General Formulations for Structures: Markov Logic

CS 6355: Structured Prediction

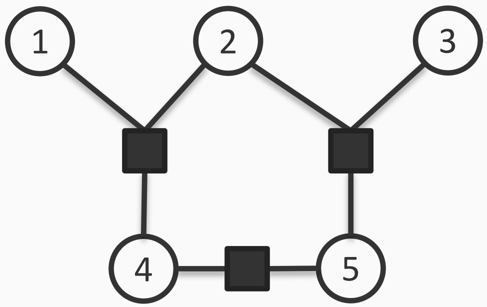
THE UNIVERSITY OF UTAH

This lecture

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

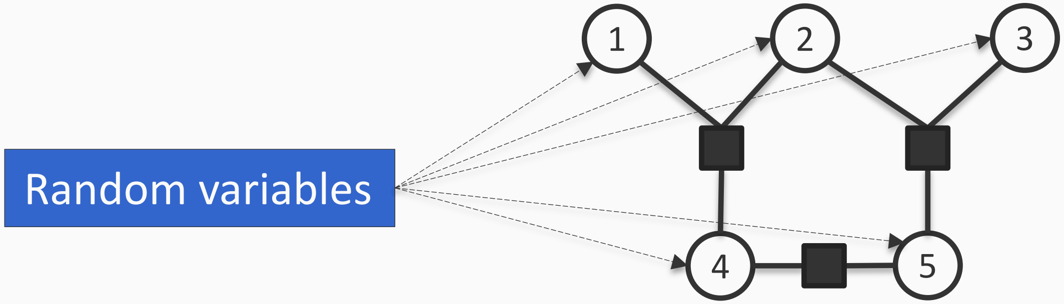
- Formulations of structured output
  – Joint models
    - Markov Logic Network
  – Conditional models
    - Conditional Random Fields (again)
    - Constrained Conditional Models

We have seen Markov networks



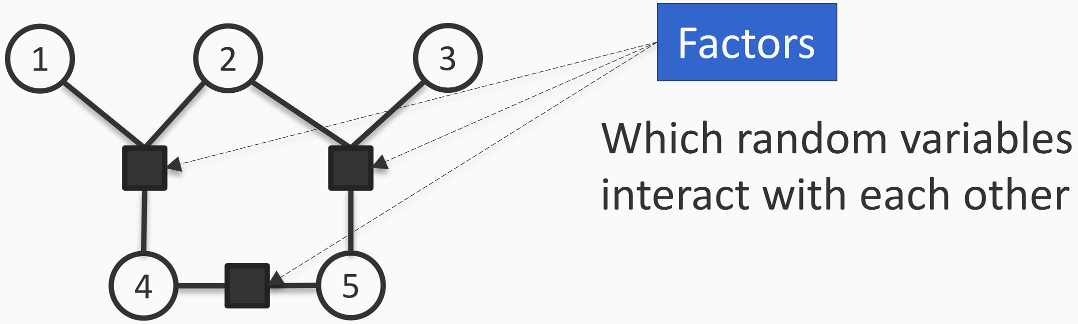
$$P(x_1, x_2, x_3, x_4, x_5) = \frac{1}{Z} f_a(x_1, x_2, x_4) f_b(x_2, x_3, x_5) f_c(x_4, x_5)$$

We have seen Markov networks



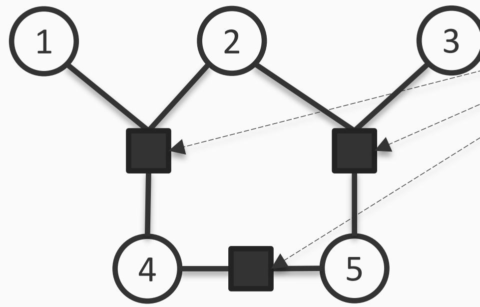
$$P(x_1, x_2, x_3, x_4, x_5) = \frac{1}{Z} f_a(x_1, x_2, x_4) f_b(x_2, x_3, x_5) f_c(x_4, x_5)$$

We have seen Markov networks



$$P(x_1, x_2, x_3, x_4, x_5) = \frac{1}{Z} f_a(x_1, x_2, x_4) f_b(x_2, x_3, x_5) f_c(x_4, x_5)$$

We have seen Markov networks



Factors

Which random variables interact with each other

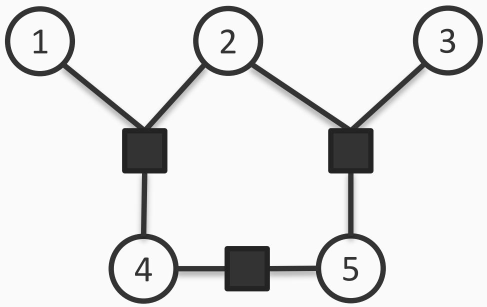
$$P(x_1, x_2, x_3, x_4, x_5) = \frac{1}{Z} f_a(x_1, x_2, x_4) f_b(x_2, x_3, x_5) f_c(x_4, x_5)$$

Factor potentials

One per factor

How are the interactions between random variables defined (via scoring)

