Intelligent Systems (AI-2)

Computer Science cpsc422, Lecture 29

Nov, 18, 2019

Slide source: from Pedro Domingos UW

422 big picture

StarAI (statistical relational AI)

Hybrid: Det +Sto

Prob CFG
Prob Relational Models
Markov Logics

Deterministic | **Stochastic**

Query

Logics
First Order Logics
Ontologies

- Full Resolution
- SAT

Belief Nets

Approx. : Gibbs

Markov Chains and HMMs

Forward, Viterbi….
Approx. : Particle Filtering

Undirected Graphical Models
Markov Networks
Conditional Random Fields

Planning

Markov Decision Processes and Partially Observable MDP

- Value Iteration
- Approx. Inference

Reinforcement Learning

Applications of AI

Representation

Reasoning Technique

Lecture Overview

- **Statistical Relational (Star-AI) Models (*for us aka Hybrid*)**
- **Recap Markov Networks and log-linear models**
- **Markov Logic**

Statistical Relational Models

Goals:

- Combine **(subsets of) logic** and **probability** into a single language (**R&R** system)
- Develop efficient **inference** algorithms
- Develop efficient **learning** algorithms
- Apply to real-world problems

L. Getoor & B. Taskar (eds.), *Introduction to Statistical Relational Learning,* MIT Press, 2007.

Plethora of Approaches

- Knowledge-based model construction [Wellman et al., 1992]
- Stochastic logic programs [Muggleton, 1996]
- Probabilistic relational models [Friedman et al., 1999]
- Relational Markov networks [Taskar et al., 2002]
- Bayesian logic [Milch et al., 2005]
- Markov logic [Richardson & Domingos, 2006]
- *And many others….!*

Prob. Rel. Models vs. Markov Logic

PRM

- Relational Skeleton
- Dependency Graph
- Parameters (CPT)

$\Rightarrow$ BNET

ML

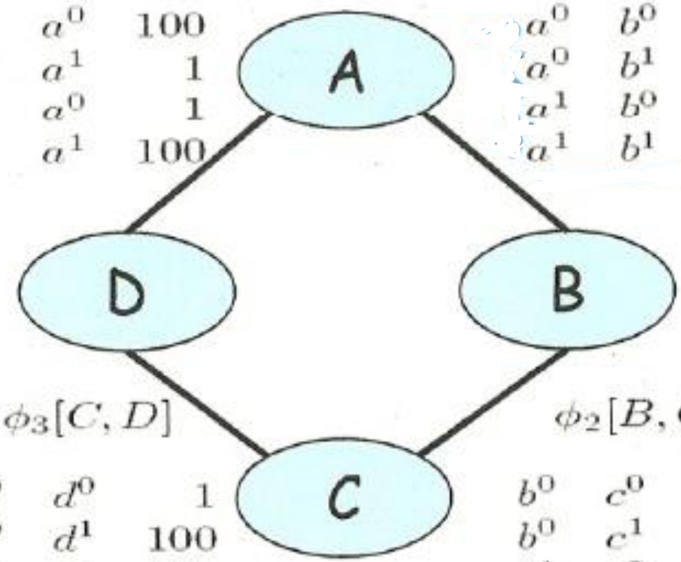
- weighted logical formulas
- set of constants

$\Rightarrow$ MARKOV LOGIC NETWORK

Lecture Overview

- Statistical Relational Models (*for us aka Hybrid*)
- **Recap Markov Networks and log-linear models**
- Markov Logic
  - Markov Logic Network (MLN)

Parameterization of Markov Networks



$\phi_4[D, A]$

d^0	a^0	100
d^0	a^1	1
d^1	a^0	1
d^1	a^1	100

$\phi_1[A, B]$

a^0	b^0	30
a^0	b^1	5
a^1	b^0	1
a^1	b^1	10

$\phi_3[C, D]$

c^0	d^0	1
c^0	d^1	100
c^1	d^0	100
c^1	d^1	1

$\phi_2[B, C]$

b^0	c^0	100
b^0	c^1	1
b^1	c^0	1
b^1	c^1	100

X set of random vars: A factor is

$\Phi(Val(X)) \rightarrow \mathbb{R}$

Factors define the local interactions (like CPTs in Bnets)

What about the global model? What do you do with Bnets?

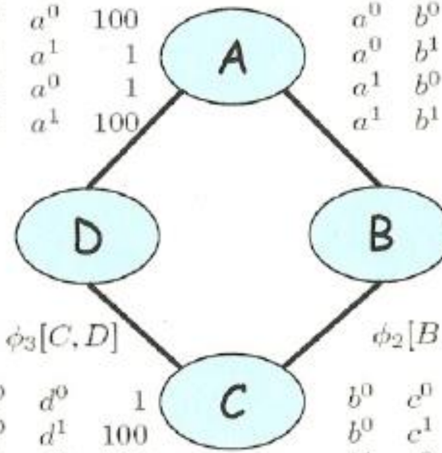
How do we combine local models?

