A Short Introduction to Probabilistic Soft Logic

Angelika Kimmig, Stephen H. Bach, Matthias Broecheler, Bert Huang and Lise Getoor

NIPS Workshop on Probabilistic Programming 2012

http://psl.umiacs.umd.edu
Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]
Probabilistic Soft Logic (PSL) [Broecheler et al, UAI 10]

Declarative language to specify graphical models
Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]

Declarative language to specify graphical models

- Logical atoms with soft truth values in $[0,1]$
Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]

Declarative language to specify graphical models

• Logical atoms with soft truth values in [0,1]

• Dependencies as weighted first order rules
Probabilistic Soft Logic (PSL)  
[Broecheler et al, UAI 10]

Declarative language to specify graphical models

- Logical atoms with soft truth values in $[0, 1]$  
- Dependencies as weighted first order rules  
- Support for similarity functions and aggregation
Probabilistic Soft Logic (PSL)  
[Broecheler et al, UAI 10]

Declarative language to specify graphical models

• Logical atoms with soft truth values in $[0,1]$  
• Dependencies as weighted first order rules  
• Support for similarity functions and aggregation  
• Linear (in)equality constraints
Probabilistic Soft Logic (PSL)

[Broecheler et al, UAI 10]

Declarative language to specify graphical models

- Logical atoms with soft truth values in $[0,1]$ 
- Dependencies as weighted first order rules 
- Support for similarity functions and aggregation 
- Linear (in)equality constraints

Efficient MPE inference: continuous convex optimization
Applications

- Collective classification
- Ontology alignment
- Entity resolution
- Link prediction
- Trust in social networks
- Social group modeling
- Personalized medicine
- ...

Ontology Alignment

similar names
Ontology Alignment

similar ranges

similar names

Service & Products
- Software
- Hardware
- IT Services

Organization
- provides
- develops

Customers
- interacts
- helps
- sells to

Employees
- work for
- interacts
- helps
- sells to

Company
- develop
- works for

Products & Services
- buys

Customer
- interacts with
- helps
- sells

Employee
- Technician
- Sales
- Accountant

Software Dev
- Hardware
- Consulting
Ontology Alignment

similar ranges

similar names

similar subconcepts
Trust Modeling

dan 0.5  0.7  0.8  0.9  bob

carla 0.2  0.8  0.8  ann

emmma 0.4  0.9  0.5  fred
**Trust Modeling**

\[
\text{trusts}(X,Y) \land \text{trusts}(Y,Z) \rightarrow \text{trusts}(X,Z)
\]
trusts(X,Y) \land trusts(Y,Z) \rightarrow trusts(X,Z)
trusts(X,Y) \land trusts(Y,Z) \rightarrow trusts(X,Z)
Voter Opinion Modeling
Voter Opinion Modeling
Voter Opinion Modeling

dan - spouse - colleague - friend - ann - spouse

bob - friend

emma - spouse

friend

carla

friend

colleague
PSL Program

friend(carla, emma) = 0.9
friend(bob, dan) = 0.4
spouse(ann, bob) = 1.0
prefers(ann, ?) = 0.8
...

Diagram with relationships and labels for individuals.
0.3: lives(A,S) ∧ majority(S,P) → prefers(A,P)
0.8: spouse(B,A) ∧ prefers(B,P) → prefers(A,P)
0.1: similarAge(B,A) ∧ prefers(B,P) → prefers(A,P)
0.4: prefers(A,P) → prefersAvg({A.friend},P)

Partial-functional: prefers
Constraints

partial-functional: prefers
Constraints

partial-functional: prefers

\[ \text{prefers}(A, \text{Republican}) + \text{prefers}(A, \text{Democrat}) \leq 1.0 \]
Local Rules

0.3: $\text{lives}(A,S) \land \text{majority}(S,P) \rightarrow \text{prefers}(A,P)$
Propagation Rules

0.8: spouse(B,A) \land prefers(B,P) \rightarrow prefers(A,P)
0.1: similarAge(B,A) \land \text{prefers}(B,P) \rightarrow \text{prefers}(A,P)

Similarity function with range $[0,1]$
0.4: prefers(A,P) → prefersAvg({A.friend},P)
Sets

0.4: prefers(A,P) → prefersAvg({A.friend},P)

all X such that friend(A,X)
0.4: prefers(A,P) \rightarrow \text{prefersAvg}\{\text{A.friend}\}, P