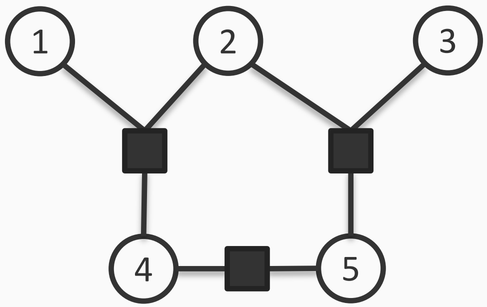
We have seen Markov networks



$$P(x_1, x_2, x_3, x_4, x_5) = \frac{1}{Z} f_a(x_1, x_2, x_4) f_b(x_2, x_3, x_5) f_c(x_4, x_5)$$

With the joint distribution, we can ask the probability of any subset of the random variables

We have seen Markov networks



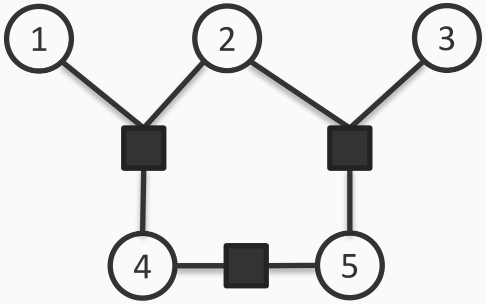
$$P(x_1, x_2, x_3, x_4, x_5) = \frac{1}{Z} f_a(x_1, x_2, x_4) f_b(x_2, x_3, x_5) f_c(x_4, x_5)$$

With the joint distribution, we can ask the probability of any subset of the random variables

Example 1

$$P(x_2) = \sum_{x_1, x_3, x_4, x_5} P(x_1, x_2, x_3, x_4, x_5)$$

We have seen Markov networks



$$P(x_1, x_2, x_3, x_4, x_5) = \frac{1}{Z} f_a(x_1, x_2, x_4) f_b(x_2, x_3, x_5) f_c(x_4, x_5)$$

With the joint distribution, we can ask the probability of any subset of the random variables

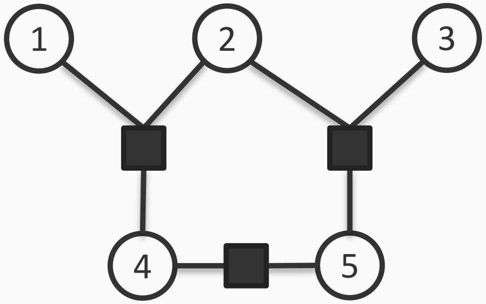
Example 1

$$P(x_2) = \sum_{x_1, x_3, x_4, x_5} P(x_1, x_2, x_3, x_4, x_5)$$

Example 2

$$P(\, x_1 \mid x_3, x_5 \,) = \frac{P(x_1, x_3, x_5)}{P(x_3, x_5)} = \frac{\sum_{x_2, x_4} P(x_1, x_2, x_3, x_4, x_5)}{\sum_{x_1, x_2, x_4} P(x_1, x_2, x_3, x_4, x_5)}$$

We have seen Markov networks

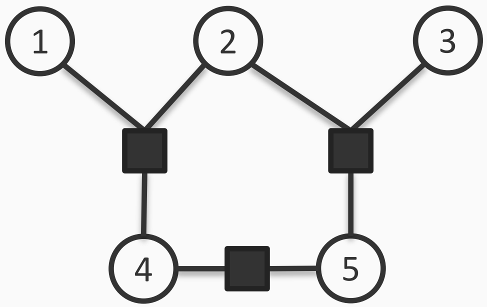


$$P(x_1, x_2, x_3, x_4, x_5) = \frac{1}{Z} f_a(x_1, x_2, x_4) f_b(x_2, x_3, x_5) f_c(x_4, x_5)$$

A graphical notation that:

1. Defines how a **joint** probability distribution is factorized over components (factors)
2. Clarifies the independence assumptions at play

We have seen Markov networks



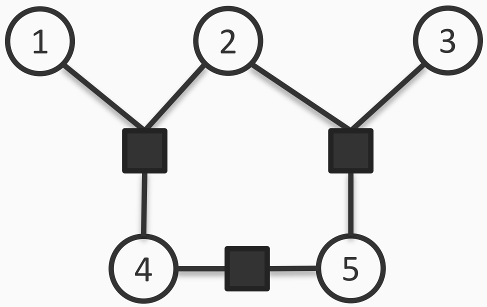
$$P(x_1, x_2, x_3, x_4, x_5) = \frac{1}{Z} f_a(x_1, x_2, x_4) f_b(x_2, x_3, x_5) f_c(x_4, x_5)$$

A graphical notation that:

1. Defines how a **joint** probability distribution is factorized over components (factors)
2. Clarifies the independence assumptions at play

But what if there are many different random variables? And what if there are groups of factors that behave similarly? The structure of the network is knowledge.

We have seen Markov networks



$$P(x_1, x_2, x_3, x_4, x_5) = \frac{1}{Z} f_a(x_1, x_2, x_4) f_b(x_2, x_3, x_5) f_c(x_4, x_5)$$

A graphical notation that:

1. Defines how a **joint** probability distribution is factorized over components (factors)
2. Clarifies the independence assumptions at play

But what if there are many different random variables? And what if there are groups of factors that behave similarly? The structure of the network is knowledge.

Can we construct Markov networks by declaratively stating such knowledge?

Representing and reasoning about knowledge

Consider the following statements

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

Questions to think about

- How do we represent this knowledge?
- How do we answer questions like: “If Anna is friends with Bob, and Bob smokes, can Anna get cancer?”