As in BNets by multiplying them!

$$\tilde{P}(A,B,C,D) = \phi_1(A,B) \times \phi_2(B,C) \times \phi_3(C,D) \times \phi_4(A,D)$$

$$P(A,B,C,D) = \frac{1}{Z}\tilde{P}(A,B,C,D)$$

Assignment				Unnormalized	Normalized
a^0	b^0	c^0	d^0	300000	.04
a^0	b^0	c^0	d^1	300000	.04
a^0	b^0	c^1	d^0	300000	.04
a^0	b^0	c^1	d^1	30	4.1×10^{-6}
a^0	b^1	c^0	d^0	500	
a^0	b^1	c^0	d^1	500	
a^0	b^1	c^1	d^0	5000000	.69
a^0	b^1	c^1	d^1	500	
a^1	b^0	c^0	d^0	100	
a^1	b^0	c^0	d^1	1000000	
a^1	b^0	c^1	d^0	100	
a^1	b^0	c^1	d^1	100	
a^1	b^1	c^0	d^0	10	
a^1	b^1	c^0	d^1	100000	
a^1	b^1	c^1	d^0	100000	
a^1	b^1	c^1	d^1	100000	



$\phi_4[D,A]$

d^0	a^0	100
d^0	a^1	1
d^1	a^0	1
d^1	a^1	100

$\phi_1[A,B]$

a^0	b^0	30
a^0	b^1	5
a^1	b^0	1
a^1	b^1	10

$\phi_3[C,D]$

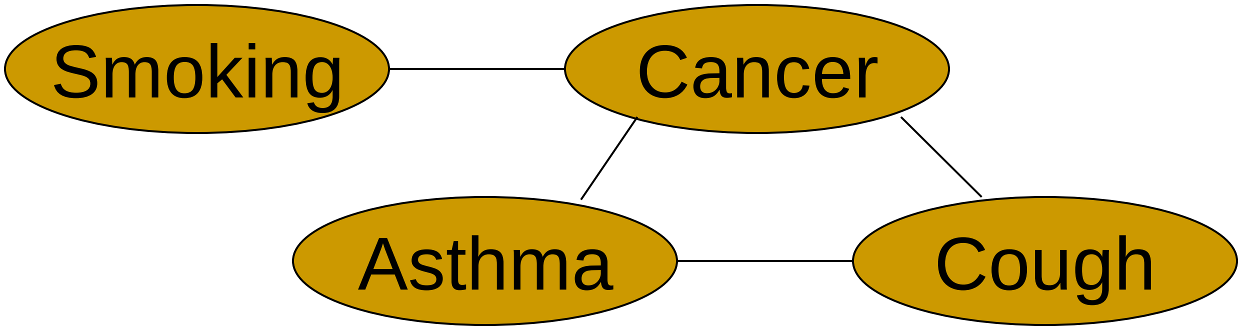
c^0	d^0	1
c^0	d^1	100
c^1	d^0	100
c^1	d^1	1

$\phi_2[B,C]$

b^0	c^0	100
b^0	c^1	1
b^1	c^0	1
b^1	c^1	100

Markov Networks

- **Undirected** graphical models



- Factors/Potential-functions defined over cliques

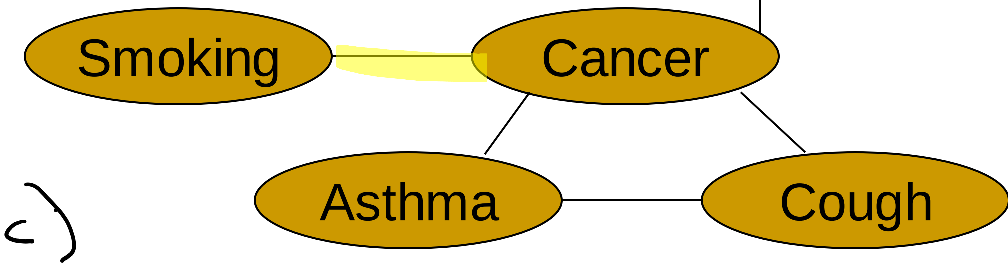
$$P(x) = \frac{1}{Z} \prod_c \Phi_c(x_c)$$

$$Z = \sum_x \prod_c \Phi_c(x_c)$$

Smoking	**Cancer**	**Φ(S,C)**
F	F	4.5
F	T	4.5
T	F	2.7
T	T	4.5

Markov Networks :log-linear model

$$P(x) = \frac{1}{Z} \prod_c \Phi_c(x_c)$$



- Log-linear model:

each $\Phi_c(x_c) = e^{w_c f_c(x_c)}$

$$w_1 = 0.51$$

$$f_1(\text{Smoking}, \text{Cancer}) = \begin{cases} 1 & \text{if } \neg \text{Smoking} \vee \text{Cancer} \\ 0 & \text{otherwise} \end{cases}$$

