General Formulations for Structures: Markov Logic

CS 6355: Structured Prediction

THE UNIVERSITY OF UTAH

So far…

- Binary and multiclass classifiers
- Graphs as structured output
- Sequence labeling

This lecture

- Graphical models
  - Bayesian Networks
  - Markov Random Fields (MRFs)

- Formulations of structured output
  - Joint models
    - Markov Logic Network
  - Conditional models
    - Conditional Random Fields (again)
    - Constrained Conditional Models

Where are we?

- Graphical models
  - Bayesian Networks
  - Markov Random Fields (MRFs)

- Formulations of structured output
  - Joint models
    - Markov Logic Network
  - Conditional models
    - Conditional Random Fields (again)
    - Constrained Conditional Models

A language for creating MRFs

Consider the following statements

- Smoking causes cancer
- If two people are friends and one smokes, so does the other

[Example from Domingos and Lowd 2009]

A language for creating MRFs

Consider the following statements

- Smoking causes cancer
- If two people are friends and one smokes, so does the other

- In logic: $\forall x; Smokes(x) \Rightarrow Cancer(x)$

$$\forall x, y; Friends(x, y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

[Example from Domingos and Lowd 2009]

A language for creating MRFs

Consider the following statements

– Smoking causes cancer
– If two people are friends and one smokes, so does the other

- In logic: $\forall x; Smokes(x) \Rightarrow Cancer(x)$

$$\forall x, y; Friends(x, y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

- The statements are not necessarily absolutely true
  – How do we associate degrees of belief to statements?

[Example from Domingos and Lowd 2009]

Markov Logic Networks

From rules to graphical models

- Convert to clauses

$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

$$\forall x, y; Friends(x,y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

→

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x,y) \vee \neg Smokes(x) \vee Smokes(y)$$

[Example from Domingos and Lowd 2009]

Markov Logic Networks

From rules to graphical models

- Convert to clauses

$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

$$\forall x, y; Friends(x,y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

→

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x,y) \vee \neg Smokes(x) \vee Smokes(y)$$

- Associate a potential function for each clause
  - Think of each formula as a factor
  - Typically, log-linear in all the variables involved

[Example from Domingos and Lowd 2009]

Markov Logic Networks

From rules to graphical models

- Convert to clauses

$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

$$\forall x, y; Friends(x, y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

→

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x, y) \vee \neg Smokes(x) \vee Smokes(y)$$

- Associate a potential function for each clause
  - Think of each formula as a factor
  - Typically, log-linear in all the variables involved

- Ground the logical expressions to all x, y that you care about

[Example from Domingos and Lowd 2009]

Example of a ground network

$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

$$\forall x, y; Friends(x, y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

→

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x, y) \vee \neg Smokes(x) \vee Smokes(y)$$

Suppose there are two people in the world: Anna (A), Bob (B)

[Example from Domingos and Lowd 2009]

Example of a ground network

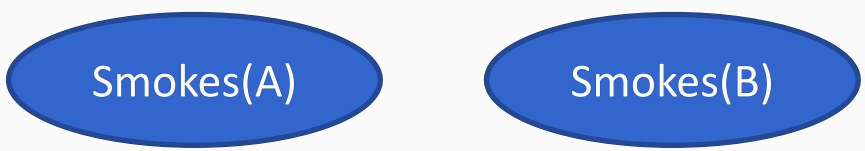
$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

$$\forall x, y; Friends(x, y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x, y) \vee \neg Smokes(x) \vee Smokes(y)$$

Suppose there are two people in the world: Anna (A), Bob (B)



[Example from Domingos and Lowd 2009]

