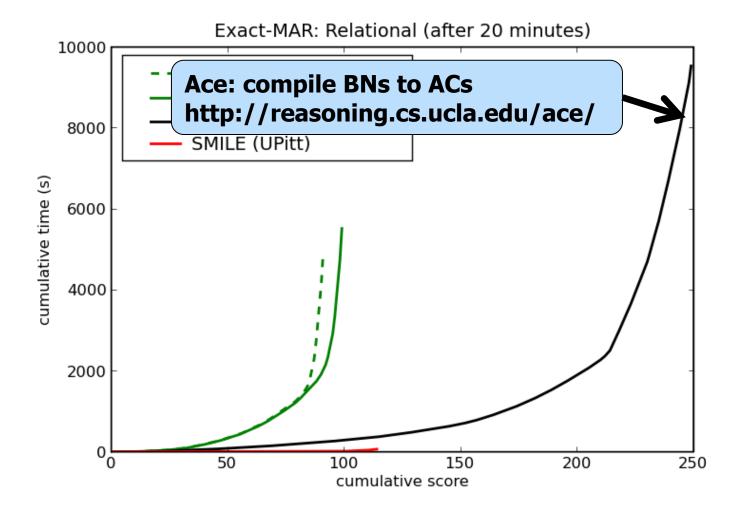
# Why Logic?

- Encoding local structure is easy:
  - Zero-parameters encoded by adding clauses:

$$\theta_{C|A} = 0$$
  $\neg A \lor \neg C$ 

Context-specific independence encoded by collapsing variables:

$$\theta_{C|AB} = \theta_{C|A \neg B}$$



- Relational networks (251 networks)
  - Average clique size is 50

### Alchemy - Open Source AI

#### Home

#### **Documentation**

Tutorial
User's Manual
Developer's Manual
APIs
Change Log
FAQs

Welcome to the Alchemy system! Alchemy is a software package providing a series of algorithms for statistical relational learning and probabilistic logic inference, based on the Markov logic representation. Alchemy allows you to easily develop a wide range of AI applications, including:

- · Collective classification
- Link prediction
- · Entity resolution
- · Social network modeling
- · Information extraction

If you are not already familiar with Markov logic, we recommend that you first read the paper Unifying Logical and Statistical AI.

#### Run: Requ

Alchemy is a software package providing a series of algorithms for statistical relational learning and probabilistic logic inference, based on Markov logic representations.

#### **Mailing Lists**

Alchemy Alchemy-announce Alchemy-update Alchemy-discuss

#### Repositories

Code Datasets MLNs Publications

#### Contributors

- Generative weight learning
- Structure learning
- MAP/MPE inference (including memory efficient)
- Probabilistic inference: MC-SAT, Gibbs Sampling, Simulated Tempering, Belief Propagation (including lifted)
- · Support for native and linked-in functions
- · Block inference and learning over variables with mutually exclusive and exhaustive values
- EM (to handle ground atoms with unknown truth values during learning)
- Specification of indivisible formulas (i.e. formulas that should not be broken up into separate clauses)
- Support of continuous features and domains
- Online inference
- · Decision Theory

In the next release we plan to include:

- · Online learning
- · Exact inference for small domains

## **Current Challenges**

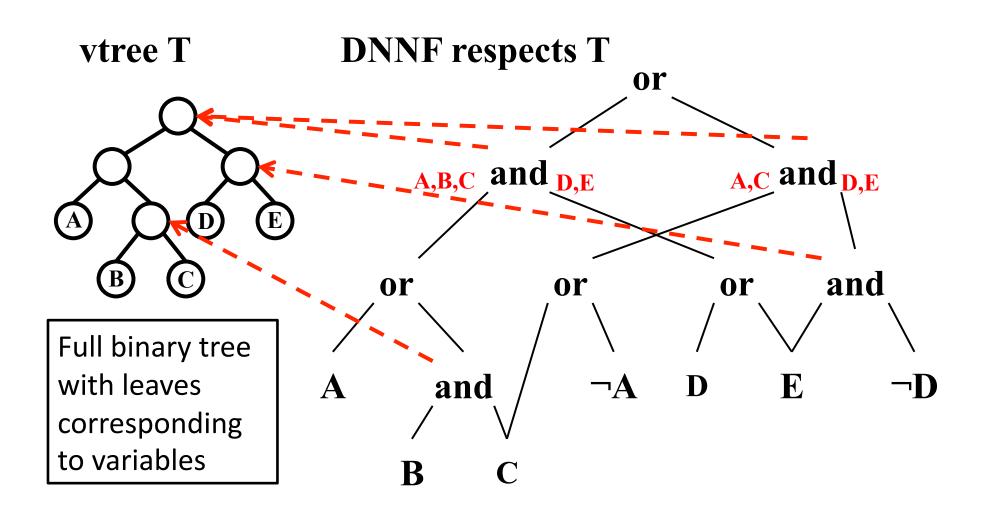
#### Incremental compilation:

- What? Current compilers monolithic: c2d (UCLA) and DSharp (Toronto)
- Need:
  - Logic: planning and verification applications
  - Probability: approximate inference
- Main insight:
  - Structured decomposability & vtrees (AAAI-08, AAAI-10)

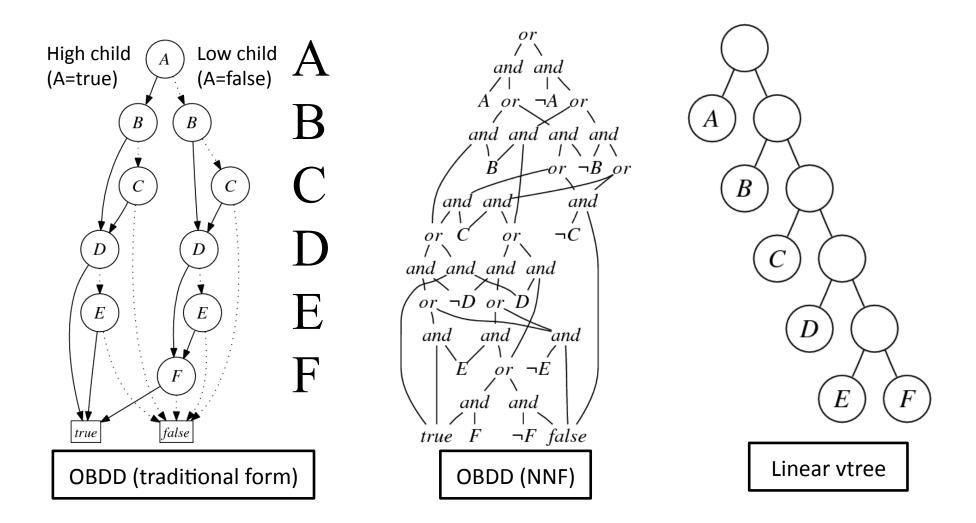
#### Guarantees and Complexity results:

- Upper & lower bounds on size of compilation (AAAI-10, ECAI-10)
- Main insights:
  - The notion of a decomposition (AAAI-10)
  - The notion of an interaction function (ECAI-10)

# **Structured Decomposability**



## **OBDD: DNNF that Respects Linear vtree**



### **Decomposition of Boolean Functions (AAAI-10)**

• Examples:  $f = (X_1 \vee X_2) \wedge (Y_1 \vee X_2) \wedge (X_1 \vee Y_2) \wedge (Y_1 \vee Y_2) \vee (X_2 \wedge Y_3)$ -  $X = \{X_1, X_2\}, Y = \{Y_1, Y_2, Y_3\}$ :