Smoking	Cancer
1	1
1	0
0	1
0	0

$$P(x) = \frac{1}{Z} \exp\left(\sum_i \boxed{w_i} \boxed{f_i(x_i)} \right)$$

Weight of Feature *i*

Feature *i*

Lecture Overview

- Statistical Relational Models (for us aka Hybrid)
- Recap Markov Networks
- **Markov Logic**

Markov Logic: Intuition(1)

- A logical KB is a set of **hard constraints** on the set of possible worlds

CONSTANT INDIVIDUALS = {a, b}

f $\boxed{\forall x\ Smokes(x) \Rightarrow Cancer(x)}$

$$\left.\begin{array}{l} Smokes(a) = T \\ Cancer(a) = F \\ Smokes(b) = F \\ Cancer(b) = F \end{array}\right\} \hat{w}$$

in FOL $\hat{w}$ is...

A. possible   B. impossible

C. cannot tell

iclicker.

Markov Logic: Intuition (2)

- The more formulas in the KB a possible world satisfies the more it should be likely
- Give each formula a **weight**
- Adopting a **log-linear model**, by design, if a possible world satisfies a formula its **probability** should go up proportionally to **exp(the formula weight)**.

$$\mathrm{P(world)} \propto \exp\left(\sum \text{weights of formulas it satisfies}\right)$$

That is, if a possible world satisfies a formula its **log probability** should go up proportionally to the formula weight.

$$\log(\mathrm{P(world)}) \propto \left(\sum \text{weights of formulas it satisfies}\right)$$

Markov Logic: Definition

- A Markov Logic Network (MLN) is
  - a set of pairs (F, w) where
    - F is a **formula** in first-order logic
    - w is a **real number**
  - Together with a set C of **constants**,
- It defines a **Markov network** with
  - One *binary node* for each **grounding** of each **predicate** in the MLN
  - One *feature/factor* for each **grounding** of each **formula** F in the MLN, with the corresponding weight w

Grounding: substituting vars with constants

Example: Friends & Smokers

Smoking causes cancer.

Friends have similar smoking habits.

Example: Friends & Smokers

$$\forall x\ Smokes(x) \Rightarrow Cancer(x)$$
$$\forall x, y\ Friends(x, y) \Rightarrow \left(Smokes(x) \Leftrightarrow Smokes(y)\right)$$

Example: Friends & Smokers

1.5	$\forall x\ Smokes(x) \Rightarrow Cancer(x)$
1.1	$\forall x, y\ Friends(x, y) \Rightarrow \left(Smokes(x) \Leftrightarrow Smokes(y)\right)$

Example: Friends & Smokers

1.5	$\forall x\ Smokes(x) \Rightarrow Cancer(x)$
1.1	$\forall x, y\ Friends(x, y) \Rightarrow \left(Smokes(x) \Leftrightarrow Smokes(y)\right)$

Two constants: **Anna** (A) and **Bob** (B)

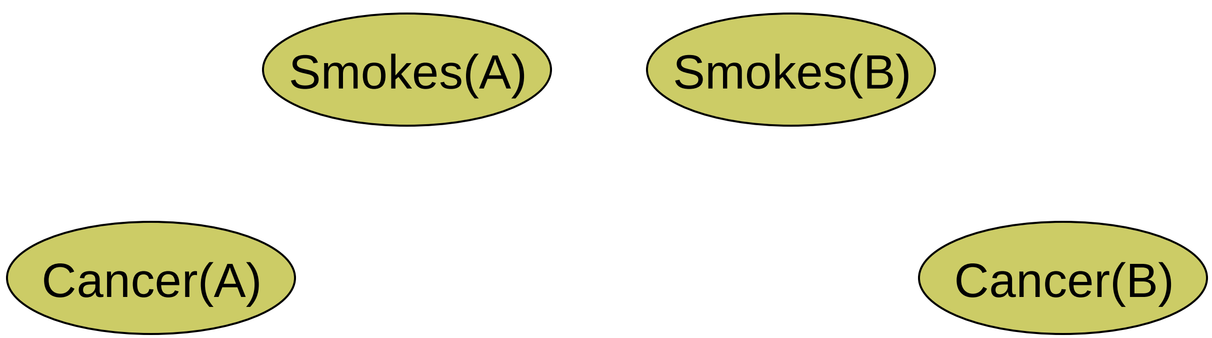
MLN nodes

1.5	$\forall x\ Smokes(x) \Rightarrow Cancer(x)$
1.1	$\forall x, y\ Friends(x, y) \Rightarrow \left(Smokes(x) \Leftrightarrow Smokes(y)\right)$

Two constants: **Anna** (A) and **Bob** (B)

