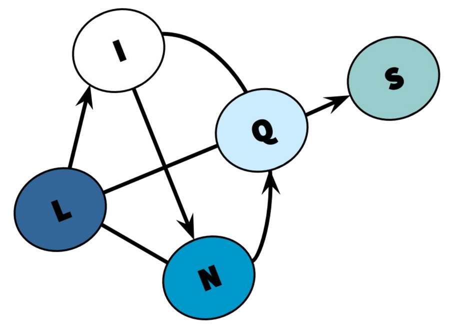
Negative Weights in Hinge-Loss Markov Random Fields

Charles Dickens*, Eriq Augustine*, Connor Pryor, & Lise Getoor
University of California, Santa Cruz
*Equal Contribution



Introduction

Hinge-loss Markov random fields (HL-MRF) are a class of probabilistic graphical models with density functions that admit tractable MAP inference. When paired with a statistical relational learning (SRL) framework, HL-MRFs are powerful tools for performing structured prediction. One such framework, probabilistic soft logic (PSL), uses weighted first-order logical statements to incorporate domain knowledge and constraints into the HL-MRF structure. Traditionally, PSL restricts weights to be non-negative to ensure MAP inference remains tractable, but this limits the types of relations PSL models can represent.

Contributions

- Introduce three novel methods for supporting negative weights in PSL.
- Leverage Gödel logical semantics to preserve convexity and scale of MAP inference problem in PSL models with negative weights.
- Develop a synthetic dataset to compare the effectiveness of the proposed and existing approaches.

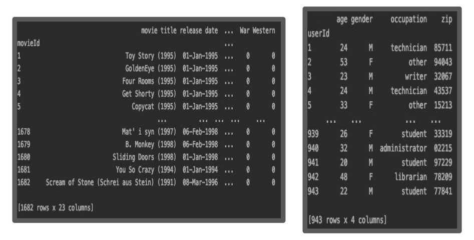
Probabilistic Soft Logic

Probabilistic Soft Logic (PSL) is a general framework for defining a **hinge-loss Markov random field (HL-MRF)**. Dependencies between variables are encoded with weighted first-order logical rules that are instantiated with data to create potentials for defining the HL-MRF density function.

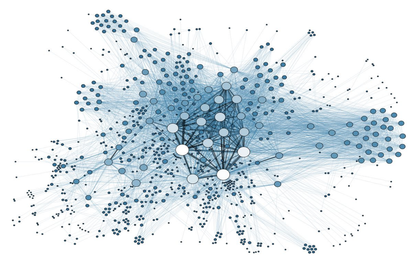
Weighted Logical Rules + Data = HL-MRF

```
w1: Friends(A, B) ∧ Likes(A, I) ->
      Likes(B, I)
w2: Likes(A, I1) ∧ Similar(I1, I2) ->
      Likes(A, I2)
```

+



=



Łukasiewicz Potentials

Instantiated rules are converted to conjunctive normal form and a random variable y_i is associated with each predicate.

```
w2: ¬Likes("Alice", "Coffee") ∨ ¬Similar("Coffee", "Tea") ∨ Likes("Alice", "Tea")
```

PSL conventionally uses Łukasiewicz logical semantics to define hinge-loss potential functions from instantiated weighted logical rules.

$$\phi(\mathbf{y}) = \left(1 - \min\left\{\sum_{i \in I^+} y_i + \sum_{i \in I^-} (1 - y_i), 1\right\}\right)^p$$

HL-MRF Distribution

All instantiated potentials together define a distribution over the random variables

$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{w},\mathbf{x})} \exp\left(-\sum_{i=1}^{m} w_i \phi_i(\mathbf{y},\mathbf{x})\right)$$

MAP Inference

Maximum-a-posteriori (MAP) inference is a tractable convex optimization problem.

$$\arg\min_{\mathbf{y} \in [0,1]^n} \sum_{i=1}^{m} w_i \phi_i(\mathbf{y},\mathbf{x})$$

Negative Weights

Weight Based Approaches

Negative Weights: Drop non negativity constraint. MAP objective is non-convex but can be expressed as a difference of convex functions.

$$\arg\min_{\mathbf{y}|(\mathbf{y},\mathbf{x}) \in \Omega} \sum_{i \in \Phi^+} w_i \phi_i(\mathbf{y},\mathbf{x}) - \sum_{i \in \Phi^-} (-w_i) \phi_i(\mathbf{y},\mathbf{x})$$

Biased Weights: All weights are sufficiently biased to ensure non-negativity.