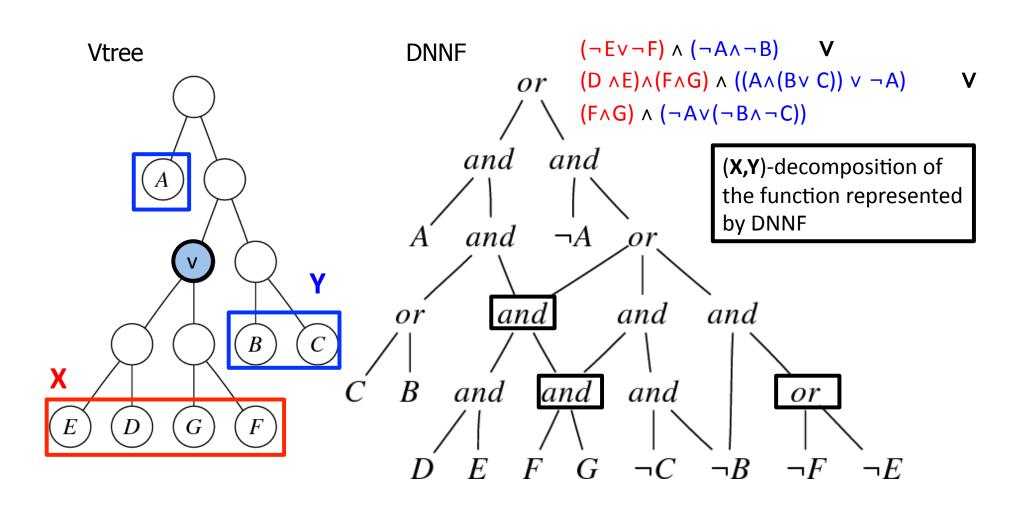
$$f(X,Y) = g(X) \wedge h(Y)$$

$$f(\mathbf{X},\mathbf{Y}) = f_1 \quad \forall \quad f_2 \quad \forall \quad f_3 \quad \forall \quad \dots \quad \forall \quad f_m$$

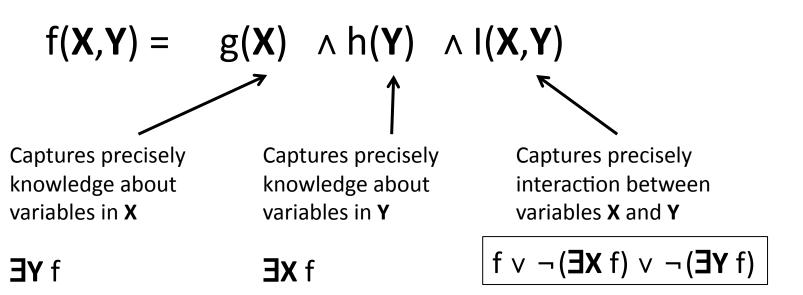
$$g_1(\mathbf{X}) \wedge h_1(\mathbf{Y}) \quad g_2(\mathbf{X}) \wedge h_2(\mathbf{Y}) \quad g_3(\mathbf{X}) \wedge h_3(\mathbf{Y}) \quad g_m(\mathbf{X}) \wedge h_m(\mathbf{Y})$$

(X,Y)-decomposition of f

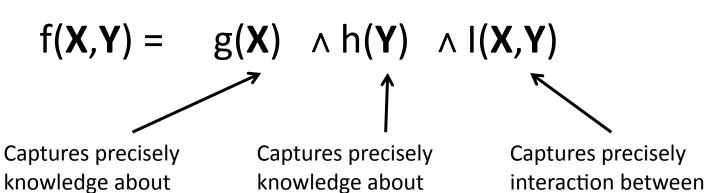
## **Lower Bounds (AAAI-10)**



## The Interaction Function (ECAI-10)



## The Interaction Function (ECAI-10)



variables in Y

**3Y** f true (B v C)

variables in X

A => (C => B)
¬A => (B => C)

variables X and Y

$$(A => B) (\neg A => C)$$
  $X = \{A\}$   $Y = \{B,C\}$ 

## **Current Research**

- Searching for good vtrees (on-going)
- Characterizing and searching for optimal decompositions
- Upper and lower bounds on size of DNNF
- Key objective: incremental compiler for DNNF and d-DNNF
- 555

## Two Main Themes

#### Exact inference as:

Enforcing decomposability and determinism on propositional knowledge bases

#### Approximate inference as:

Relaxing, compensating for, and recovering equivalence constraints (equalities)



#### AUTOMATED REASONING GROUP



#### UNIVERSITY OF CALIFORNIA LOS ANGELES

HOME :: PUBLICATIONS :: TALKS :: MEMBERS :: SOFTWARE :: RELATED COURSES



Reasoning Group Download Online Help Report a Bug Credits Contact

SENSITIVITY ANALYSIS, MODELING, INFERENCE AND MORE

Try ACE - a companion system for networks exhibiting

BatchTool Code Bandit Editing Models EM Learning File Formats Inference MAP

AR Group, UCLA

modeling and reasoning with Bayesian networks, developed in Java by the Automated Reasoning Group of Professor Adnan Darwiche at UCLA.

local structure: determinism and CSI

SamIam is a comprehensive tool for

Samiam includes two main components: a graphical user interface and a reasoning

engine. The graphical interface allows users to develop Bayesian network models and to save them in a variety of formats. The reasoning engine supports many tasks including: classical inference; parameter estimation; time-space tradeoffs; sensitivity analysis; and explanation-generation based on MAP and

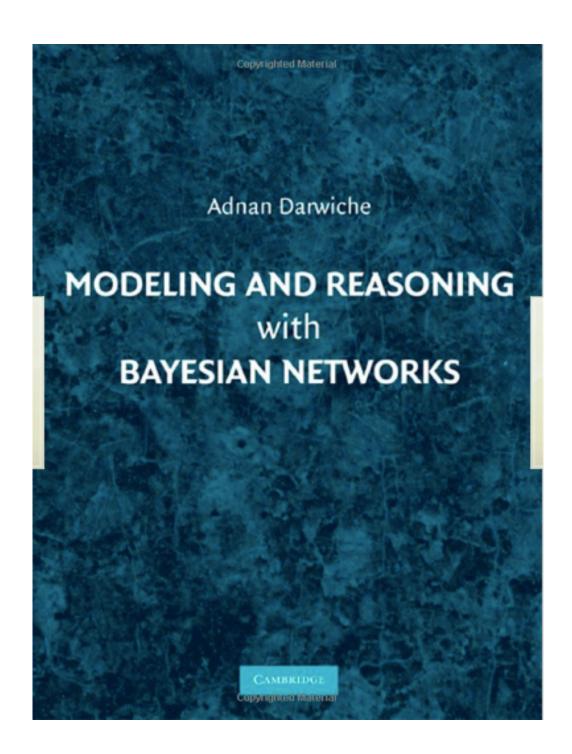
MPF Relational Models Sensitivity Analysis Time-Space Tradeoffs Timing MAP

MPE.

Home | Screenshots | Sponsors | Suggestions | Videos | FAQ | Links

Copyright @ 2004-2005, Automated Reasoning Group, University Of California, Los Angeles, All Rights Reserved.

# http://reasoning.cs.ucla.edu/samiam/





#### AUTOMATED REASONING GROUP



#### UNIVERSITY OF CALIFORNIA LOS ANGELES

HOME :: PUBLICATIONS :: TALKS :: MEMBERS :: SOFTWARE :: RELATED COURSES



Reasoning Group Download Online Help Report a Bug Credits Contact

SENSITIVITY ANALYSIS, MODELING, INFERENCE AND MORE

BatchTool
Code Bandit
Editing Models
EM Learning
File Formats
Inference
MAP
MPE
Relational Models

MPE
Relational Models
Sensitivity Analysis
Time-Space Tradeoffs
Timing MAP

Try ACE - a companion system for networks exhibiting local structure: determinism and CSI

SamIam is a comprehensive tool for modeling and reasoning with Bayesian networks, developed in Java by the Automated Reasoning Group of Professor Adnan Darwiche at UCLA.

Samiam includes two main components: a graphical user interface and a reasoning engine. The graphical interface allows users



to develop Bayesian network models and to save them in a variety of formats. The reasoning engine supports many tasks including: classical inference; parameter estimation; time-space tradeoffs; sensitivity analysis; and explanation-generation based on MAP and MPE.

AR Group, UCLA

Home | Screenshots | Sponsors | Suggestions | Videos | FAQ | Links

Copyright @ 2004-2005, Automated Reasoning Group, University Of California, Los Angeles, All Rights Reserved.