Representing and reasoning about knowledge

Consider the following statements

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

Questions to think about

- How do we represent this knowledge?
- How do we answer questions like: "If Anna is friends with Bob, and Bob smokes, can Anna get cancer?"

Logic is a natural language for declaratively stating knowledge and making inferences.

Representing knowledge

"Smoking causes cancer."

$$\forall x, \mathrm{Smokes}(x) \Rightarrow \mathrm{Cancer}(x)$$

We use predicates Smokes and Cancer in this universally quantified statement.

Representing knowledge

"Smoking causes cancer."

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

"If two people are friends and one smokes, so does the other."

$$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

Reasoning about knowledge

"Smoking causes cancer."

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

"If two people are friends and one smokes, so does the other."

$$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

Suppose we have two friends Anna and Bob, and Bob smokes. What can we infer about Anna?

Reasoning about knowledge

"Smoking causes cancer."

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

"If two people are friends and one smokes, so does the other."

$$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

Suppose we have two friends Anna and Bob, and Bob smokes. What can we infer about Anna?

1. Anna and Bob are friends: $\text{Friends}(\text{Anna}, \text{Bob})$

Reasoning about knowledge

"Smoking causes cancer."

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

"If two people are friends and one smokes, so does the other."

$$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

Suppose we have two friends Anna and Bob, and Bob smokes. What can we infer about Anna?

1. Anna and Bob are friends: $\text{Friends}(\text{Anna}, \text{Bob})$
2. Bob smokes: $\text{Smokes}(\text{Bob})$

Reasoning about knowledge

"Smoking causes cancer."

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

"If two people are friends and one smokes, so does the other."

$$\forall x, y\ \text{Friends}(x, y) \land \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

Suppose we have two friends Anna and Bob, and Bob smokes. What can we infer about Anna?

1. Anna and Bob are friends: $\text{Friends}(\text{Anna}, \text{Bob})$
2. Bob smokes: $\text{Smokes}(\text{Bob})$
3. We know that: $\text{Friends}(\text{Anna}, \text{Bob}) \land \text{Smokes}(\text{Bob}) \Rightarrow \text{Smokes}(\text{Anna})$

Reasoning about knowledge

"Smoking causes cancer."

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

"If two people are friends and one smokes, so does the other."

$$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

Suppose we have two friends Anna and Bob, and Bob smokes. What can we infer about Anna?

1. Anna and Bob are friends: $\text{Friends}(\text{Anna}, \text{Bob})$
2. Bob smokes: $\text{Smokes}(\text{Bob})$
3. We know that: $\text{Friends}(\text{Anna}, \text{Bob}) \wedge \text{Smokes}(\text{Bob}) \Rightarrow \text{Smokes}(\text{Anna})$
4. And we also know that: $\text{Smokes}(\text{Anna}) \Rightarrow \text{Cancer}(\text{Anna})$

Reasoning about knowledge

"Smoking causes cancer."

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

"If two people are friends and one smokes, so does the other."

$$\forall x, y \ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

Suppose we have two friends Anna and Bob, and Bob smokes. What can we infer about Anna?

1. Anna and Bob are friends: $\text{Friends}(\text{Anna}, \text{Bob})$
2. Bob smokes: $\text{Smokes}(\text{Bob})$
3. We know that: $\text{Friends}(\text{Anna}, \text{Bob}) \wedge \text{Smokes}(\text{Bob}) \Rightarrow \text{Smokes}(\text{Anna})$
4. And we also know that: $\text{Smokes}(\text{Anna}) \Rightarrow \text{Cancer}(\text{Anna})$

Logic is an expressive language, but how do we deal with uncertainty?

From logic to Markov networks

Consider the following statements

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

- In logic:

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$
$$\forall x, y\ \text{Friends}(x, y) \land \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

[Example from Domingos and Lowd 2009]

From logic to Markov networks

Consider the following statements

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

- In logic:

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$
$$\forall x, y\ \text{Friends}(x, y) \land \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

- 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

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

$$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

[Example from Domingos and Lowd 2009]

Markov Logic Networks

From rules to graphical models

- Convert to clauses

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

$$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

Recall:

- A literal is predicate or its negation
- A clause is a disjunction of literals
- Any implication $A \Rightarrow B$ is equivalent to $\neg A \vee B$

[Example from Domingos and Lowd 2009]

Markov Logic Networks

From rules to graphical models

- Convert to clauses

$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$	→	$\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$
$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$		$\forall x, y, \neg\text{Friends}(x, y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$

Recall:

- A literal is predicate or its negation
- A clause is a disjunction of literals
- Any implication $A \Rightarrow B$ is equivalent to $\neg A \vee B$

[Example from Domingos and Lowd 2009]

Markov Logic Networks

From rules to graphical models

- Convert to clauses

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

$$\forall x, y\ \text{Friends}(x,y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

→

$$\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$$

$$\forall x, y, \neg\text{Friends}(x,y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$$

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

[Example from Domingos and Lowd 2009]

Markov Logic Networks

From rules to graphical models

- Convert to clauses

$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$

$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$

→

$\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$

$\forall x, y, \neg\text{Friends}(x, y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$

- Associate a potential function for each clause
  - Think of each formula as a factor
  - Could be 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

$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$

1.5

$\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$

$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$

1.0

$\forall x, y, \neg\text{Friends}(x, y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$

Each rule is associated with a weight

[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

Weight	Formula	Clausal form
1.5	$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$	$\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$
1.0	$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$	$\forall x, y, \neg\text{Friends}(x, y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$

[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$

1.5

$\forall x, \neg\text{Smokes}(x) \lor \text{Cancer}(x)$

$\forall x, y \ \text{Friends}(x, y) \land \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$

1.0

$\forall x, y, \neg\text{Friends}(x, y) \lor \neg\text{Smokes}(y) \lor \text{Cancer}(y)$

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

[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

Formula	Weight	Clausal form
$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$	1.5	$\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$
$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$	1.0	$\forall x, y, \neg\text{Friends}(x, y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$

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

Each predicate gets grounded a random variable, one for each object in the world.