Example of a ground network

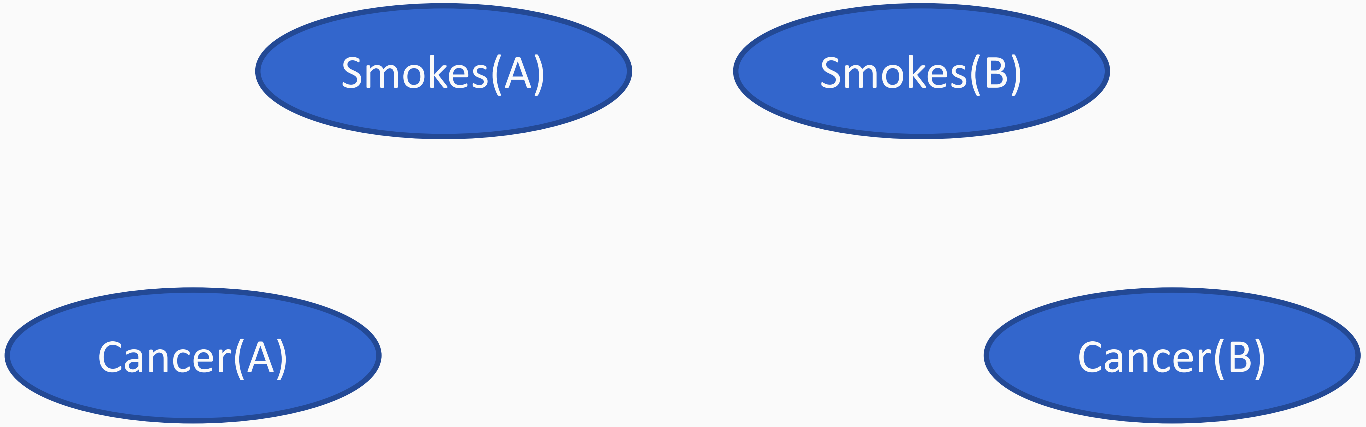
$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

$$\forall x, y; Friends(x, y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x, y) \vee \neg Smokes(x) \vee Smokes(y)$$

Suppose there are two people in the world: Anna (A), Bob (B)



[Example from Domingos and Lowd 2009]

Example of a ground network

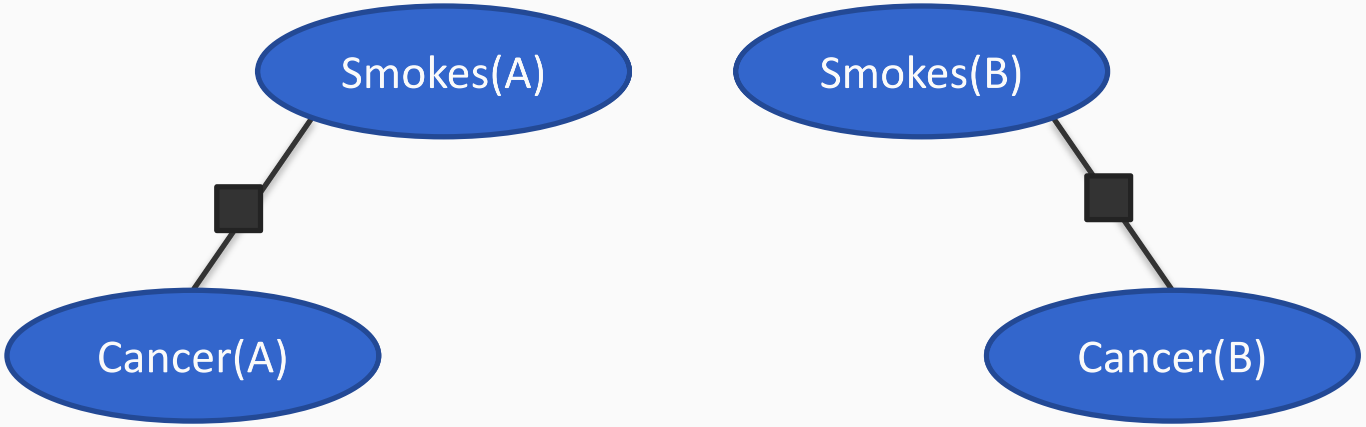
$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

$$\forall x, y; Friends(x, y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x, y) \vee \neg Smokes(x) \vee Smokes(y)$$

Suppose there are two people in the world: Anna (A), Bob (B)



[Example from Domingos and Lowd 2009]