## http://reasoning.cs.ucla.edu

## Two Main Themes

#### Exact inference as:

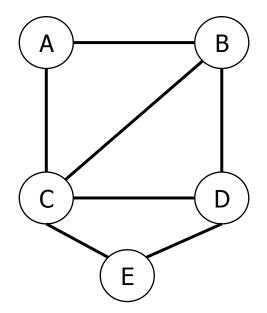
Enforcing decomposability and determinism on propositional knowledge bases

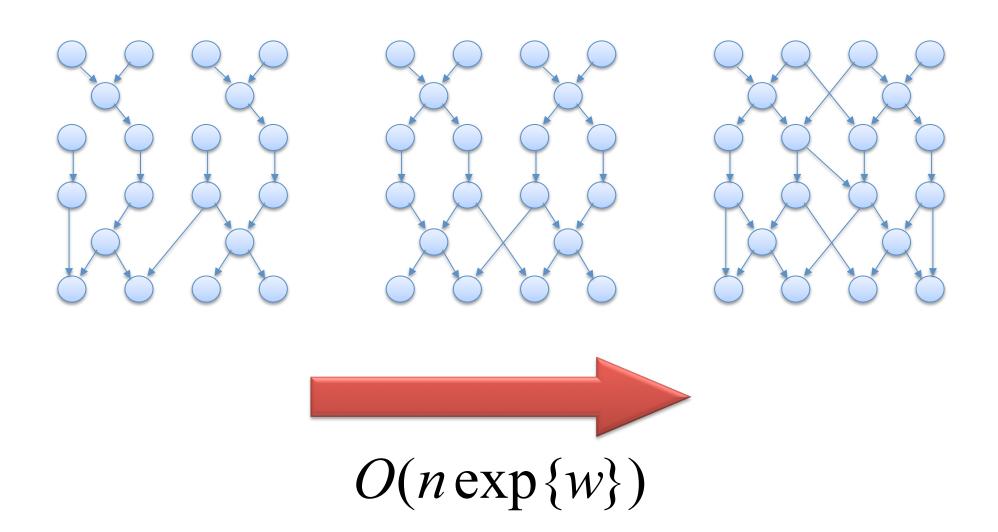
#### Approximate inference as:

Relaxing, compensating for, and recovering equivalence constraints (equalities)

**CNF** 

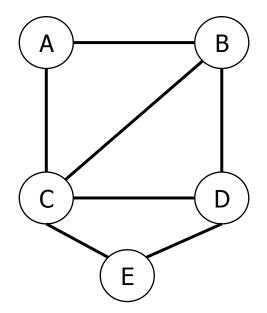
**Constraint Graph** 





**CNF** 

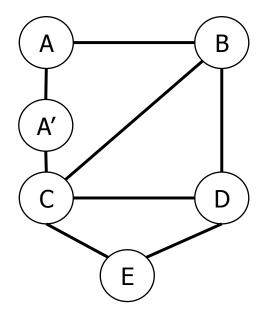
**Constraint Graph** 



**CNF** 

$$(A \lor B \lor \neg C) \land$$
  
 $(A' \lor \neg B \lor C) \land$   
 $(C \lor \neg D \lor E) \land$   
 $(B \lor D) \land$   
 $(D \lor E) \land$   
 $(A = A')$ 

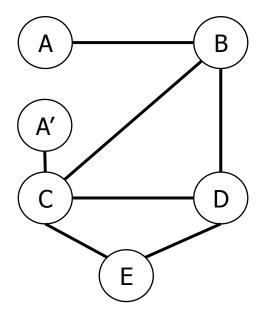
**Constraint Graph** 



### **Treewidth**

**CNF** 

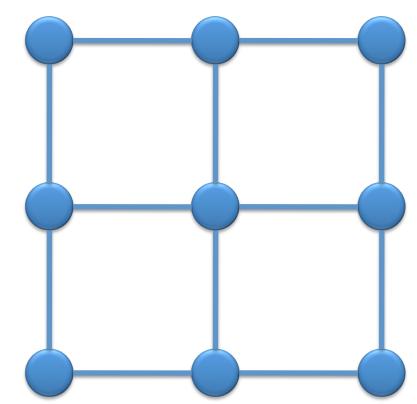
**Constraint Graph** 



## **Equivalence Constraints**

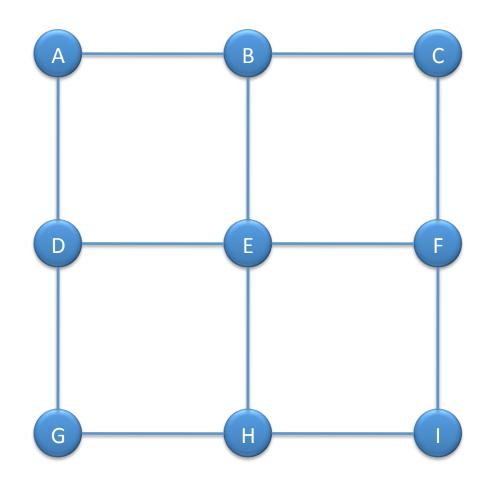
$$\psi_{eq}(X_i = x_i, X_j = x_j) =$$

$$\begin{cases} 1 & \text{if } x_i = x_j \\ 0 & \text{otherwise} \end{cases}$$

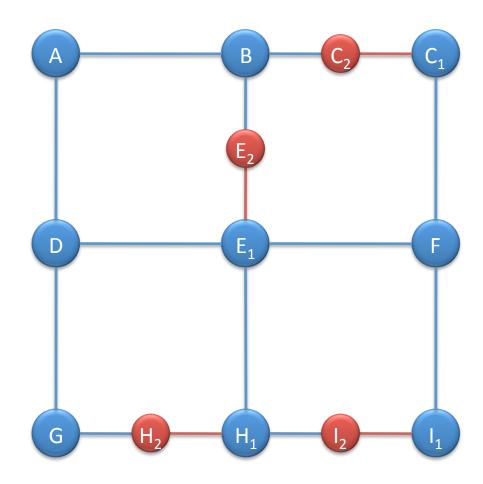


$$Pr(x_1,...,x_n) = \frac{1}{Z} \psi(x_1,x_5)...\psi(x_2,x_4)...\theta(x_1)...\theta(x_n)$$

• Model  ${\mathcal M}$ 

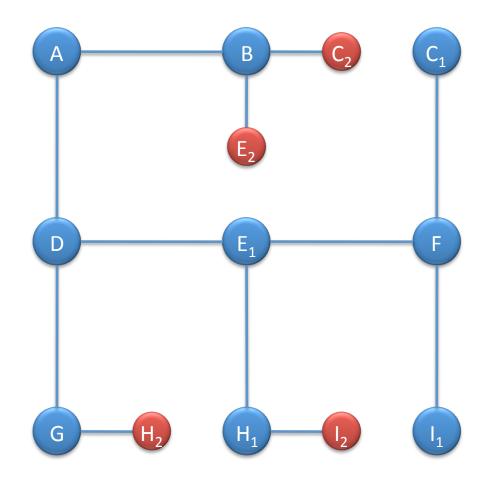


Model + Eq.