So we will have predicates such as Smokes(A), Cancer(A), Smokes(B), Friends(A, B)…

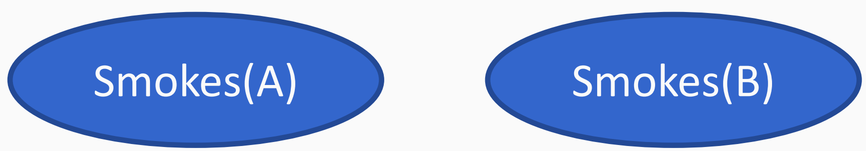
[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

1.5 $\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$ → $\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$

1.0 $\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$ → $\forall x, y, \neg\text{Friends}(x, y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$

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

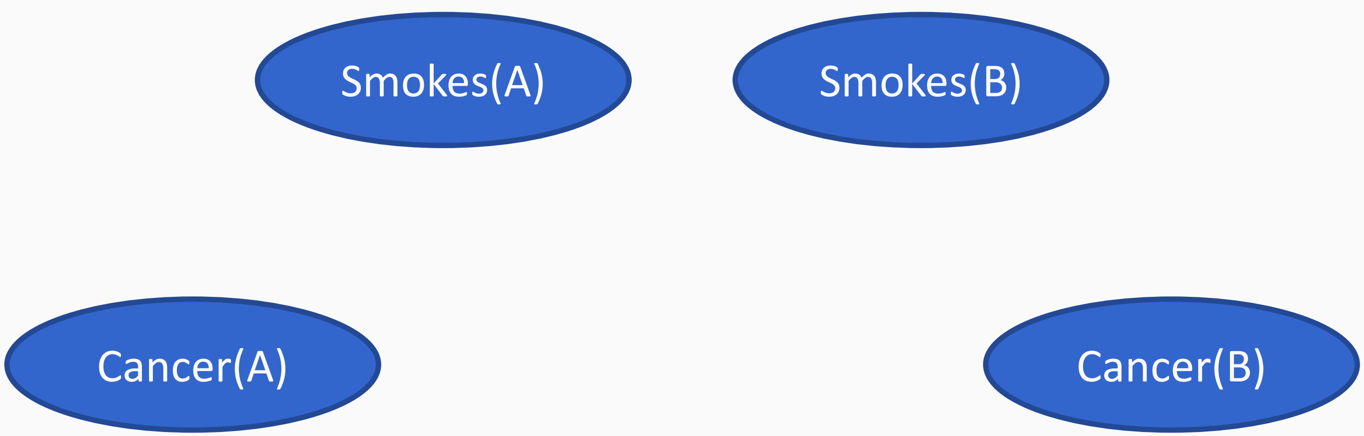


[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

Formula	Weight	Clausal form
$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$	1.5	$\forall x, \neg\text{Smokes}(x) \lor \text{Cancer}(x)$
$\forall x, y\ \text{Friends}(x, y) \land \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$	1.0	$\forall x, y, \neg\text{Friends}(x, y) \lor \neg\text{Smokes}(y) \lor \text{Cancer}(y)$

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



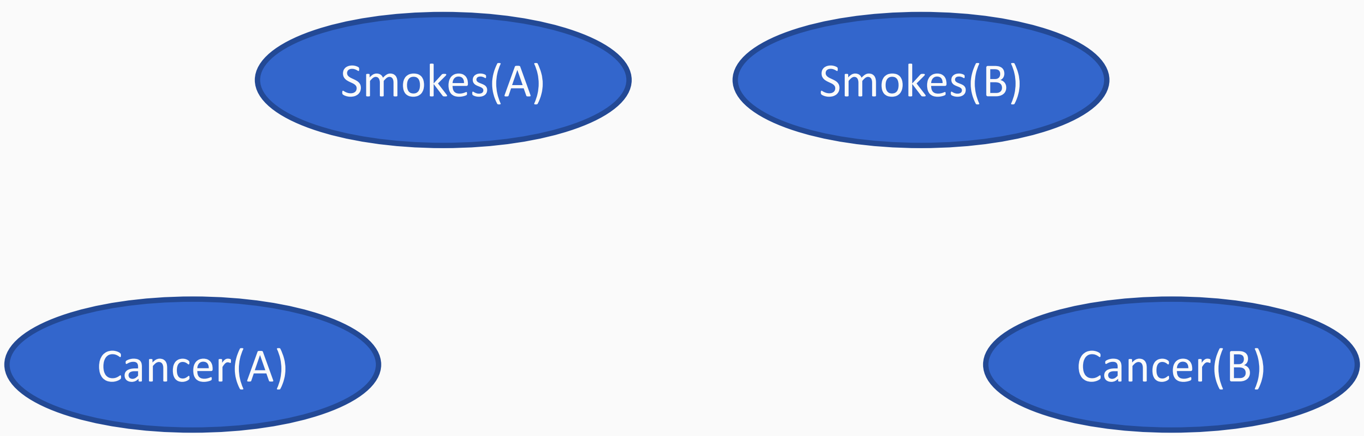
[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$	1.5	$\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$
$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$	1.0	$\forall x, y, \neg\text{Friends}(x, y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$