Example of a ground network

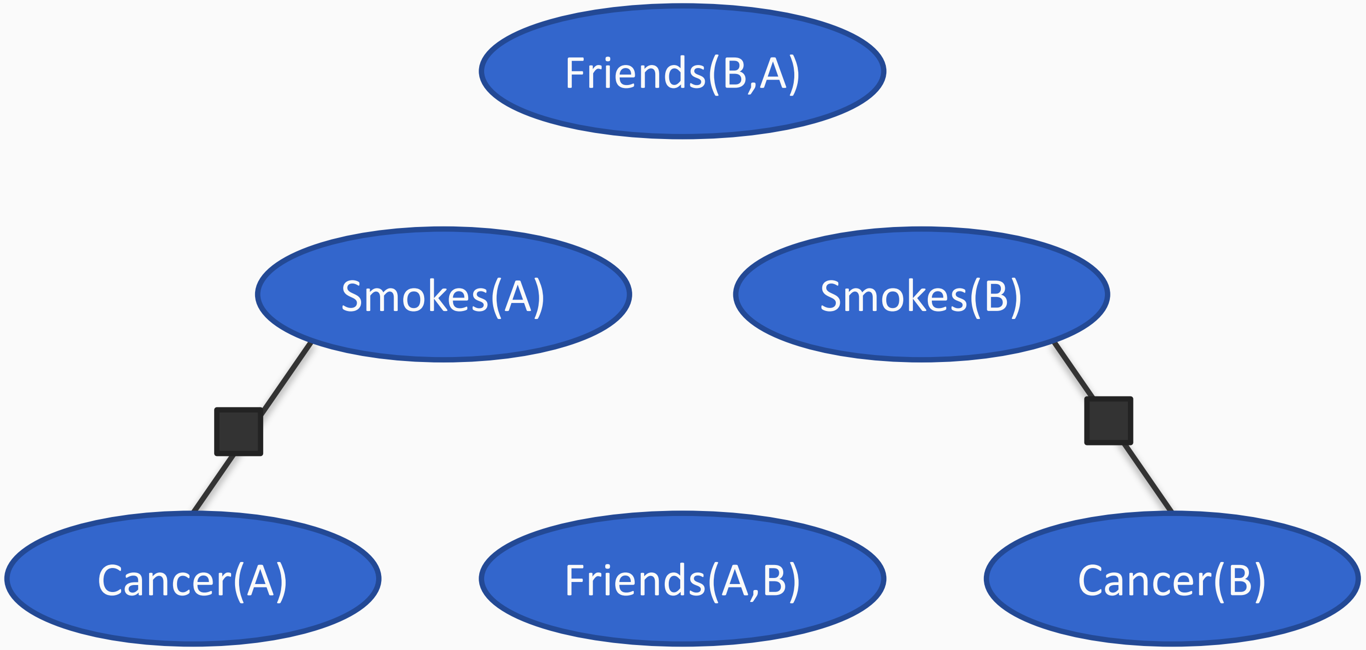
$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

$$\forall x, y; Friends(x,y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x,y) \vee \neg Smokes(x) \vee Smokes(y)$$

Suppose there are two people in the world: Anna (A), Bob (B)



[Example from Domingos and Lowd 2009]

Example of a ground network

$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

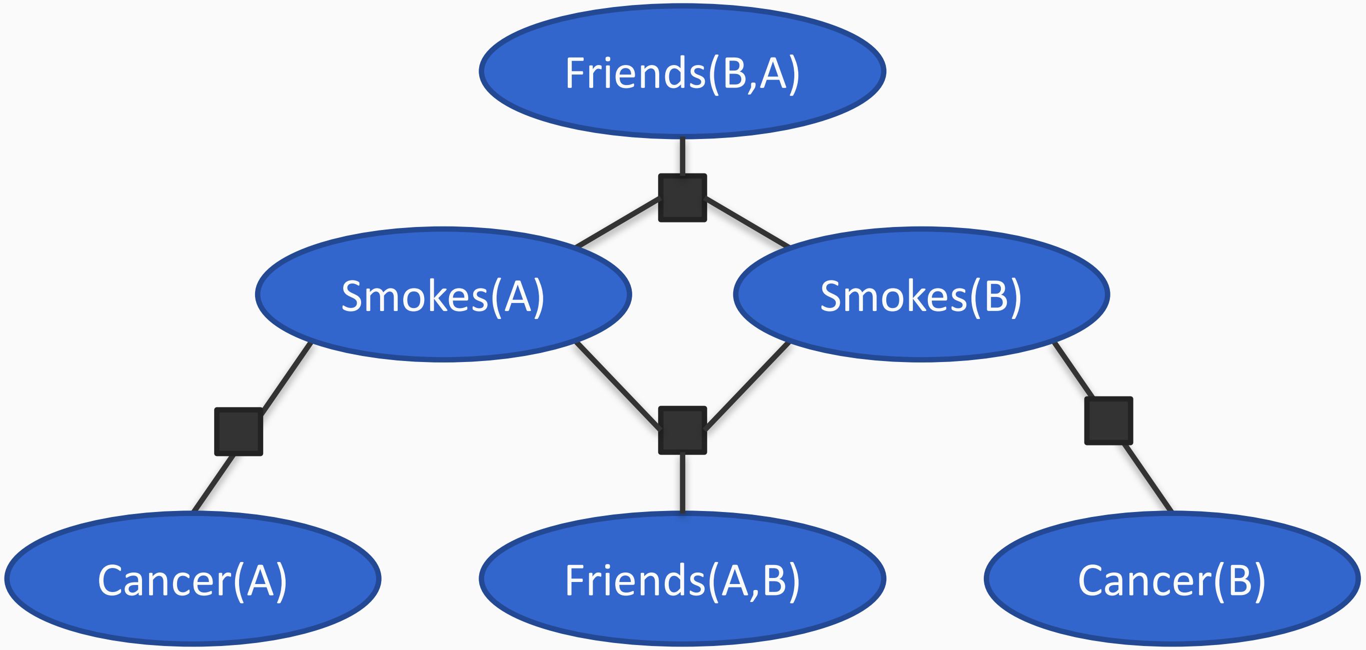
$$\forall x, y; Friends(x,y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