- One *binary node* for each grounding of each predicate in the MLN

Grounding: substituting vars with constants



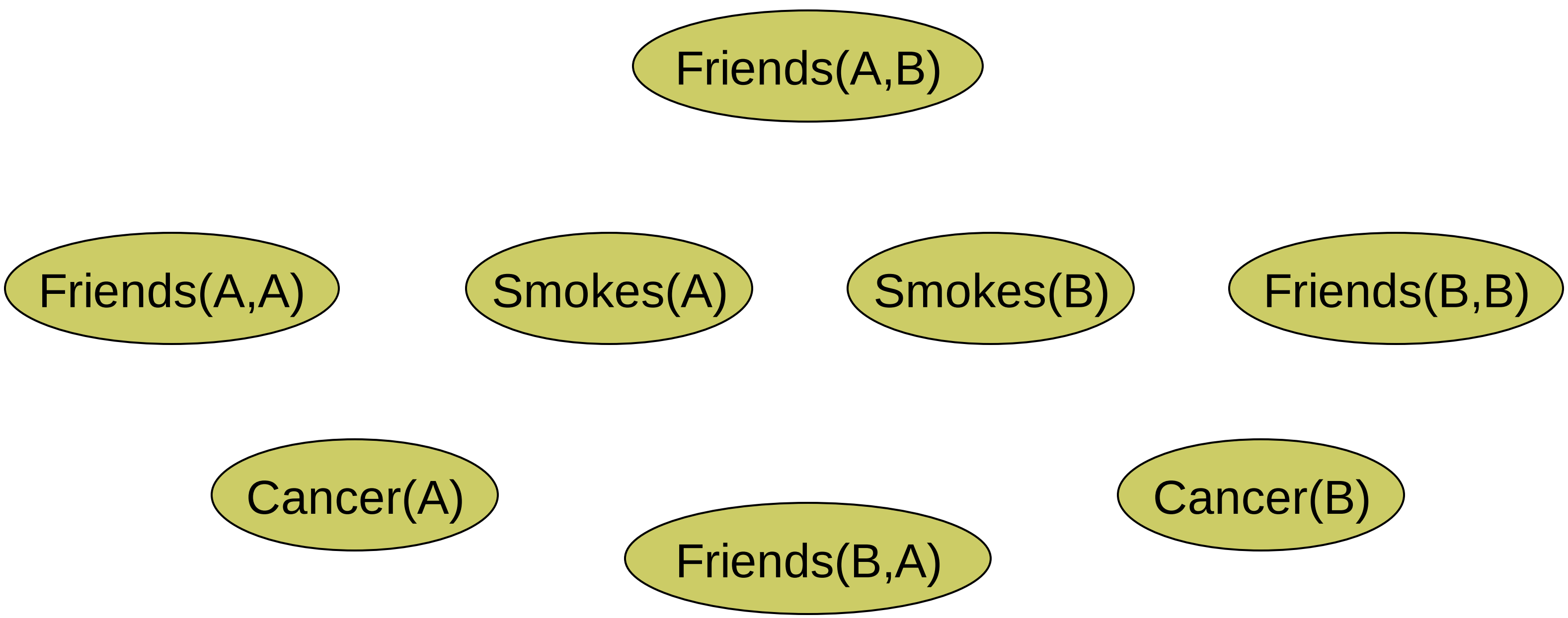
- Any nodes missing?

MLN nodes (complete)

1.5	$\forall x\ Smokes(x) \Rightarrow Cancer(x)$
1.1	$\forall x, y\ Friends(x, y) \Rightarrow \left(Smokes(x) \Leftrightarrow Smokes(y)\right)$

Two constants: **Anna** (A) and **Bob** (B)