\text{truth value} \equiv \text{average truth value of preferring}(X,P)
PSL Program

- Ground atoms = random variables
- Soft truth value assignments
- Assignment satisfying more rules more likely
- Constraints to rule out unwanted assignments
Probabilistic Model

\[ f(I) = \frac{1}{Z} \exp \left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right) \]
Probabilistic Model

\[ f(I) = \frac{1}{Z} \exp \left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right) \]

Interpretation
Probabilistic Model

\[ f(I) = \frac{1}{Z} \exp \left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right) \]

Interpretation

Set of rule groundings
Probabilistic Model

\[ f(I) = \frac{1}{Z} \exp \left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right) \]

- **Interpretation**
- **Rule’s weight**
- **Set of rule groundings**
Probabilistic Model

Ground rule’s distance from satisfaction given $I$

\[ f(I) = \frac{1}{Z} \exp \left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right) \]

- Interpretation
- Rule’s weight
- Set of rule groundings
Probabilistic Model

Ground rule’s distance from satisfaction given $I$

$$f(I) = \frac{1}{Z} \exp \left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right)$$

- Interpretation
- Rule’s weight
- Set of rule groundings
- $\in \{1, 2\}$
Probabilistic Model

Ground rule’s distance from satisfaction given \( I \)

\[
f(I) = \frac{1}{Z} \exp \left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(I))^k \right)
\]

Interpretation

Rule’s weight

Set of rule groundings

Normalization constant

\[
Z = \int_{J \in \mathcal{J}} \exp \left( - \sum_{r \in P} \sum_{g \in G(r)} w_r (d_g(J))^k \right)
\]
Distance from Satisfaction

\[ d_r(I) = \max\{0, I(\text{body}) - I(\text{head})\} \]
Distance from Satisfaction

\[ d_r(I) = \max\{0, I(body) - I(head)\} \]

\[ body \rightarrow head \text{ satisfied} \iff \]

truth value of body \(\leq\) truth value of head
Distance from Satisfaction

\[ d_r(I) = \max\{0, I(body) - I(head)\} \]

body \rightarrow head  \text{ satisfied}  
\iff 

\text{truth value of body} \leq \text{truth value of head}

Lukasiewicz t-norm

\[
I(v_1 \land v_2) = \max\{0, I(v_1) + I(v_2) - 1\}
\]
\[
I(v_1 \lor v_2) = \min\{I(v_1) + I(v_2), 1\}
\]
\[
I(\neg l_1) = 1 - I(v_1)
\]
Distance from Satisfaction

\[ \text{similarAge}(\text{bob}, \text{ann}) \land \text{prefers}(\text{bob}, \text{DEM}) \rightarrow \text{prefers}(\text{ann}, \text{DEM}) \]

\[ 0.8 \quad 0.5 \quad 0.5 \]
Distance from Satisfaction

\[ \text{similarAge}(\text{bob}, \text{ann}) \land \text{prefers}(\text{bob}, \text{民主}) \rightarrow \text{prefers}(\text{ann}, \text{民主}) \]

\[
d_r(I) = \max\{0, I(\text{body}) - I(\text{head})\}
\]

\[
I(v_1 \land v_2) = \max\{0, I(v_1) + I(v_2) - 1\}
\]
Distance from Satisfaction

\[ \text{similarAge}(\text{bob}, \text{ann}) \land \text{prefers}(\text{bob}, \text{donald}) \rightarrow \text{prefers}(\text{ann}, \text{donald}) \]

\[
\begin{align*}
0.8 & \quad 0.5 & \quad 0.5 \\
\end{align*}
\]

\[
d_r(I) = \max\{0, I(\text{body}) - I(\text{head})\}
\]

\[
I(v_1 \land v_2) = \max\{0, I(v_1) + I(v_2) - 1\}
\]

\[
\max\{0, 0.8 + 0.5 - 1\} \leq 0.5 \quad \text{d}=0.0
\]
Distance from Satisfaction

\[ \text{similarAge}(\text{bob}, \text{ann}) \land \text{prefers}(\text{bob}, \text{绝不}) \rightarrow \text{prefers}(\text{ann}, \text{绝不}) \]

\[
\begin{align*}
&0.8 & 0.5 & 0.2 \\
&d_r(I) = \max\{0, I(\text{body}) - I(\text{head})\} \\
&I(v_1 \land v_2) = \max\{0, I(v_1) + I(v_2) - 1\} \\
\end{align*}
\]

\[
\begin{align*}
&\max\{0, 0.8 + 0.5 - 1\} \leq 0.5 \quad d=0.0 \\
&\max\{0, 0.8 + 0.5 - 1\} \leq 0.2 \quad d=0.1
\end{align*}
\]
Distance from Satisfaction