→

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x,y) \vee \neg Smokes(x) \vee Smokes(y)$$

Suppose there are two people in the world: Anna (A), Bob (B)



[Example from Domingos and Lowd 2009]

Example of a ground network

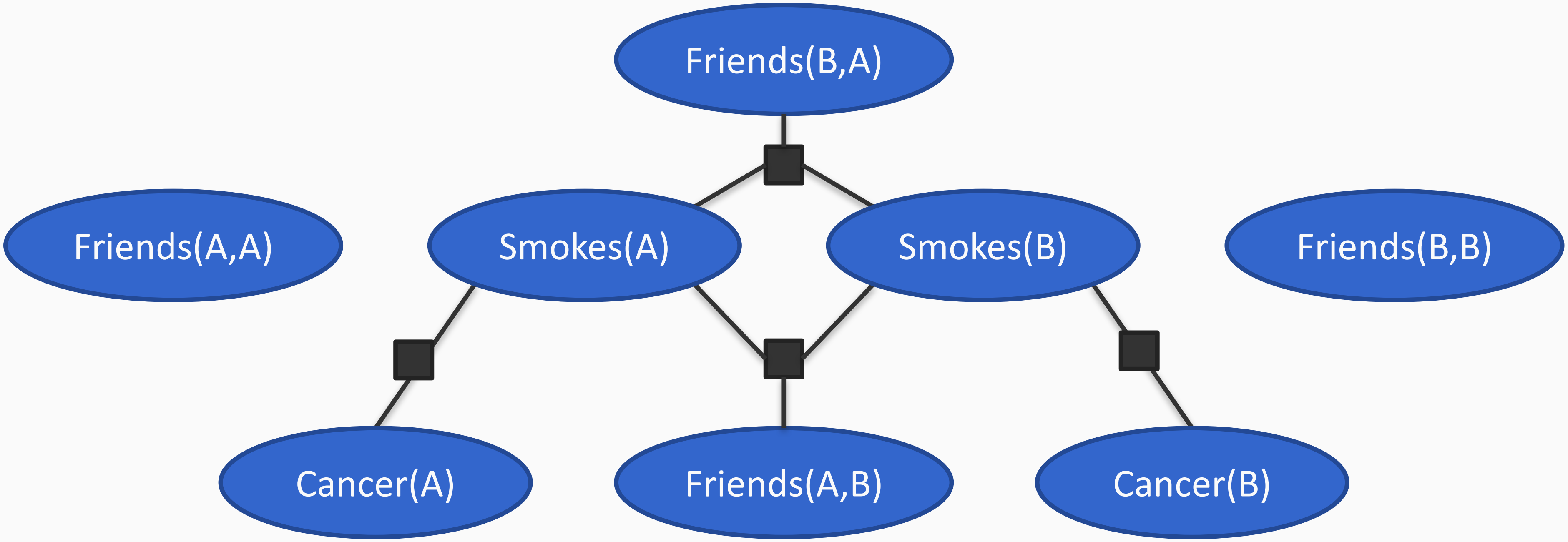
$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

$$\forall x, y; Friends(x,y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x,y) \vee \neg Smokes(x) \vee Smokes(y)$$

Suppose there are two people in the world: Anna (A), Bob (B)



