# Relational Logistic Regression: the Directed Analog of Markov Logic Networks

Seyed Mehran Kazemi[1], David Buchman[1], Kristian Kersting[2], Sriraam Natarajan[3], David Poole[1]

1- http://cs.ubc.ca/~{smkazemi,davidbuc,poole}, 2-http://www-ai.cs.uni-dortmund.de/PERSONAL/kersting.html, 3-http://homes.soic.indiana.edu/natarasr/

## LR vs. RLR vs. MLN

| | Logistic Regression (LR) | Relational Logistic Regression (RLR) [1] | MLN |
|---|---|---|---|
| **Model:** | Directed | **Directed** | **Undirected** |
| | **Non-relational** | **Relational** | Relational |
| **Defines:** | Conditional prob. dist. (CPD) | CPD (aggregator) | Joint probability |
| **Using:** | Pairwise terms | Weighted logic formulae | Weighted logic formulae |
| **Definition:** | $P(\, var \mid parents\,) \propto \prod_i e^{w_i \; parent_i}$ | $P(\, ground\ var \mid ground\ parents\,) \propto \prod_{all\ true\ formula\ instances} e^{w_{formula}}$ <br> $P(\, ground\ var \mid ground\ parents\,) = sigmoid(\sum_{all\ true\ f.ins.} w_{formula})$ <br> $sigmoid(x) = \frac{1}{1+e^{-x}}$ | $P(\, all\ ground\ vars\,) \propto \prod_{all\ true\ formula\ instances} e^{w_{formula}}$ |

- RLR = relational extension of logistic regression (LR)
- RLR = directed analog of MLNs
  - When all parents observed: MLN's CPD = RLR's CPD
  - When using RLR for all PRVs: MLN's joint distribution ≠ RLRs' joint distribution

## Weighted Formulae (WFs)

| | RLR [1] | MLN |
|---|---|---|
| **Notation:** | $\langle L, F, w \rangle$ <br> $F$ = logic formula <br> $w$ = weight <br> $free\_vars(F) \subseteq L$ = set of logical variables | $\langle F, w \rangle$ <br> $F$ = logic formula <br> $w$ = weight |
| **F instantiated for:** | all bindings of individuals for $L$ | all bindings of individuals for $free\_vars(F)$ |

| Example: RLR WF | Equivalent MLN WF | #Instantiations |
|---|---|---|
| $\langle \{x\}, R(x) \wedge Q, w \rangle$ | $\langle R(x) \wedge Q, w \rangle$ | $\lvert x \rvert = \lvert population(x) \rvert$ |
| $\langle \{x,x'\}, R(x) \wedge Q, w \rangle$ | $\langle R(x) \wedge Q \wedge true(x'), \ w \rangle$ <br> or $\langle R(x) \wedge Q \wedge (R(x') \vee \neg R(x')), \ w \rangle$ | $\lvert x \rvert^2$ |

## Unquantified RLRs: Canonical Forms

- Unquantified RLR formulae may use: **{¬,∧,∨,⊕ (XOR), …}**
- Prop 3: RLRs using only **{∧}** are equally expressive.
  - Corresponds to [2]'s "canonical parametrization with reference state 'true' "
- Prop 4: RLRs using only **{∨}** are equally expressive.
  - Related to [2]'s "canonical parametrization with reference state 'false' "
- Note: RLRs using only **{⊕} (XOR)** are equally expressive.
  - Corresponds to [2]'s "spectral parametrization"

## RLR: Characterization

- Props 6,7 (if and only if):

  CPD representable by RLR ⟺ CPD has the form:

  $$P(\, ground\ var = true \mid ground\ parents\,) = sigmoid(\, polynomial\ of\ counts\,)$$

  count = #tuples (of individuals), for which a formula of the parents is true.

## RLR: Representing Common Aggregators

Unquantified RLRs can represent common aggregators.

- E.g.: AND, OR, noisy-AND, noisy-OR, majority, #trues > constant, %trues > constant
  - Example – $Q \equiv OR(R(*))$: $\langle \{\ \}, Q, \ -M \rangle$
    $\langle \{x\}, Q \wedge R(x), \ 2M \rangle$

    $error_{M \to \infty} \to 0.$
- With non-binary values: mean > constant, max > constant, max, mode=constant, mode

## Examples

"A party is usually fun if you know at least one social person in the party."

As RLR: $\langle \{x\}, \ funFor(x), \ -5 \rangle$
$\langle \{x,y\}, \ funFor(x) \wedge knows(x,y) \wedge social(y), \ 10 \rangle$

$\Longrightarrow P(\, funFor(x) \mid \text{knows}(x,*), \text{social}(*)\,) = sigmoid(-5 + 10 n_T)$

$n_T$ = #individuals $y$ for which $knows(x,y) \wedge social(y)$.

The RLR: $\langle \{x,x'\}, \ Q \wedge R(x) \wedge R(x'), \ 10 \rangle$
$\langle \{x\}, \ Q \wedge \neg R(x), \ -20 \rangle$
$\langle \{\ \}, \ Q, \ -45 \rangle$

Represents: $P(Q = true \mid R(*)) = sigmoid(\, 10 n_T^2 - 20 n_F - 45\,)$

$n_T, n_F$ = #$x$'s such that $R(x) = true, false$

$\Longrightarrow \ P(Q = true \mid R(*)) \cong \{\ 0\ if\ n_T^2 \le 2n_F + 4, \ \ 1\ if\ n_T^2 \ge 2n_F + 5\ \}.$

In canonical {∧} form: $\langle \{x,x'\}, \ Q \wedge R(x) \wedge R(x'), \ 10 \rangle$
$\langle \{x\}, \ Q, \ -20 \rangle$
$\langle \{x\}, \ Q \wedge R(x), \ 20 \rangle$
$\langle \{\ \}, \ Q, \ -45 \rangle$

## REFERENCES

1. Kazemi, S.M., Buchman, D., Kersting, K., Natarajan, S., and Poole, D., "Relational logistic regression", In *Proc. 14th International Conference on Principles of Knowledge Representation and Reasoning (KR)*, 2014.
   **The paper covered here.**
2. Buchman, D., Schmidt, M., Mohamed, S., Poole, D., de Freitas, N., "On Sparse, Spectral and Other Parameterizations of Binary Probabilistic Models", in *15th International Conference on Artificial Intelligence and Statistics (AI-STATS)*, 2012.
   **Canonical models (and learning) for the non-relational, undirected case.**
3. Popescul, A., Ungar, L.H., Lawrence, S., and Pennock, D.M., "Towards structural logistic regression: Combining relational and statistical learning", In *KDD Workshop on Multi-relational Data Mining*, 2002.
   **Learning LR for relational domains using propositionalization.**
4. Huynh, T.N., and Mooney R.J., "Discriminative structure and parameter learning for Markov logic networks", In *Proc. of the International Conference on Machine Learning (ICML)*, 2008.
   **Can be considered as an approach for learning RLR.**