Negation Based Approaches

Negative weighted rules are first negated in the potential instantiation process.

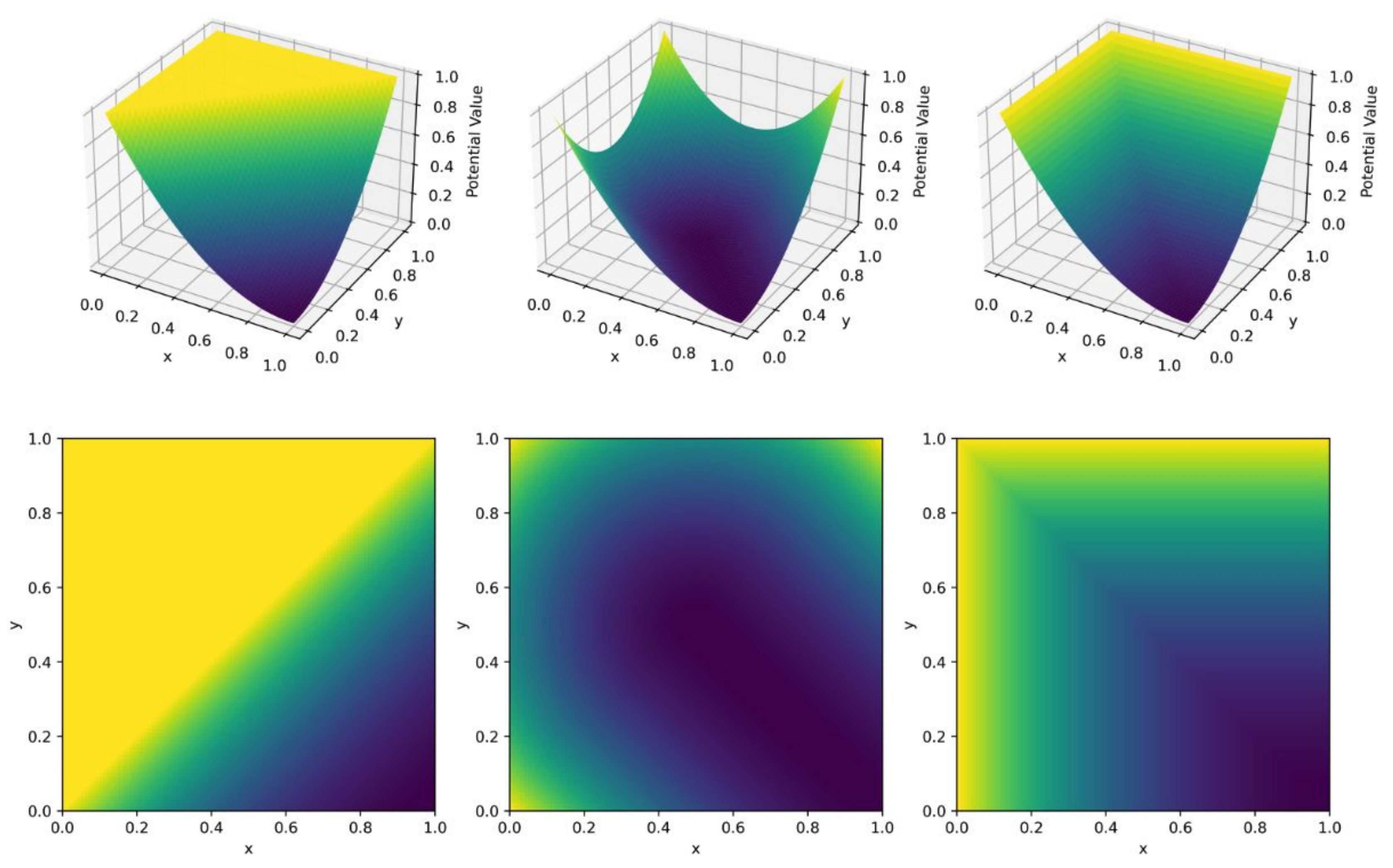
```
W: ¬P("a") ∨ Q("a")   →   |W|: P("a") ∧ ¬Q("a")
```

Suppose P("a") is observed in the example above. The three negation based approaches instantiate the following potentials.