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

Each clause becomes a factor that connects the associated random variables



[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

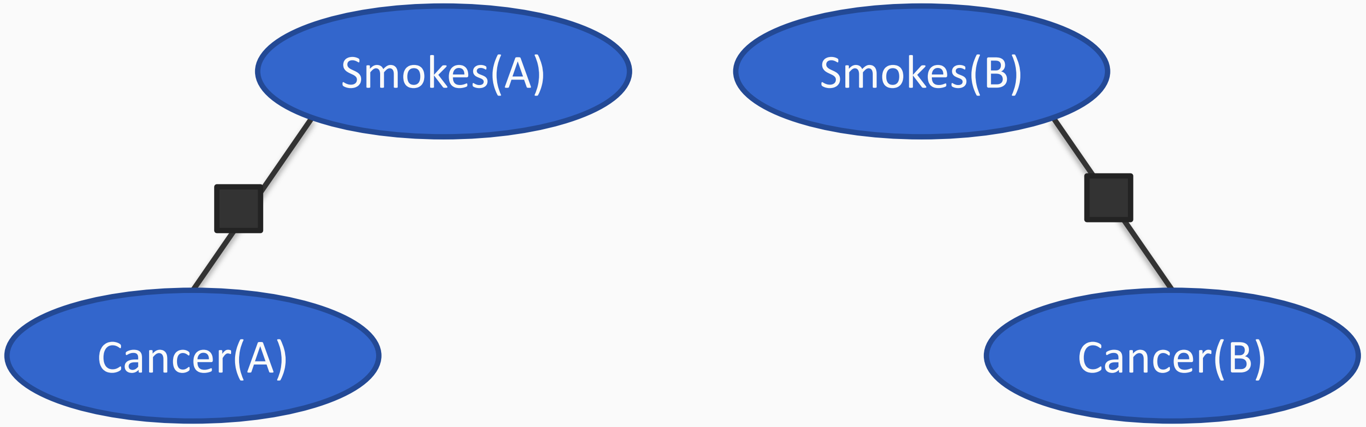
$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$

$\forall x, y \; \text{Friends}(x, y) \land \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$

1.5 $\forall x, \neg\text{Smokes}(x) \lor \text{Cancer}(x)$

1.0 $\forall x, y, \neg\text{Friends}(x, y) \lor \neg\text{Smokes}(y) \lor \text{Cancer}(y)$

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

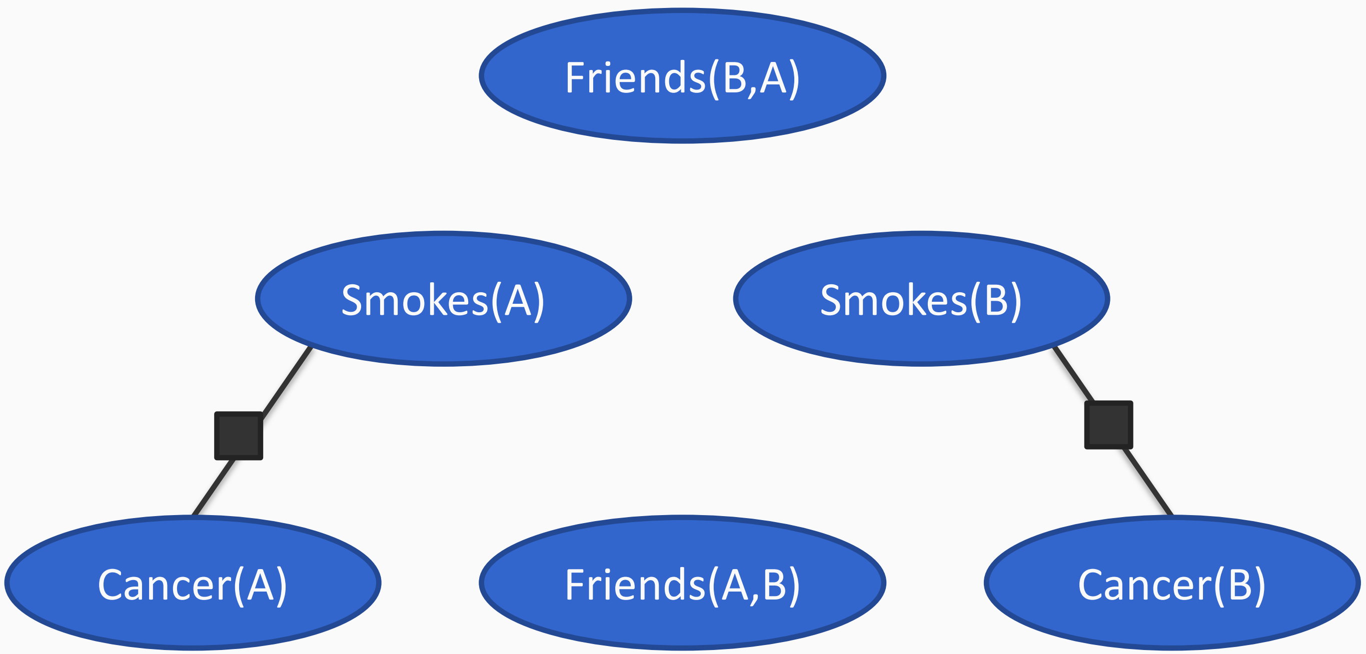


[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$	1.5	$\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$
$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$	1.0	$\forall x, y, \neg\text{Friends}(x, y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$

