

# Representing Aggregators in Relational Probabilistic Models

David Buchman<sup>1</sup> David Poole<sup>1</sup>

<sup>1</sup> Department of Computer Science, University of British Columbia, Vancouver, B.C., Canada

29th AAAI Conference on Artificial Intelligence (AAAI 2015)



## Aggregators

Aggregators: dependency on unbounded #vars.

Adding aggregators to MLNs – the simplest case:

$$\begin{array}{c} \text{MLN over } \{A(x)\}, \\ \text{defining } \mathbf{P}_{\text{old}}(A_{1..n}) \end{array} \xrightarrow{\text{Add } B \text{ \& weighted formulae } \mathbb{F}} \begin{array}{c} \text{MLN over } \{A(x), B\}, \\ \text{defining } \mathbf{P}_{\text{new}}(A_{1..n}, B). \\ \sum_{B \in \{0,1\}} \mathbf{P}_{\text{new}}(A_{1..n}, B) = \mathbf{P}_{\text{new}}(A_{1..n}) \end{array}$$

“Adding aggregators as **dependent vars**” = the original distribution does not change:

$$\mathbf{P}_{\text{new}}(A_{1..n}) = \mathbf{P}_{\text{old}}(A_{1..n})$$

**Theorem 2:** If  $\mathbb{F}$  has no quantifiers ( $\exists$  and  $\forall$ ),

then  $B$  is independent of  $A_{1..n}$  in all possible groundings.

$\implies$  No useful aggregator can be represented!

◇  $n_0 :=$  #vars among  $A_{1..n}$  that are **0** (“*false*”)

◇  $n_1 :=$  #vars among  $A_{1..n}$  that are **1** (“*true*”)

◇ An aggregator is **saturated** if for any  $n$ ,  $\mathbf{P}(B \mid A_{1..n})$  is equal for all assignments to  $A_{1..n}$  for which  $n_0, n_1 \geq \text{threshold}$ .

**Theorem 4:** General case –  $\mathbb{F}$  may contain quantifiers:

then  $\mathbf{P}(B \mid A_{1..n})$  is a saturated aggregator.

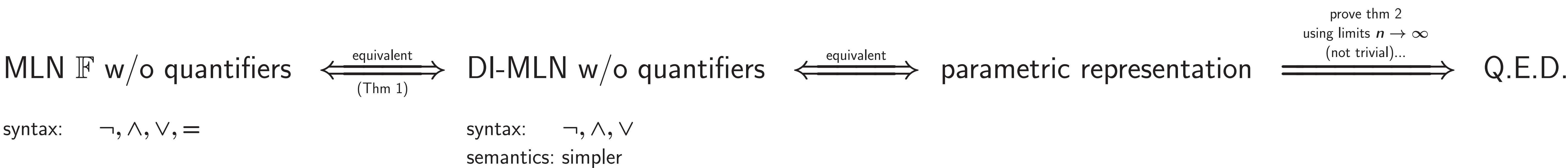
## Overview of Theorem 2 & 4’s Proofs

◇ Ignore the original MLN’s formulae; treat  $\mathbb{F}$  as an MLN over  $\{A(x), B\}$ .

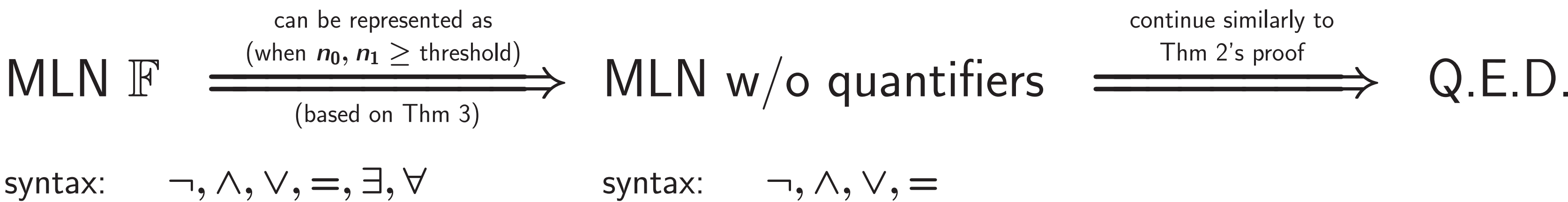
◇ Theorems 2 & 4 are **negative** results  $\implies$  must show **for all possible MLNs**  $\mathbb{F}$ .

◇ Approach: **simplify representation until results can be proved analytically.**

Theorem 2:



Theorem 4:



## Implications

Our Results for Common Binary Aggregators:

	Aggregator	$\mathbf{P}(B = 1 \mid n_0, n_1)$	Parameters	Thm 2	Thm 4
Deterministic:	AND	$\mathbb{I}_{n_0=0}$	$t \geq 0$	no	yes
	OR	$\mathbb{I}_{n_1 \geq 1}$		no	yes
	At least $t$ 1’s	$\mathbb{I}_{n_1 \geq t}$		no	yes
	XOR	$\frac{1 + (-1)^{n_1}}{2}$		no	no
	At least $t\%$ 1’s	$\mathbb{I}_{n_1 / (n_0 + n_1) \geq t}$		no	no
“Hybrid”:	Majority	$\frac{1}{2} \mathbb{I}_{n_1 = n_0} + \mathbb{I}_{n_1 > n_0}$	$0 < t < 1$	no	no
	Noisy AND	$\alpha^{n_0}$		no	approximated
Probabilistic:	Noisy OR	$1 - \alpha^{n_1}$	$\alpha \in [0, 1]$	no	approximated
	Random mux (or “average”)	$\frac{n_1}{n_0 + n_1}$	$w, w_0, w_1$	no	no
	Logistic regression	$1 / (1 + e^{-(w + w_0 n_0 + w_1 n_1)})$		no	no (approx if $w_0 w_1 \geq 0$ )
	Relational logistic regr. (RLR)	$1 / (1 + e^{-\sum_{L \in \mathcal{F}, w} L F w \sum_{x \in \mathcal{X}} F_{n,x} x})$		no	no (approx if ...)

Our mathematical tools may facilitate further theoretic results. E.g.,

**Theorem 5:** Relational logistic regression (RLR) with quantifiers for  $\mathbf{P}(B \mid A_{1..n})$  represents a “sigmoid of a polynomial of counts” when  $n_0, n_1 \geq \text{threshold}$ .

◇ Kazemi, S. M.; Buchman, D.; Kersting, K.; Natarajan, S.; Poole, D. *Relational logistic regression*. In 14th International Conference on Principles of Knowledge Representation and Reasoning (KR 2014).

## DI-MLNs: A Normal Form for MLNs

**DI-MLN (Distinct-Individuals MLN):** A model similar to MLNs, but:

1. Simplified formulae:

syntax: no “=”

semantics: simplified semantics for  $\exists$  &  $\forall$ , denoted by  $\exists_{\neq}$  and  $\forall_{\neq}$

2. Its formulae are only instantiated when all free variables are assigned distinct individuals.

**Theorem 1:**

$$\text{MLNs} \xLeftrightarrow[\text{equivalent (equally expressive)}]{\text{equivalent (equally expressive)}} \text{DI-MLNs}$$

$$\text{MLNs with no } \exists, \forall \xLeftrightarrow[\text{equivalent (equally expressive)}]{\text{equivalent (equally expressive)}} \text{DI-MLNs with no } \exists_{\neq}, \forall_{\neq}$$

## Quantifiers = “Exceptions”

**Theorem 3:**

For any [DI-]formula  $\varphi$  over  $\{A(x), B\}$  there is a [DI-]formula  $\varphi_{\text{unq}}$  **with no quantifiers**, such that

$$\text{whenever } n_0, n_1 \geq \text{threshold}_{\varphi} \implies \varphi_{\text{unq}} \equiv \varphi$$

(I.e.:  $\exists$  and  $\forall$  only allow “exceptions” when  $n_0 < \text{threshold}$  or  $n_1 < \text{threshold}$ .)

**Example 3:**  $\underbrace{\forall_{\neq} x A(x)}_{\varphi} \equiv \underbrace{\text{false}}_{\varphi_{\text{unq}}} \text{ whenever } n_0, n_1 \geq \underbrace{1}_{\text{threshold}_{\varphi}}.$

## Results vs. Literature

Paper/Setting:	Poole et al. (2012)	Thm 2	Thm 4	Natarajan et al. (2010)
Model:	Basic Model:	MLN	MLN	MLN
	Population size:	unbounded	unbounded	unbounded
	Hard constraints allowed?	yes	yes	yes
	# free vars per formula:	<b>0 – 1</b>	<b>0 – <math>\infty</math></b>	<b>0 – <math>\infty</math></b>
	“=” allowed?	<b>no</b> (1 lo. var)	yes	yes
	$\exists, \forall$ allowed?	<b>no</b>	<b>yes</b>	<b>yes</b>
Auxiliary vars:	Weights depend on $n$ ?	no	no	no
		<b>0</b>	<b>0</b>	<b><math>\Theta(n)</math></b>
Summary:	Maximal factor scope:	$\{B, A_i\}$	<b>fixed, arbitrarily large</b>	$\{B, A_{1..n}\}$
	Flexibility:	<b>very limited flex.</b>	<b>flexible</b>	<b>highly flexible</b>
	Result type <sup>1</sup> :	negative	negative	negative
Result:	Which aggregators can be added as dependent PRVs?	<b>none</b> <sup>2</sup>	<b>only saturated aggregators</b>	<b>positive</b>
				at least some

<sup>1</sup> A positive result (e.g., “some”) is stronger when allowing less flexibility. A negative result (e.g., “none” or “only saturated”) is stronger when allowing more flexibility.

<sup>2</sup> Some dependencies on  $n$  itself, but not on  $A_{1..n}$ , can be modeled.

◇ Poole, D.; Buchman, D.; Natarajan, S.; Kersting, K. *Aggregation and population growth: The relational logistic regression and Markov logic cases*. In Proc. UAI Workshop on Statistical Relational AI, 2012.

◇ Natarajan, S.; Khot, T.; Lowd, D.; Tadepalli, P.; Kersting, K. *Exploiting causal independence in Markov logic networks: Combining undirected and directed models*. In European Conference on Machine Learning (ECML), 2010.