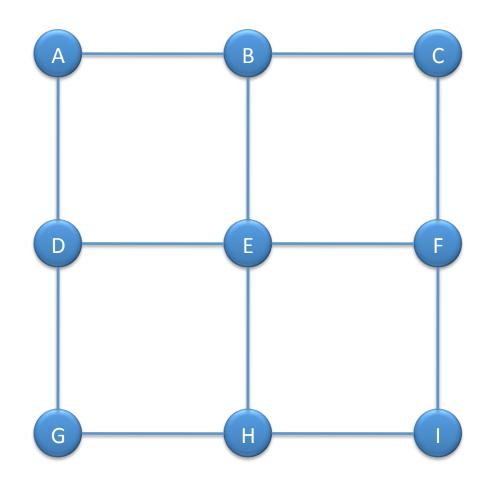


Relaxed

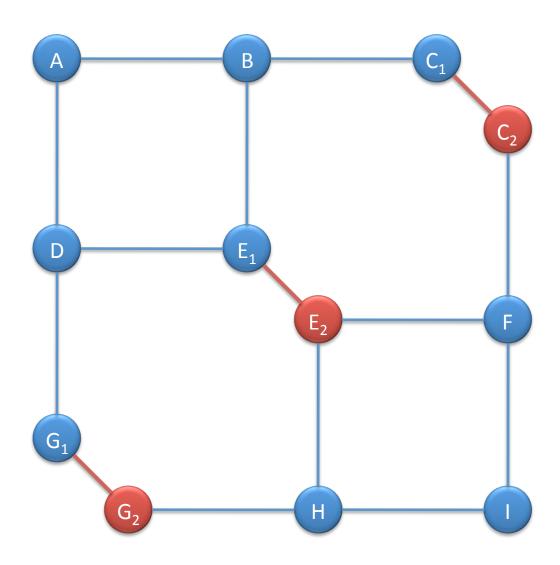
• Treewidth 1



• Model  ${\mathcal M}$ 

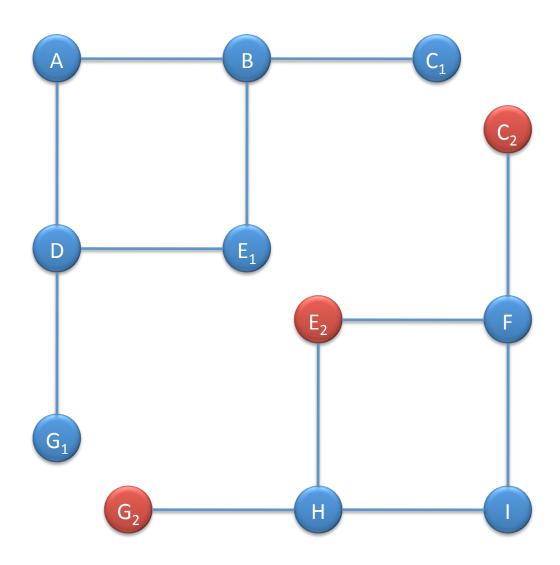


• Model + Eq.

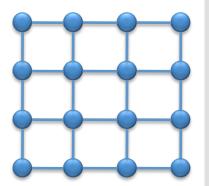


Relaxed

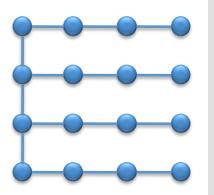
Decomposed



#### Model + Eq



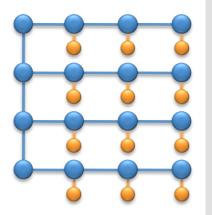
Relax



Intractable model, augmented with equivalence constraints

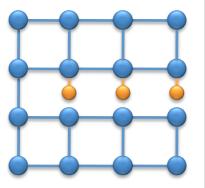
Simplify network structure: Relax equivalence constraints

### Compensate



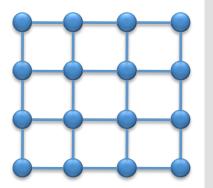
Compensate for relaxation: Restore a weaker equivalence

#### Recover

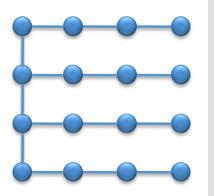


Recover structure, identify an improved approximation

#### Model + Eq



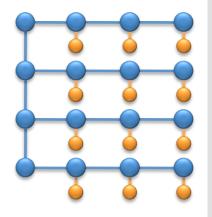
Relax



Intractable model, augmented with equivalence constraints

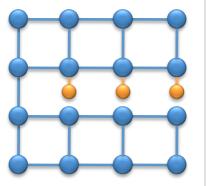
Usually gives upper/lower bounds: mini-buckets, MAXSAT

#### Compensate



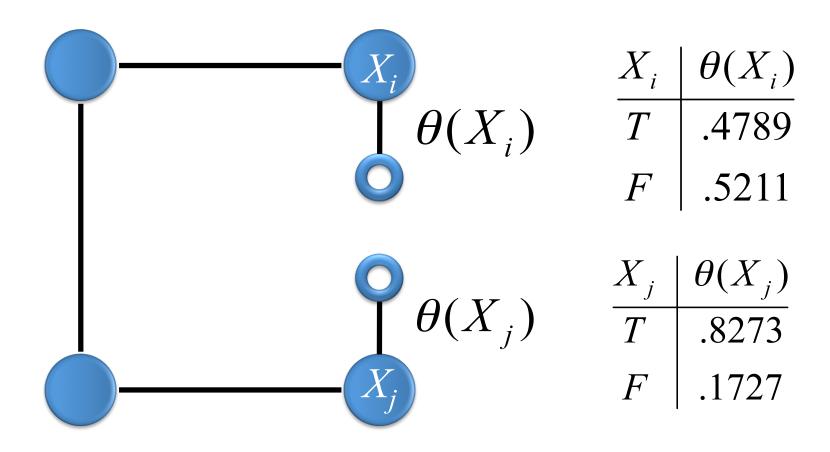
Compensate for relaxation: Restore a weaker equivalence