\( \text{similarAge(bob,ann) \land prefers(bob, \text{\textbullet}) \rightarrow prefers(ann, \text{\textbullet})} \)

\[
\begin{align*}
0.8 & \quad 0.9 & \quad 0.2 \\
\end{align*}
\]

\[
d_r(I) = \max\{0, I(\text{body}) - I(\text{head})\}
\]

\[
I(v_1 \land v_2) = \max\{0, I(v_1) + I(v_2) - 1\}
\]

\[
\begin{align*}
\max\{0, 0.8+0.5-1\} & \leq 0.5 & d=0.0 \\
\max\{0, 0.8+0.5-1\} & \leq 0.2 & d=0.1 \\
\max\{0, 0.8+0.9-1\} & \leq 0.2 & d=0.5
\end{align*}
\]
Tasks
Tasks

- MPE Inference
  \[ \text{prefers}(\text{bob, } \text{民主党 }) \geq \text{prefers}(\text{bob, } \text{共和党 }) \]
Tasks

• MPE Inference
  \[ \text{prefers(bob, } 
  \text{驴)} \geq \text{prefers(bob, 驴)} \] ?

• Computing Marginals
  \[ \text{P(prefers(bob, } 
  \text{驴)} \geq 0.8) \] ?
Tasks

• MPE Inference
  \( \text{prefers(bob, } \) \geq \text{prefers(bob, } \)?

• Computing Marginals
  \( P(\text{prefers(bob, } \) \geq 0.8) \)?

• Weight Learning

• Structure Learning
Geometric Intuition:
MPE Inference

\[ d_1(I) = \max\{0, I(x_1) - I(x_2)\} \]
\[ d_2(I) = \max\{0, I(x_2) - I(x_3)\} \]
\[ I(x_1) + I(x_3) \leq 1 \]

most likely interpretations given \( x_1 = 0 \)
MPE Inference

- Convex optimization problem
- New solver [Bach et al, NIPS 12]
  - Consensus optimization
  - Linear time in practice
  - Closed form solutions for subproblems
Consensus Optimization

[Bach et al, NIPS 12]
Consensus Optimization

\[ r_1(V') \]
\[ r_2(V'') \]
\[ r_3(V''') \]

original random variables

[Bach et al, NIPS 12]
Consensus Optimization

[Bach et al, NIPS 12]

rules with local copies of random variables

original random variables

$r_1(V')$

$r_2(V'')$

$r_3(V''')$

$v_1$

$v_2$

$v_3$

$v_4$
Consensus Optimization

[Bach et al, NIPS 12]

rules with local copies of random variables

optimize truth values & agreement with original variables per rule

original random variables
Consensus Optimization

rules with local copies of random variables

optimize truth values & agreement with original variables per rule

update variables to average of copies

[Bach et al, NIPS 12]
Consensus Optimization

[Bach et al, NIPS 12]

rules with local copies of random variables

optimize truth values & agreement with original variables per rule

original random variables

update variables to average of copies
Consensus Optimization

[Bach et al, NIPS 12]

optimize truth values & agreement with original random variables per rule

new: fast solutions

rules with local copies of random variables

original random variables

update variables to average of copies
Geometric Intuition: Marginals

\[ d_1(I) = \max\{0, I(x_1) - I(x_2)\} \]
\[ d_2(I) = \max\{0, I(x_2) - I(x_3)\} \]
\[ I(x_1) + I(x_3) \leq 1 \]
Computing Marginals

Histogram sampling using hit-and-run Monte Carlo scheme

[Broecheler and Getoor, NIPS 10]
Hit-and-Run
Hit-and-Run
Hit-and-Run
Hit-and-Run
Hit-and-Run
Probabilistic Soft Logic (PSL)

Declarative language to specify graphical models

- Logical atoms with soft truth values in \([0,1]\)
- Dependencies as weighted first order rules
- Support for similarity functions and aggregation
- Linear (in)equality constraints

Inference

- MPE: consensus optimization
- Marginals: hit-and-run histogram sampling
Thank you!

http://psl.umiacs.umd.edu