- One *binary node* for each grounding of each predicate in the MLN



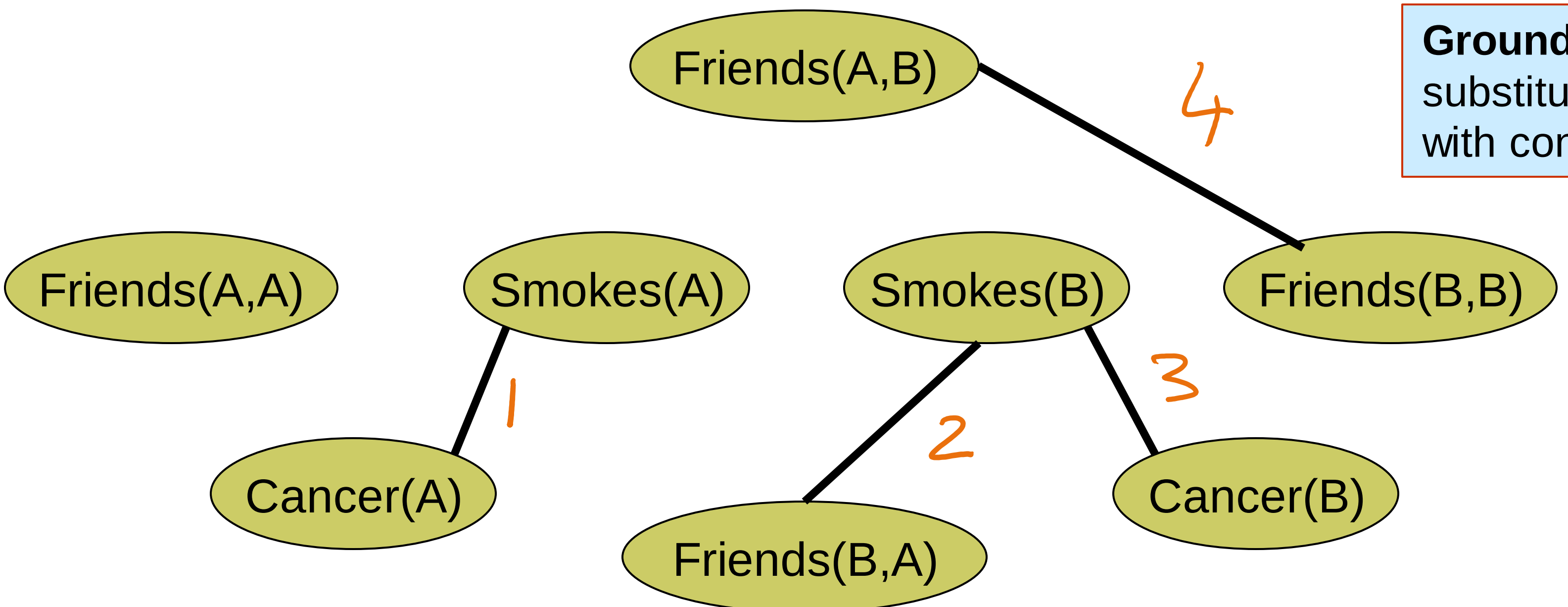
MLN features

1.5	$\forall x\ Smokes(x) \Rightarrow Cancer(x)$
1.1	$\forall x, y\ Friends(x, y) \Rightarrow \left(Smokes(x) \Leftrightarrow Smokes(y)\right)$

Two constants: **Anna** (A) and **Bob** (B)

Edge between two nodes iff the corresponding ground predicates appear together in at least one grounding of one formula

Grounding: substituting vars with constants



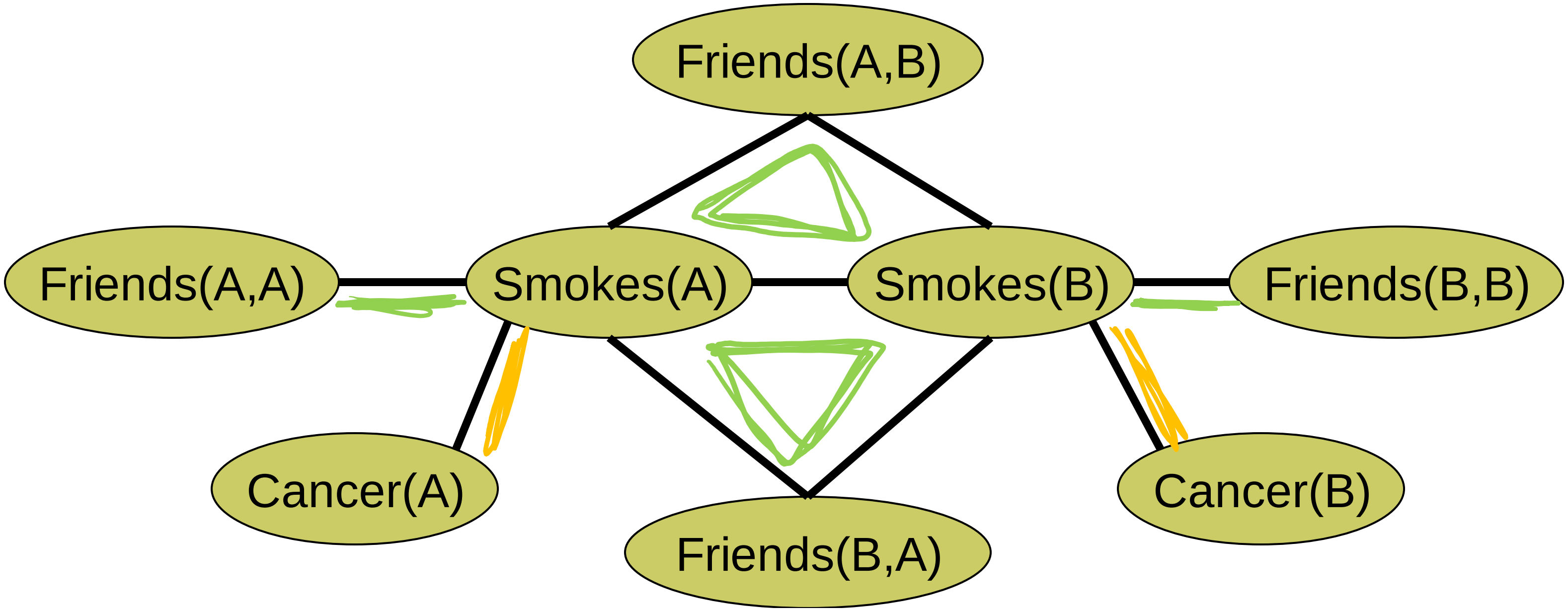
Which edge should not be there?