#### Recover

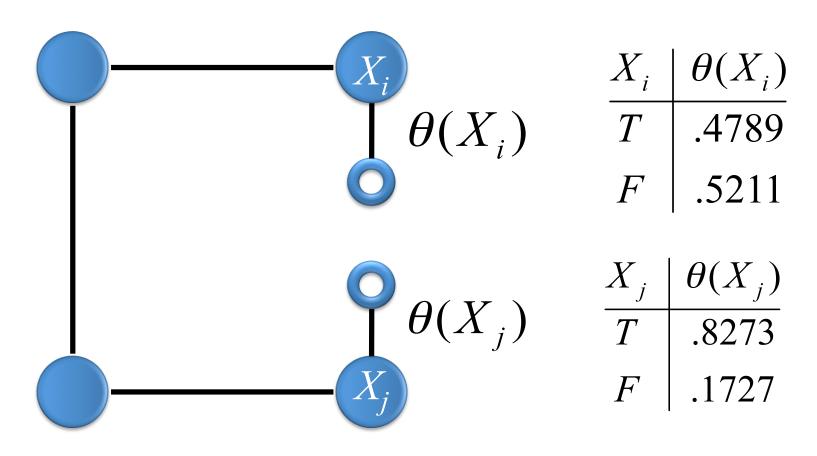


Recover structure, identify an improved approximation

### Compensating for an Equivalence

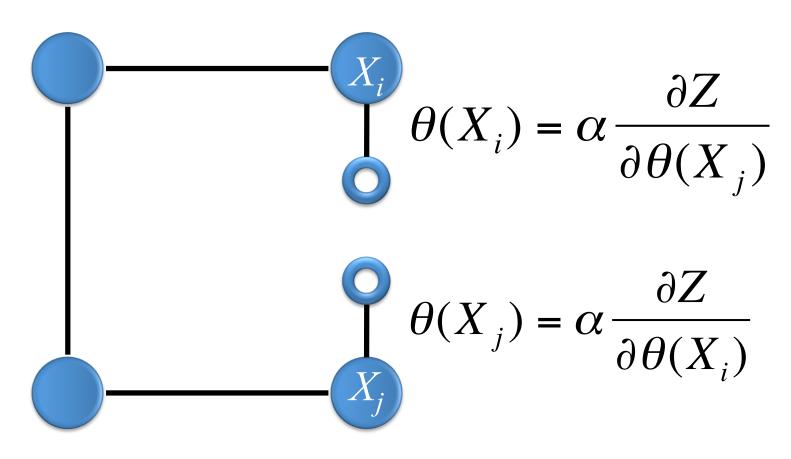


### Compensating for an Equivalence

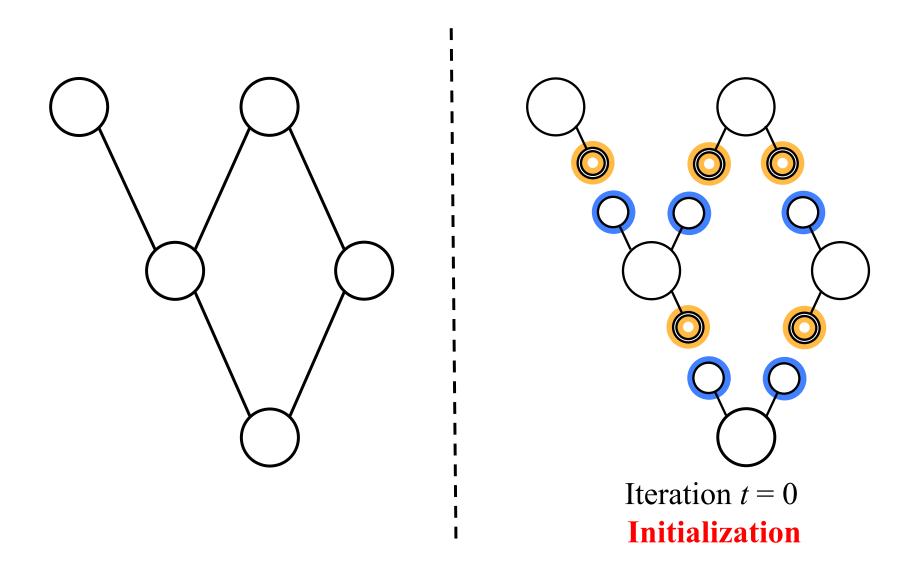


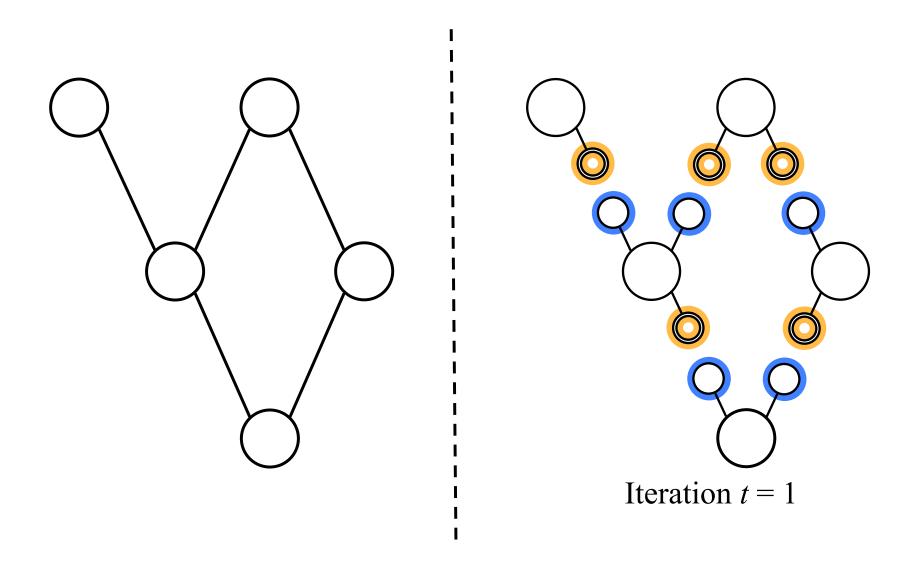
$$\Pr(X_i = x) = \Pr(X_j = x)$$

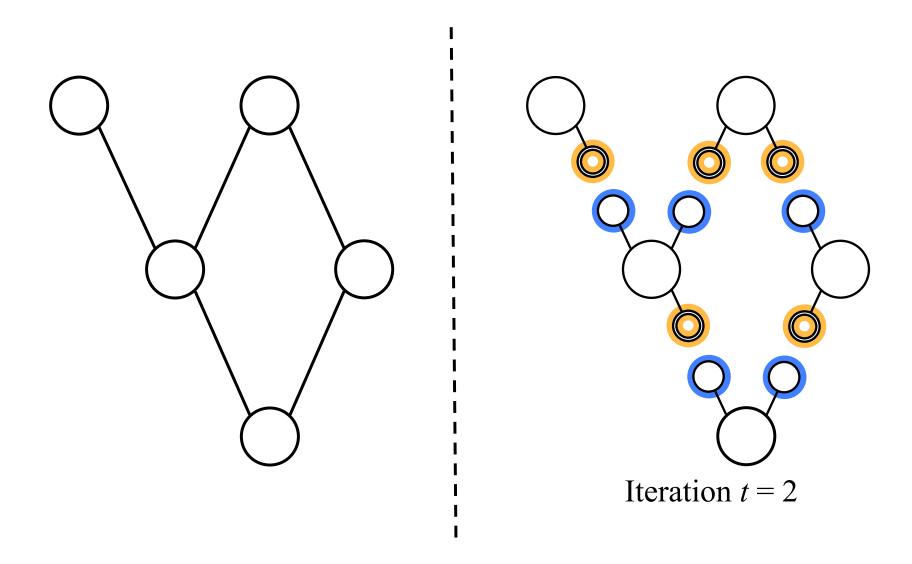
### Compensating for an Equivalence

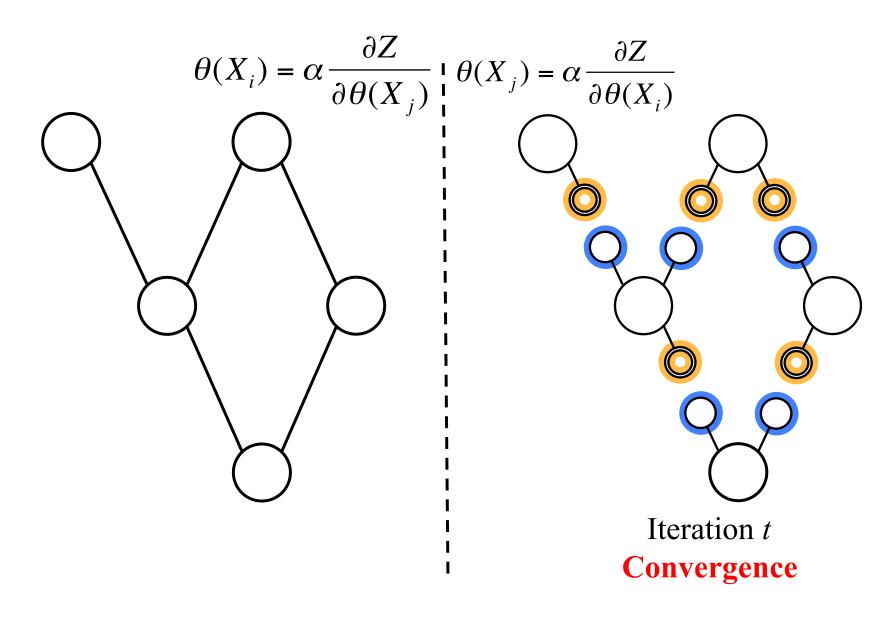


$$\Pr(X_i = x) = \Pr(X_j = x)$$

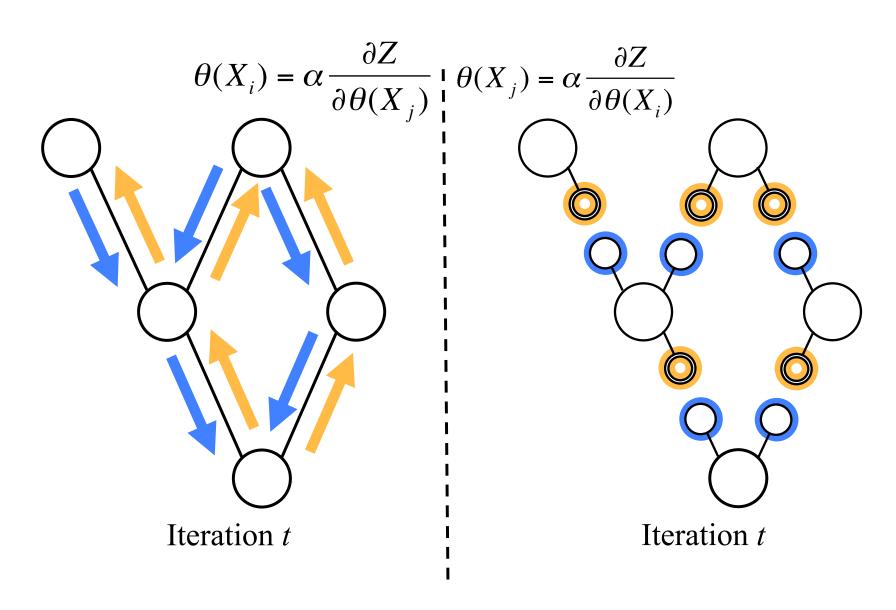




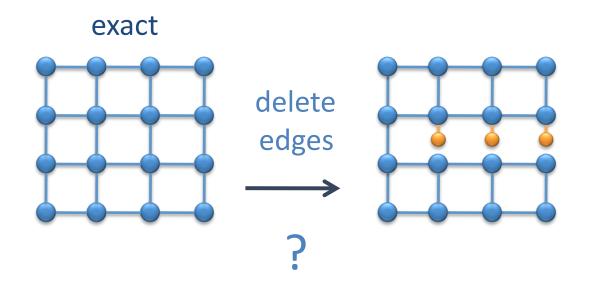




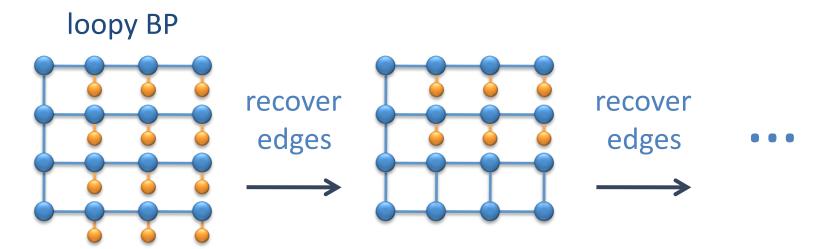
### Characterizing Loopy Belief Propagation