Łukasiewicz Negation: $\phi(y,x) = \min\{1 + y - x, 1\}^2$

Sum of Disjunctions: $\phi(y,x) = (1 - \min\{x + y, 1\})^2 + (1 - \min\{x + (1 - y), 1\})^2$
$+ (1 - \min\{(1 - x) + (1 - y), 1\})^2$

Gödel Negation: $\phi(y,x) = \max\{1 - x, y\}^2$



Surface plots and heat maps of the three potentials instantiated by the negated rule.

Empirical Evaluation

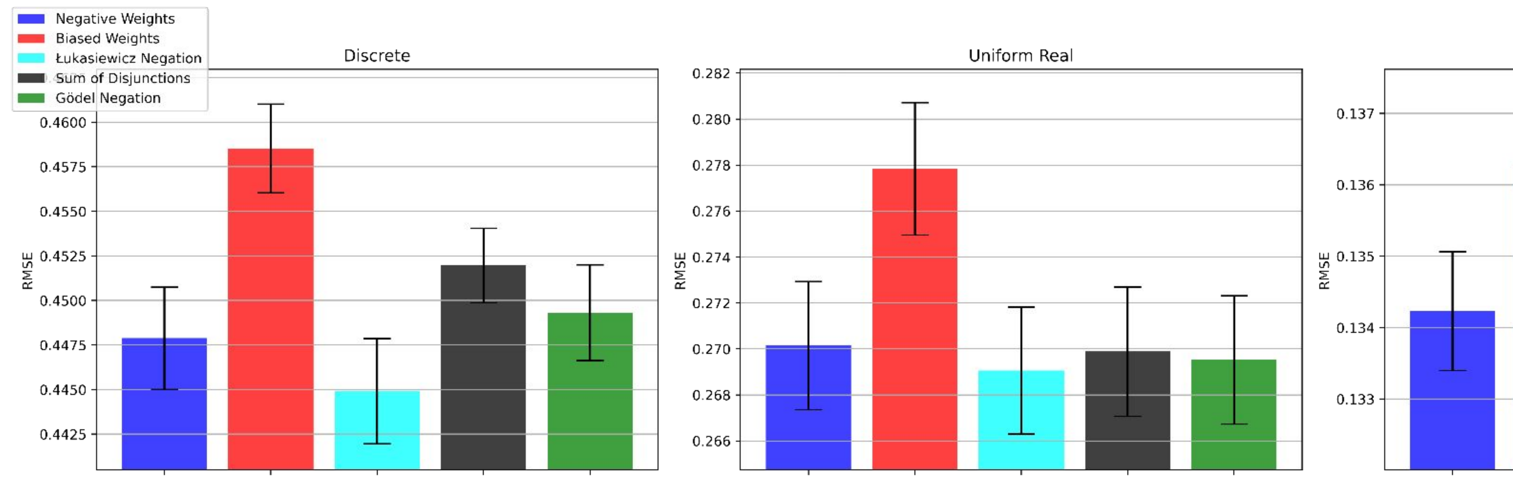
All methods are implemented in PSL and evaluated across 10 folds and 3 variants of a synthetic dataset.

Probabilistic Soft Logic

Evaluation Model

```
-1.0: Predictor1(X, Y) -> Target(X, Y) ^2
0.1: Target(X1, Y) ∧ Similar(X1, X2) -> Target(X2, y) ^2
1.0: Predictor2(X, Y) = Target(X2, y) ^2
0.1: ¬Target(X1, Y) ^2
```

RMSE by Dataset and Method



MAP Inference Convergence

Method	Discrete Epochs	Uniform Real Epochs	Centered Real Epochs
Negative Weights	33.7 (5.48)	22.3 (3.83)	16.4 (3.38)
Biased Weights	37.6 (6.13)	44.1 (5.20)	32.6 (5.06)
Łukasiewicz Negation	44.4 (5.08)	21.4 (4.14)	22.8 (8.88)
Sum of Disjunctions	24.9 (6.28)	26.0 (5.68)	19.5 (3.41)
Gödel Negation	24.6 (3.86)	28.0 (4.99)	20.6 (4.70)

Conclusions and Future Work

Conclusions

- Leveraging alternative real-valued logical semantics can increase expressivity and preserve both convexity of PSL MAP inference and model size.
- SGD on non-convex HL-MRF MAP inference problems from negative weights and Łukasiewicz negation can still converge to good local minimum.

Future Work

- Explore modelling applications for negative weights.
- Find other areas where alternative real-valued logical semantics improves scalability of PSL.
- Integrate negative weight semantics into weight and structure learning algorithms.