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

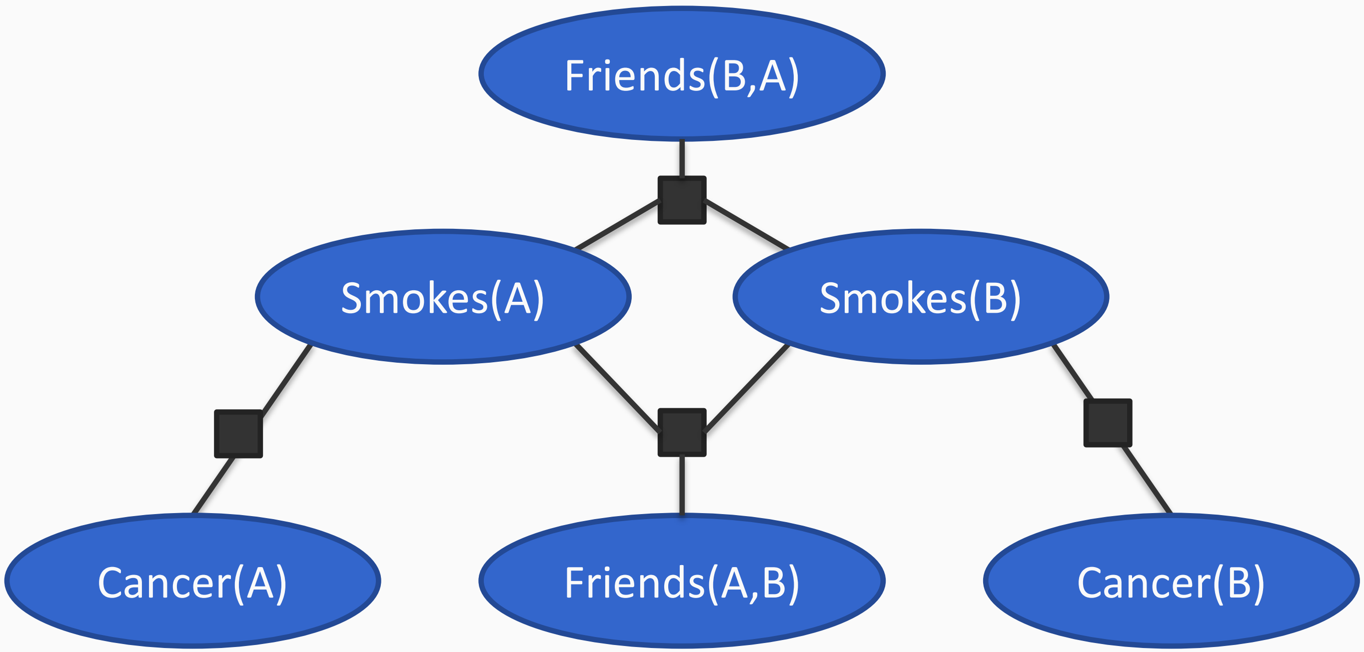


[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

Formula	Weight	Clausal form
$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$	1.5	$\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$
$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$	1.0	$\forall x, y, \neg\text{Friends}(x, y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$

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



[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$

$\forall x, y\ \text{Friends}(x, y) \land \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$