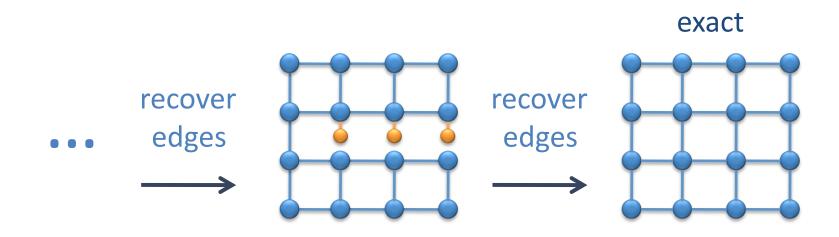


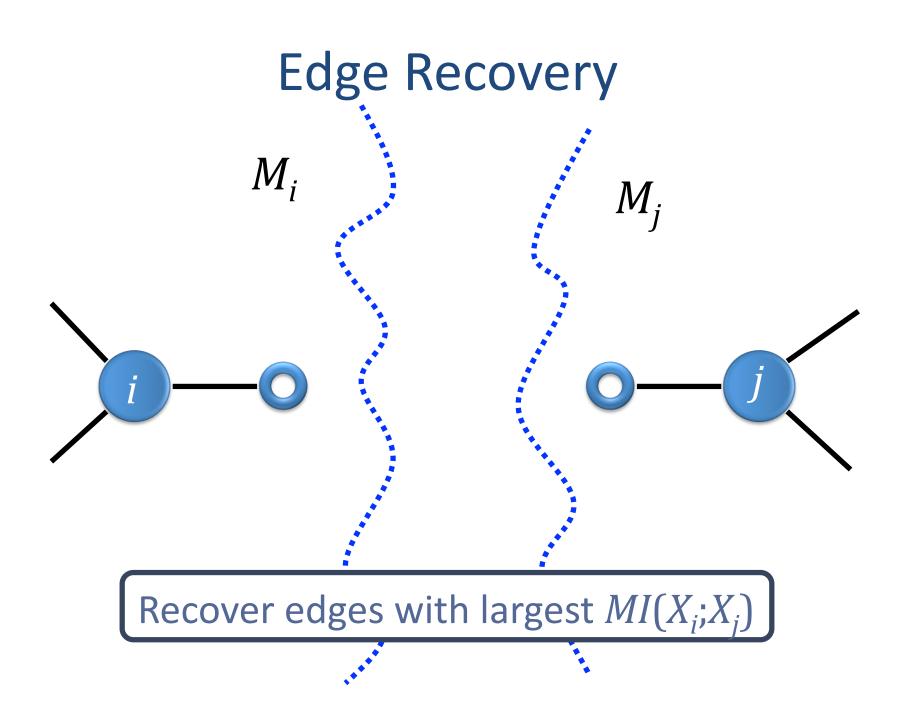
# Which Edges to Delete?



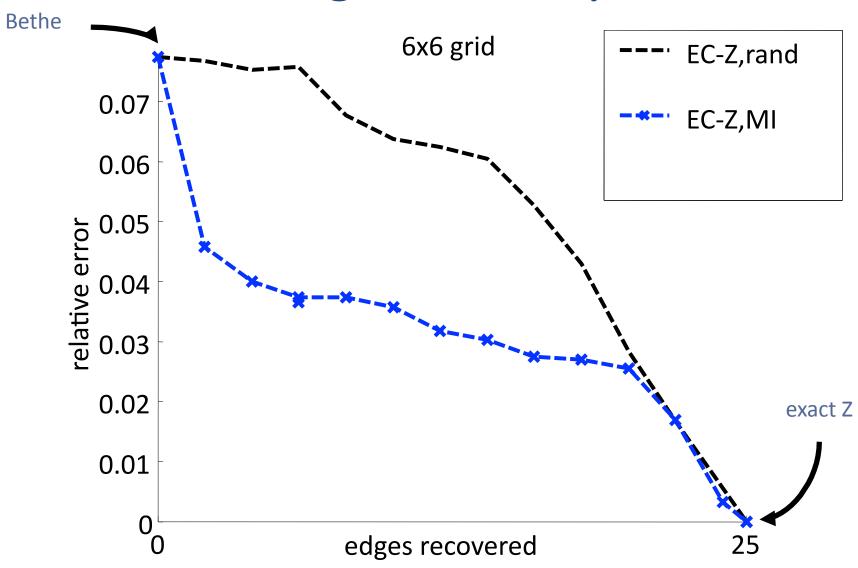
# Edge Recovery







# Edge Recovery



## **Evaluation Benchmarks**

Benchmark	PR	MAR	MPE
CSP	8	8	55
Grids	20	20	40
Image Alignment			10
Medical Diagnosis	26	26	
Object Detection	96	96	92
Pedigree	4	4	
Protein Folding			21
Protein-Protein Interaction			8
Segmentation	50	50	50
TOTAL	204	204	287

## **Overall Results**

PR Task: 20 Seconds MAR Task: 20 Seconds

Solver	Score
edbr	1.7146
vgogate	2.1620
libDai	2.2775

Solver	Score
edbq	0.2390
libDai2	0.3064
vgogate	0.4409

# Ideally...

Exact inference based on compiling CNFs

- Edge recovery using incremental compilation:
  - conjoin recovered equivalence constraint with current compilation

Not there yet: more engineering needed!

## Key Ideas

- Approximate inference: formulated as exact inference in an approximate model
- Approximate models: obtained by relaxing and compensating for equivalence constraints
- Anytime inference: selective recovery of equivalence constraints
- Exact inference: formulated in terms of enforcing decomposability and determinism of propositional knowledge bases



#### AUTOMATED REASONING GROUP



#### UNIVERSITY OF CALIFORNIA LOS ANGELES

HOME :: PUBLICATIONS :: TALKS :: MEMBERS :: SOFTWARE :: RELATED COURSES



Reasoning Group Download Online Help Report a Bug Credits Contact

SENSITIVITY ANALYSIS, MODELING, INFERENCE AND MORE

Try ACE - a companion system for networks exhibiting

BatchTool Code Bandit Editing Models EM Learning File Formats Inference MAP

AR Group, UCLA

modeling and reasoning with Bayesian networks, developed in Java by the Automated Reasoning Group of Professor Adnan Darwiche at UCLA.

local structure: determinism and CSI

SamIam is a comprehensive tool for

Samiam includes two main components: a graphical user interface and a reasoning

engine. The graphical interface allows users to develop Bayesian network models and to save them in a variety of formats. The reasoning engine supports many tasks including: classical inference; parameter estimation; time-space tradeoffs; sensitivity analysis; and explanation-generation based on MAP and