[Example from Domingos and Lowd 2009]

Example of a ground network

$$\forall x; Smokes(x) \Rightarrow Cancer(x)$$

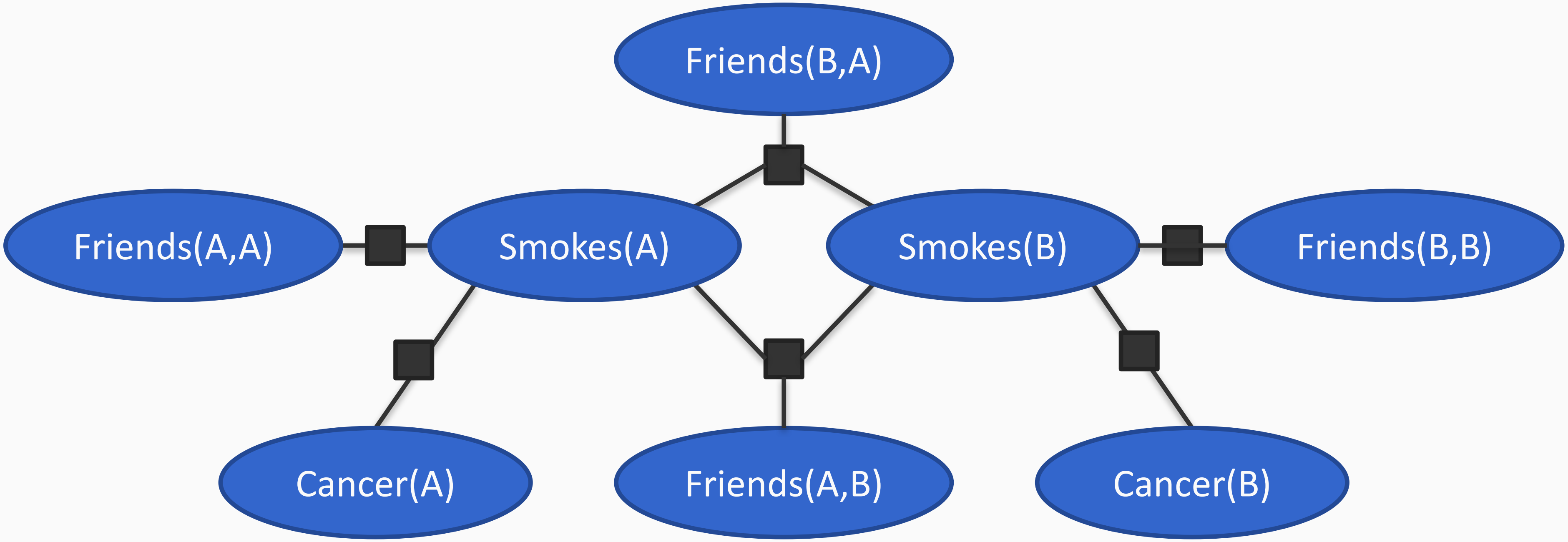
$$\forall x, y; Friends(x,y) \wedge Smokes(x) \Rightarrow Smokes(y)$$

→

$$\forall x; \neg Smokes(x) \vee Cancer(x)$$

$$\forall x, y; \neg Friends(x,y) \vee \neg Smokes(x) \vee Smokes(y)$$

Suppose there are two people in the world: Anna (A), Bob (B)



[Example from Domingos and Lowd 2009]

Learning in MLNs

Two kinds of learning (true for all formulations, actually)

1. Given a network/collection of formulas, learn the weights that define the potential functions
   - Use maximum likelihood method
   - Other training methods exist
     - Approximate the likelihood with pseudo-likelihood

Learning in MLNs

Two kinds of learning (true for all formulations, actually)

1. Given a network/collection of formulas, learn the weights that define the potential functions
   - Use maximum likelihood method
   - Other training methods exist
     - Approximate the likelihood with pseudo-likelihood

2. Learn the formulas themselves
   - Much harder
   - Uses ideas from inductive logic programming

Markov Logic Networks

- Specifies a undirected graphical model template
  - "Unroll" the network to get the full MRF
    - And then use any standard graphical model algorithms
  - Requires us to ground the network
    - There has been work on inference at the first order level too, though
  - **Note**: Each formula corresponds to a factor in the factor graph
  - Other ways of specifying templates exist

- Creates a joint model
  - Unlike conditional random fields, for example