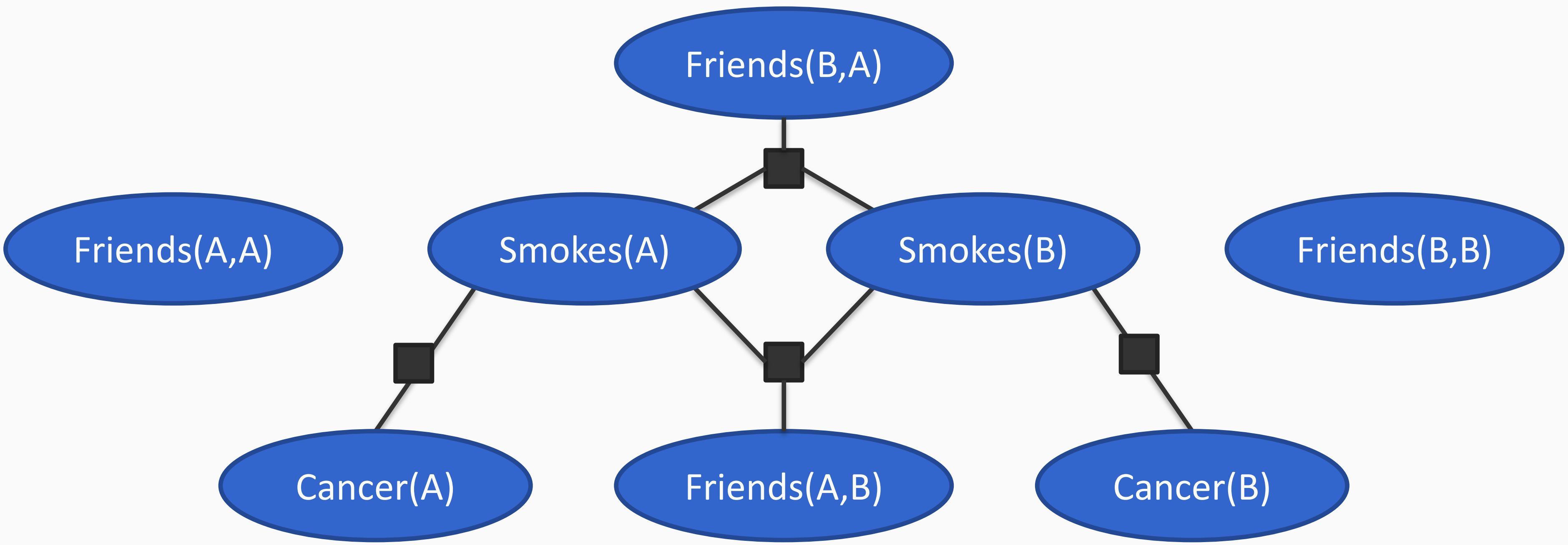
1.5

1.0

$\forall x, \neg\text{Smokes}(x) \lor \text{Cancer}(x)$

$\forall x, y, \neg\text{Friends}(x, y) \lor \neg\text{Smokes}(y) \lor \text{Cancer}(y)$

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

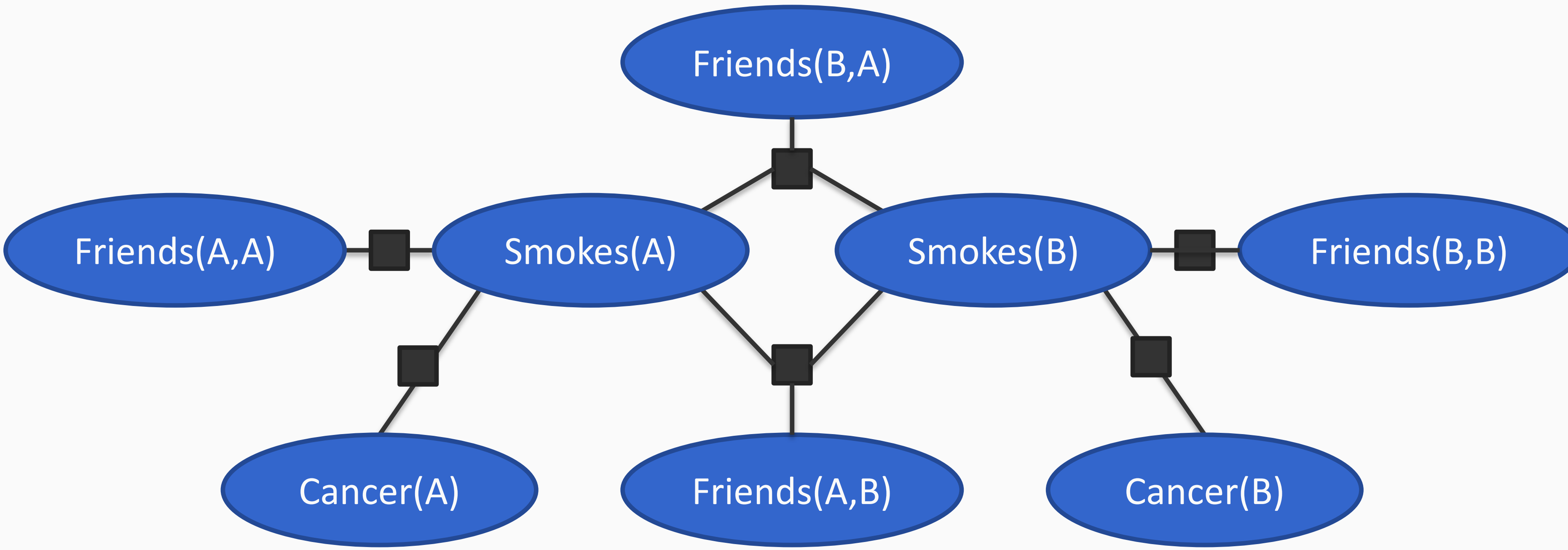


[Example from Domingos and Lowd 2009]

Weighted formulas →ground network

Formula	Weight	Clausal form
$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$	1.5	$\forall x, \neg\text{Smokes}(x) \vee \text{Cancer}(x)$
$\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$	1.0	$\forall x, y, \neg\text{Friends}(x, y) \vee \neg\text{Smokes}(y) \vee \text{Cancer}(y)$

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



[Example from Domingos and Lowd 2009]

Markov logic: Templated MRFs

1.5 $\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$

1.0 $\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$

World = {Anna (A), Bob (B)}

[Example from Domingos and Lowd 2009]