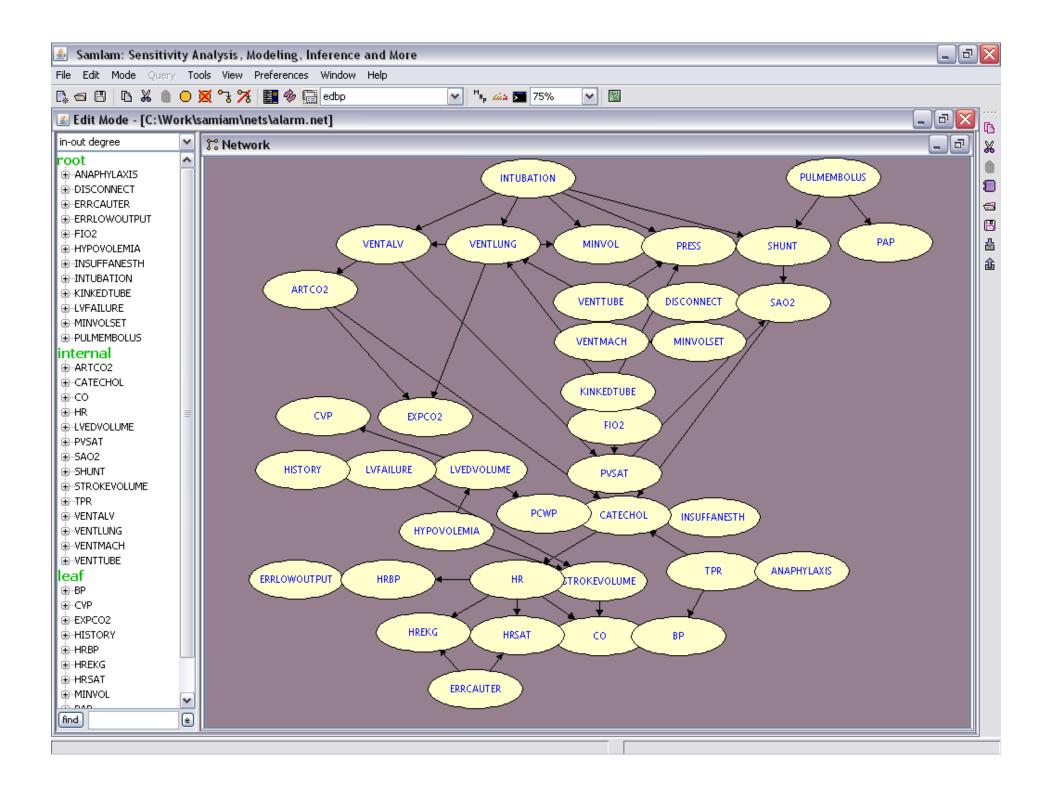
MPF Relational Models Sensitivity Analysis Time-Space Tradeoffs Timing MAP

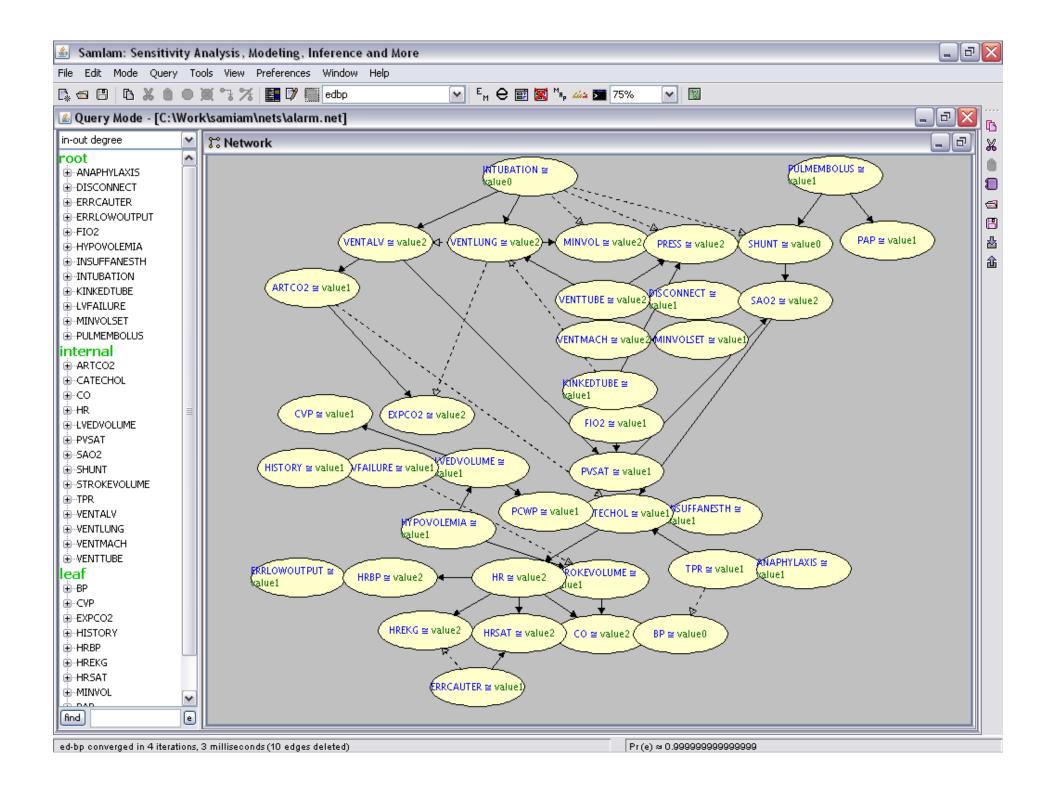
MPE.

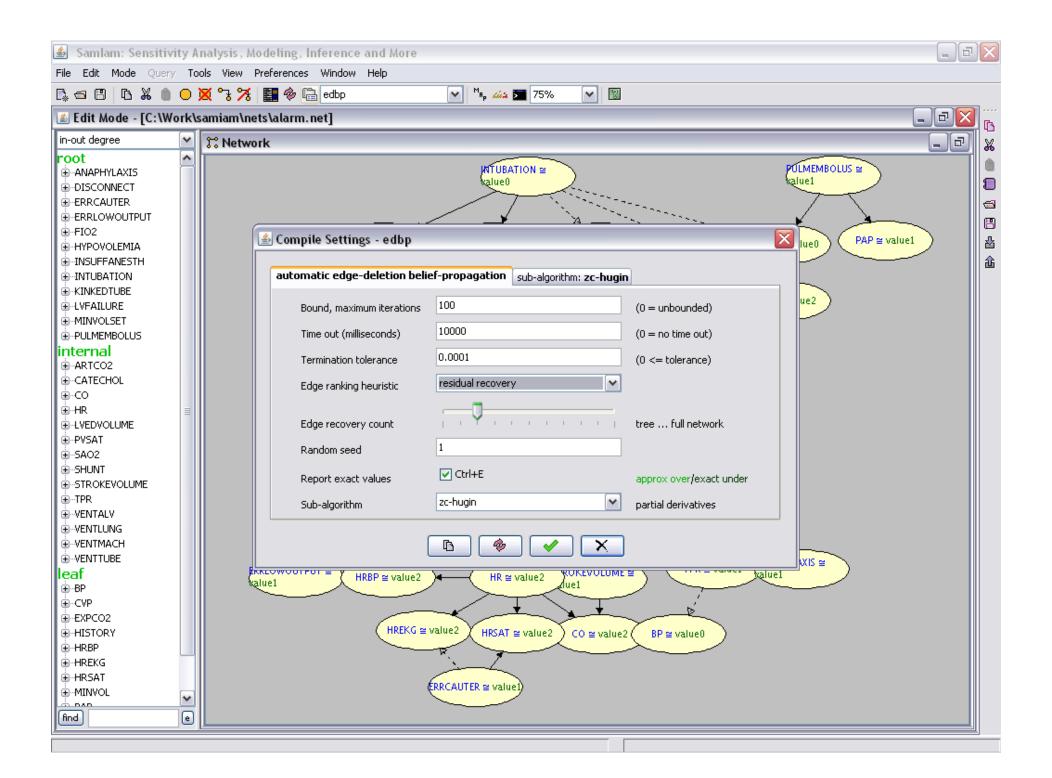
Home | Screenshots | Sponsors | Suggestions | Videos | FAQ | Links