A. 1 B. 2 C. 3 D. 4

MLN features

1.5	$\forall x\ Smokes(x) \Rightarrow Cancer(x)$
1.1	$\forall x, y\ Friends(x,y) \Rightarrow \left(Smokes(x) \Leftrightarrow Smokes(y)\right)$

Two constants: **Anna** (A) and **Bob** (B)



One *feature/factor* for each **grounding** of each **formula** F in the MLN

MLN: parameters

- For each formula *i* we have a factor

$\Phi_i(pw) = e^{w_i f_i(pw)}$ ← possible world

w_i weight of formula

$$f_i(pw) = \begin{cases} 1 & \text{when formula is true in } pw \\ 0 & \text{otherwise} \end{cases}$$

$$1.5 \quad \forall x\ Smokes(x) \Rightarrow Cancer(x)$$

$$f\,(\text{Smokes(x)},\ \text{Cancer(x)}) = \begin{cases} 1 & \text{if Smokes (x)} \Rightarrow \text{Cancer (x)} \\ 0 & \text{otherwise} \end{cases}$$

pw_1 ....

Smokes(A) T

Cancer(A) F $e^0 = 1$

pw_2 ....

Smokes(A) T

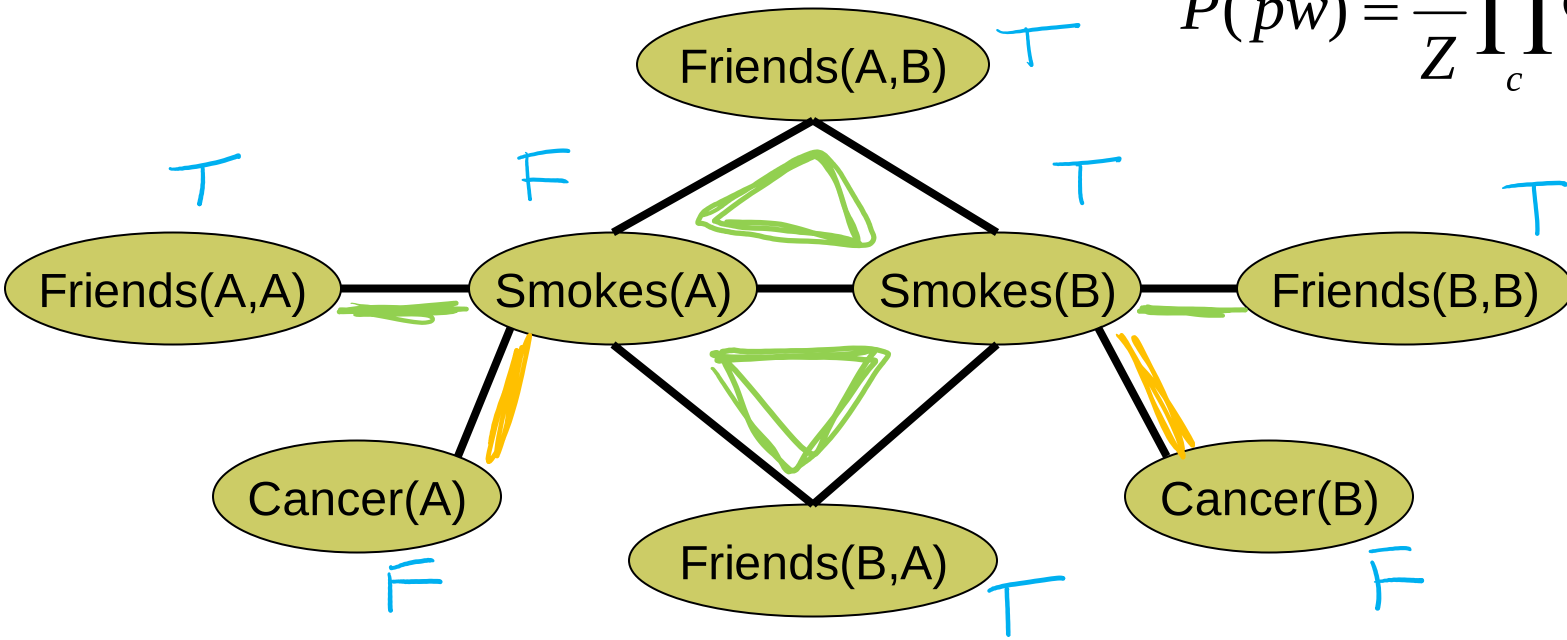
Cancer(A) T $e^{1.5}$

MLN: prob. of possible world

1.5	$\forall x\ Smokes(x) \Rightarrow Cancer(x)$
1.1	$\forall x, y\ Friends(x, y) \Rightarrow \big(Smokes(x) \Leftrightarrow Smokes(y)\big)$