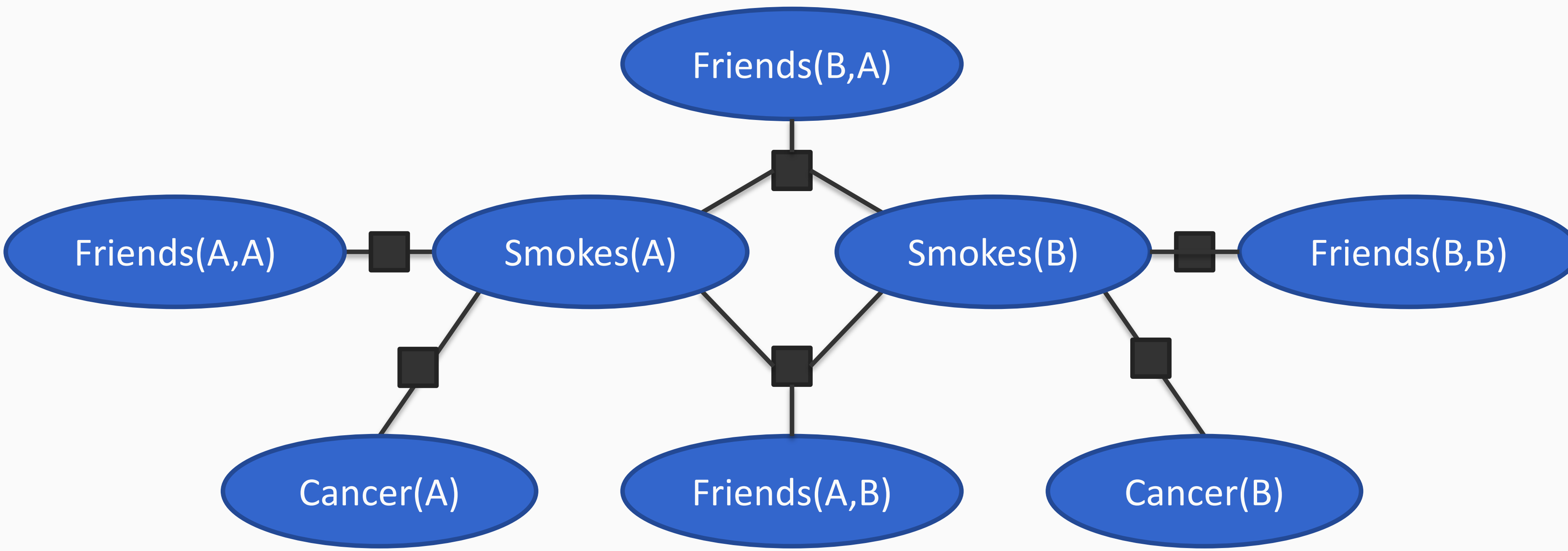
Markov logic: Templated MRFs

1.5 $\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$

1.0 $\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$

World = {Anna (A), Bob (B)}

Short hand notation for a large factor graph



[Example from Domingos and Lowd 2009]

Markov logic: Templated MRFs

1.5 $\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$

1.0 $\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$

World = {Anna (A), Bob (B)}

Short hand notation for a large factor graph

$$P(\text{assignment}) \propto \exp\left(\sum_i w_i n_i(\text{assignment})\right)$$

[Example from Domingos and Lowd 2009]

Markov logic: Templated MRFs

1.5 $\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$

1.0 $\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$

World = {Anna (A), Bob (B)}

Short hand notation for a large factor graph

$$P(\text{assignment}) \propto \exp\left(\sum_i w_i n_i(\text{assignment})\right)$$

Weight for the i^{th} formula

[Example from Domingos and Lowd 2009]

Markov logic: Templated MRFs

1.5 $\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$

1.0 $\forall x, y\ \text{Friends}(x, y) \wedge \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$

World = {Anna (A), Bob (B)}

Short hand notation for a large factor graph

Friend ,B)

$$P(\text{assignment}) \propto \exp\left(\sum_i w_i n_i(\text{assignment})\right)$$

Weight for the i^{th} formula

Number of factors where the i^{th} formula holds (i.e. number of true groundings for the formula)

[Example from Domingos and Lowd 2009]

Markov Logic Networks

From rules to graphical models

- Convert to clauses

- Ground the logical expressions to all variables that you care about

- Associate a potential function for each clause
  - Each formula is a factor
    - Could be log-linear in all the variables involved

[Example from Domingos and Lowd 2009]

Markov Logic: A different perspective

Suppose we have a knowledge base (KB)

- Standard logic: The KB constraints the set of possible worlds
  - The rules in the knowledge base are hard constraints that rule out certain assignments to the predicates

Markov Logic: A different perspective

Suppose we have a knowledge base (KB)

- Standard logic: The KB constraints the set of possible worlds
  - The rules in the knowledge base are hard constraints that rule out certain assignments to the predicates

$$\forall x, \text{Smokes}(x) \Rightarrow \text{Cancer}(x)$$

$$\forall x, y \ \text{Friends}(x, y) \land \text{Smokes}(x) \Rightarrow \text{Smokes}(y)$$

In a world where Smokes(Bob) and Friends(Anna, Bob) holds,
Cancer(Anna) is *forced* to be true

Markov Logic: A different perspective

Suppose we have a knowledge base (KB)

- Standard logic: The KB constraints the set of possible worlds
  - The rules in the knowledge base are hard constraints that rule out certain assignments to the predicates

- Markov logic: Each rule is a soft constraint
  - Worlds that violate these rules are allowed, but improbable
  - Each formula has a weight, formulas with higher weights are stronger constraints
    - Formulas with infinite weights are hard constraints
  - The probability of a world (i.e an assignment to the random variables) is proportional to $\exp(\sum \text{weights of clauses it satisfies})$

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

Summary: 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 model for the joint distribution
  - There is no separation of the variables as “inputs” and “outputs”
  - Unlike conditional random fields, for example