KRR with Uncertainty

Markov Logic

Ira A. Fulton Schools of Engineering
Arizona State University

Objectives

Objective

Explain the different roles of logic and probability in KR

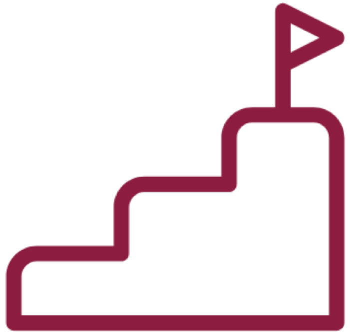
Objective

Explain the benefits of combining logic and probability in KR

Statistical Relational Learning

Goals

- Combine logic and probability into a single language
- Develop efficient inference algorithms
- Develop efficient learning algorithms
- Apply to real-world problems



Applications

Collective classification

- Determine labels for a set of objects (such as Web pages) based on their attributes as well as their relations to one another

Social network analysis and link prediction

- Predict relations between people based on attributes, attributes based on relations, cluster entities based on relations, etc. (smoker example)

Entity resolution

- Determine which observations imply real-world objects (Deduplicating a database)

etc.

Markov Logic: Intuition (1 of 3)

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

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

- Consider the Herbrand universe {a,b} and the Herbrand interpretation w = {Smokes(a)}

Smokes(a) → cancer(a)

smokes (b) → cancer (b)

- **Q**: In FOL, w is
  - A: Possible
  - B: Impossible
  - C: Cannot tell

Markov Logic: Intuition (2 of 3)

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

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

- Consider the Herbrand universe is is {a,b} and the Herbrand interpretation w = {Smokes(a)}

- Let's make them soft constraints:
  - When a world violates a formula, the world becomes less probable, not impossible

Markov Logic: Intuition (3 of 3)

- The more formulas in the KB a possible world satisfies the more it should be likely

- Give each formula a **weight**

- By design, if a possible world satisfies a formula its **log probability** should go up proportionally to the formula weight.

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

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

Markov Logic

- Markov logic combines first-order logic with Markov networks
- A Markov Logic Network (MLN) consists of set of weighted first-order formulas
- The probability of a world is proportional to the exponential of the sum of the formulae that are true in the world.
- The idea is to view logical formulas as soft constraints on the set of possible worlds.
  - A positive weight on a formula increases the probability of the worlds that satisfy the formula,
  - A negative weight on a formula decreases the probability of the worlds that satisfy the formula.

Markov Logic: Syntax (1 of 2)

- Assume first-order signature σ that has no function constants of positive arity and finitely many object constants
- A MLN L of signature σ is a finite set of weighted formulas of the form $\mathrm{w}{:}\mathrm{F}$ where F is a first-order formula of σ and w is a real number
- We say that L is ground if its formulas contain no variables

Markov Logic: Syntax (2 of 2)

- **Any (non-ground) MLN $\mathbb{L}$ of signature σ can be identified with the ground MLN by turning each formula in $\mathbb{L}$ into a set of "propositional images": (roughly)**
  - replacing $\exists x\, F(x)$ with $F(c_1) \vee \cdots \vee F(c_n)$
  - replacing $\forall x\, F(x)$ with $F(c_1) \wedge \cdots \wedge F(c_n)$

  and then
  - replacing w: $F(x)$ with w: $F(c_1), \dots, w{:}\, F(c_n)$

Markov Logic as Weighted Logic

- An MLN L of signature σ is a finite set of "weighted formula" $\mathrm{w} : \mathrm{F}$, where F is a first-order formula of σ and w is a real number.

- We say that an MLN is **ground** if its formulas contain no variables.

- We assume that the underlying Herbrand universe is finite.

- The weight of an interpretation I for ground MLN is
  - $\mathrm{W(I)} = \exp\left(\sum_{\substack{w:F \\ I \models F}} w\right)$
  - $\mathrm{P(I)} = \cancel{\lim_{\alpha\to\infty}} \frac{W(I)}{\Sigma_J W(J)}$
    - where J ranges over all Herbrand interpretations of σ

Example

- w: $\boxed{\forall x\ Smokes(x) \Rightarrow Cancer(x)}$

ω : Smokes(a) → Cancer(a)

- Consider the Herbrand universe is {a}
  - I1 = {} P(I1) = $\frac{e^{\omega}}{3 \cdot e^{\omega} + 1}$
  - I2= {Cancer(a)} P(I2) = $P(I_1)$
  - I3= {Smokes(a)} P(I3) = $\frac{e^{0}}{3 \cdot e^{\omega} + 1}$
  - I4= {Smokes(a), Cancer(a)} P(I4) = $P(I_1)$

$\omega(I_1) = e^{\omega}$

$\omega(I_2) = e^{\omega}$

$\omega(I_3) = e^{0}$

$\omega(I_4) = e^{\omega}$

- Z = $3 \cdot e^{\omega} + 1$
- P(Cancer(a)| Smokes(a)) = $\frac{e^{\omega}}{e^{\omega} + 1}$
- P(Smokes(a)∧ ¬Cancer(a)) = $\frac{1}{3 \cdot e^{\omega} + 1}$

Example: Friends & Smokers (1 of 3)

Smoking causes cancer.

Friends have similar smoking habits.

Example: Friends & Smokers (2 of 3)

1.5	$\cancel{\forall x}\ Smokes(x) \Rightarrow Cancer(x)$
1.1	$\cancel{\forall x,y}\ Friends(x,y) \Rightarrow (Smokes(x) \Leftrightarrow Smokes(y))$

- Two constants: **Anna** (A) and **Bob** (B)
  - 1.5 Smokes(A) => Cancer(A) ✓
  - 1.5 Smokes(B) => Cancer(B) ✗
  - 1.1 Friends(A,A) => (Smokes(A) <=> Smokes(A) ✓
  - 1.1 Friends(B,B) => (Smokes(B) <=> Smokes(B)) ✓
  - 1.1 Friends(A,B) => (Smokes(A) <=> Smokes(B)) ✗
  - 1.1 Friends(B,A) => (Smokes(B) <=> Smokes(A)) ✗

$$W(I) = \exp\left(\sum_{\substack{w:F \\ I \vDash F}} w\right)$$

I = {Friends(A,A), Friends(A,B), Friends(B,A), Friend(B,B), Smokes(B)}

$$P(I) = \frac{e^{1.5+1.1+1.1}}{Z}$$

Example: Friends & Smokers (3 of 3)

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)

- 1.5 Smokes(A) => Cancer(A) ✓
- 1.5 Smokes(B) => Cancer(B) ✓
- 1.1 Friends(A,A) => (Smokes(A) <=> Smokes(A) ✓
- 1.1 Friends(B,B) => (Smokes(B) <=> Smokes(B)) ✓
- 1.1 Friends(A,B) => (Smokes(A) <=> Smokes(B)) ✓
- 1.1 Friends(B,A) => (Smokes(B) <=> Smokes(A)) ✓

I = {Friends(A,A), Friends(A,B), Friend(B,B), Cancer(B)}

P(I) = $\frac{e^{1.5\times 2 + 1.1\times 4}}{Z}$

$$W(I) = \exp(\sum_{\substack{w:F \\ I \models F}} w)$$

Wrap-Up