Two constants: **Anna** (A) and **Bob** (B)

$$P(pw) = \frac{1}{Z} \prod_c \Phi_c(pw_c)$$



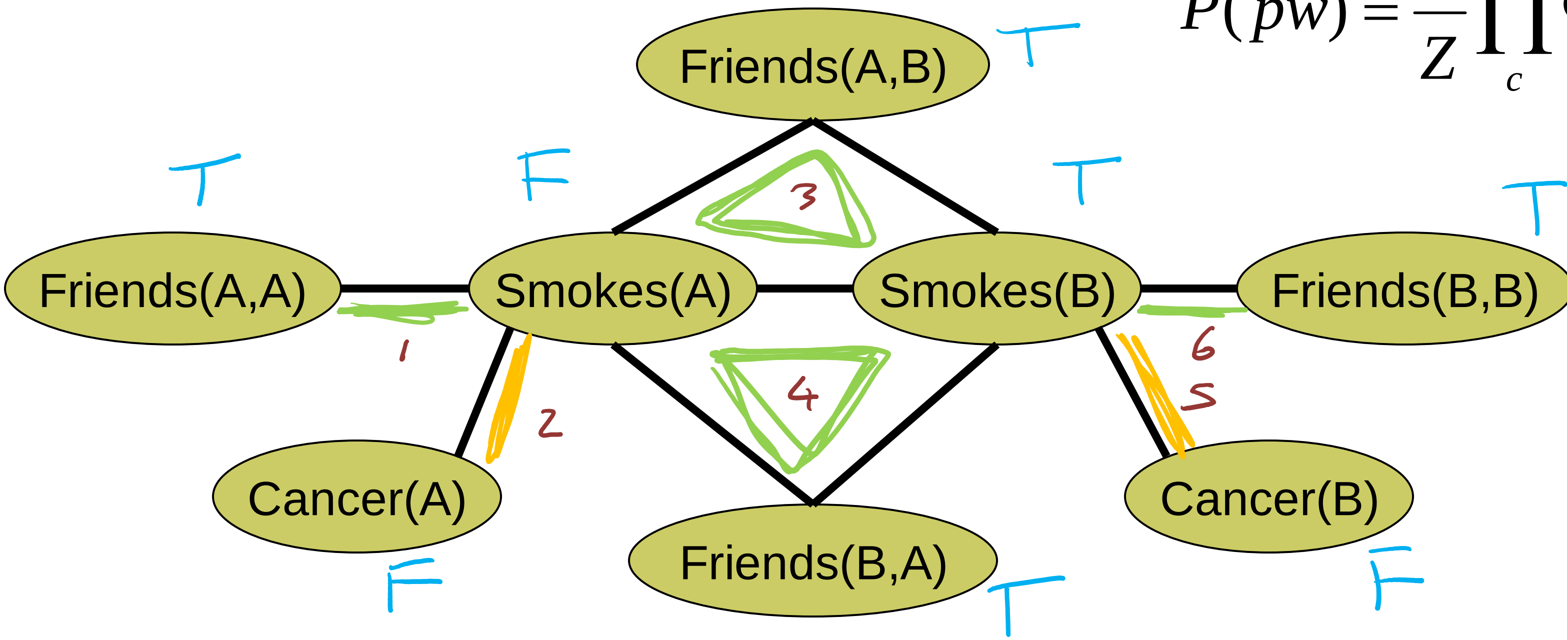
$$P(pw) = \left(e^{1.1} * e^{1.1} * e^{0} * e^{0} * e^{1.5} * e^{0}\right) / Z$$

MLN: prob. of possible world

1.5	$\forall x\ Smokes(x) \Rightarrow Cancer(x)$
1.1	$\forall x, y\ Friends(x,y) \Rightarrow \big(Smokes(x) \Leftrightarrow Smokes(y)\big)$

Two constants: **Anna** (A) and **Bob** (B)