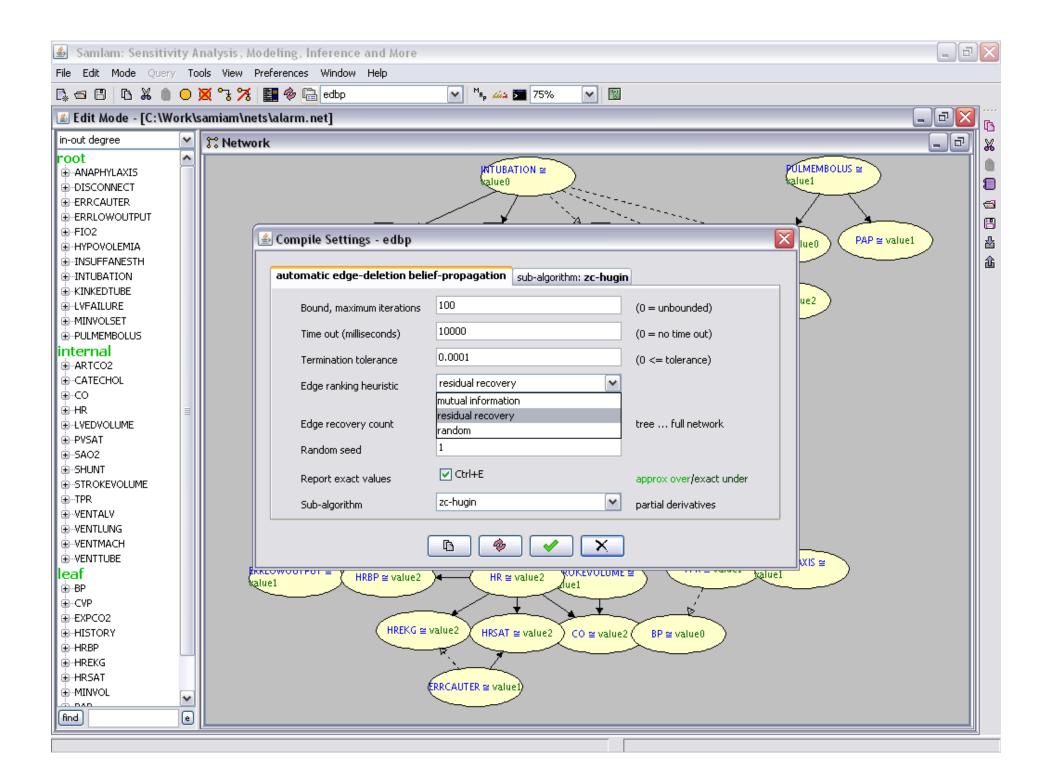
Copyright @ 2004-2005, Automated Reasoning Group, University Of California, Los Angeles, All Rights Reserved.

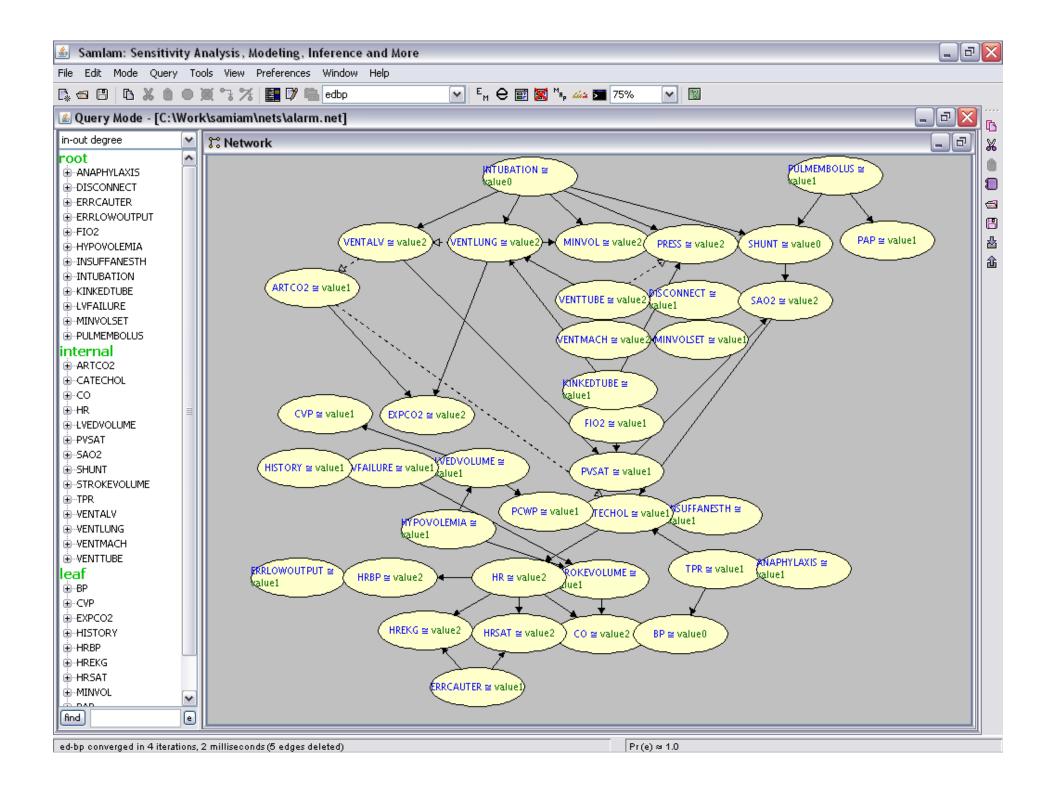
# http://reasoning.cs.ucla.edu/samiam/

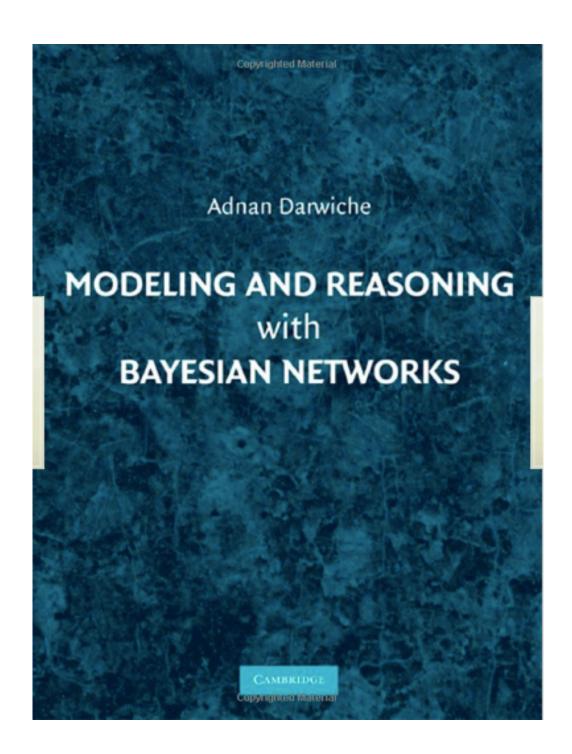














#### AUTOMATED REASONING GROUP



#### UNIVERSITY OF CALIFORNIA LOS ANGELES

HOME :: PUBLICATIONS :: TALKS :: MEMBERS :: SOFTWARE :: RELATED COURSES



Reasoning Group Download Online Help Report a Bug Credits Contact

SENSITIVITY ANALYSIS, MODELING, INFERENCE AND MORE

BatchTool
Code Bandit
Editing Models
EM Learning
File Formats
Inference
MAP
MPE
Relational Models

MPE
Relational Models
Sensitivity Analysis
Time-Space Tradeoffs
Timing MAP

Try ACE - a companion system for networks exhibiting local structure: determinism and CSI

SamIam is a comprehensive tool for modeling and reasoning with Bayesian networks, developed in Java by the Automated Reasoning Group of Professor Adnan Darwiche at UCLA.

Samiam includes two main components: a graphical user interface and a reasoning engine. The graphical interface allows users



to develop Bayesian network models and to save them in a variety of formats. The reasoning engine supports many tasks including: classical inference; parameter estimation; time-space tradeoffs; sensitivity analysis; and explanation-generation based on MAP and MPE.

AR Group, UCLA

Home | Screenshots | Sponsors | Suggestions | Videos | FAQ | Links

Copyright @ 2004-2005, Automated Reasoning Group, University Of California, Los Angeles, All Rights Reserved.

## http://reasoning.cs.ucla.edu