$$P(pw) = \frac{1}{Z} \prod_c \Phi_c(pw_c)$$



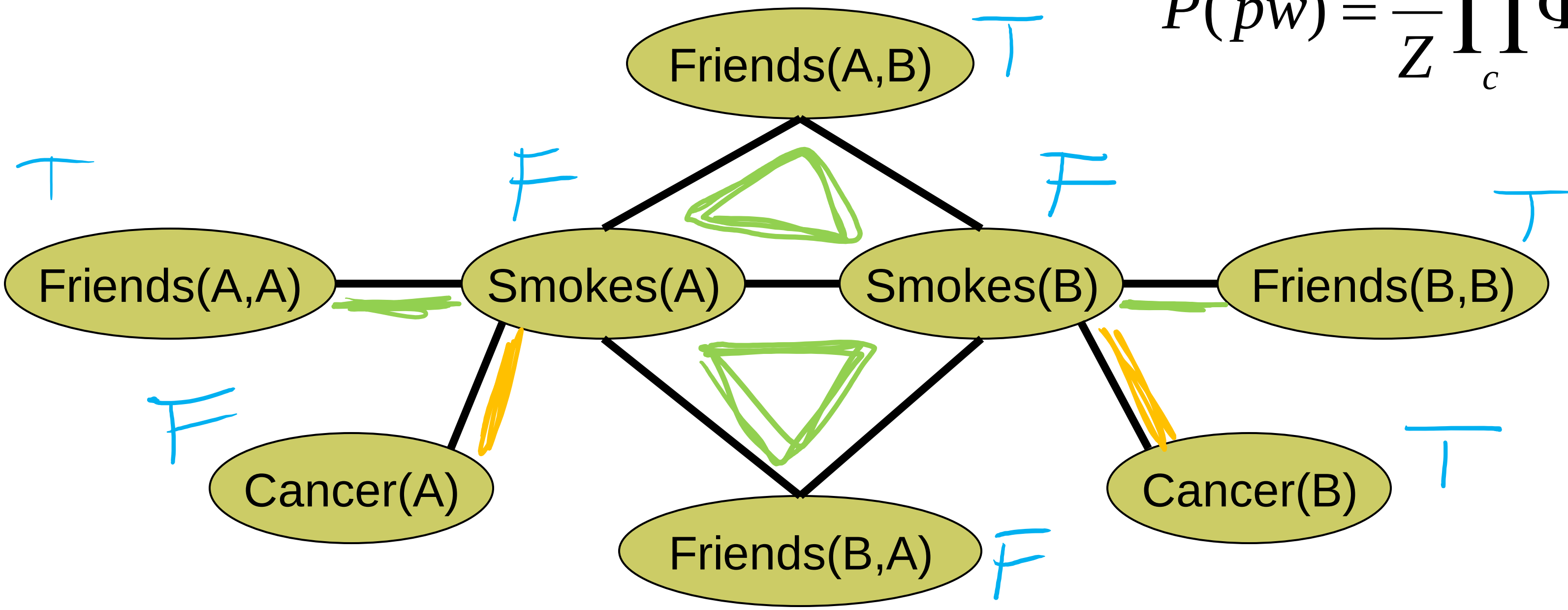
$$P(pw) = \left(\underset{1}{e^{1.1}} * \underset{6}{e^{1.1}} * \underset{3}{e^{0}} * \underset{4}{e^{0}} * \underset{2}{e^{1.5}} * \underset{5}{e^{0}} \right) / Z$$

MLN: prob. of possible world

1.5	$\forall x \; Smokes(x) \Rightarrow Cancer(x)$
1.1	$\forall x, y \; Friends(x, y) \Rightarrow (Smokes(x) \Leftrightarrow Smokes(y))$

Two constants: **Anna** (A) and **Bob** (B)

$$P(pw) = \frac{1}{Z} \prod_c \Phi_c(pw_c)$$



$$\left(\qquad \right) \frac{1}{Z}$$

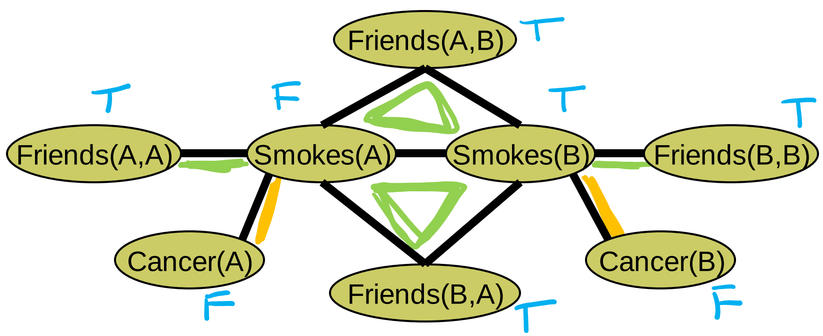
MLN: prob. Of possible world

- Probability of a world pw:

$$P(pw) = \frac{1}{Z}\exp\left(\sum_i \boxed{w_i}\,\boxed{n_i(pw)}\right)$$

Weight of formula *i*

No. of true groundings of formula *i* in *pw*



$$P(pw) = \left(\underbrace{e^{1.1} * e^{1.1}}_{n_2(pw)=2} * e^{0} * e^{0} * \underbrace{e^{1.5}}_{n_1(pw)=1} * e^{0}\right) / Z$$

$$\mathrm{P(world)} \propto \exp\left(\sum \text{weights of grounded formulas it satisfies}\right)$$

Learning Goals for today's class

You can:

- Describe the intuitions behind the design of a Markov Logic
- Define and Build a Markov Logic Network
- Justify and apply the formula for computing the probability of a possible world

Next class on Wed

Markov Logic

-relation to FOL

- Inference (MAP and Cond. Prob)

Assignment-4 will be posted this evening, due on Nov 29