THE EFFECT OF COMBINATION FUNCTIONS ON THE COMPLEXITY OF RELATIONAL BAYESIAN NETWORKS

Denis D. Mauá and Fabio G. Cozman

Universidade de São Paulo, Brazil

September 7, 2016
Solution to many tasks comprises building a concrete Bayesian network from some template model for each given domain.

- Object recognition
- Viral marketing
- Entity resolution
- Social network analysis
- ...
**Relational Bayesian Networks**

- **declarative** approach for specifying **abstract** models
- as **expressive** as other probabilistic relational languages
- explicit representation of repetition, determinism and context-specific independence, which can be used to speed up inference
- clear and **rigorously defined** syntax/semantics
 Complexity of inference with RBNs has not been thoroughly examined yet

In particular, the effect of combination functions, which allow summarizing information from different individuals

This work: Complexity of marginal inference as parametrized by combination functions
**Relational Bayesian Network**

Acyclic directed graph where each node is a relation symbol annotated with a probability formula over atoms, numbers and combination functions.

\[
\begin{align*}
\text{difficult} & \rightarrow \text{passes} & \text{intelligent}
\end{align*}
\]

- \( \Pr(\text{difficult}) = 0.3 \)
- \( \Pr(\text{intelligent}) = 0.6 \)
- \( \Pr(\text{passes} | \text{difficult, intelligent}) = \\
(0.4 \cdot \text{difficult}(X) + 0.95(1 - \text{difficult}(X))) \cdot \text{intelligent}(Y) + \\
(0.1 \cdot \text{difficult}(X) + 0.5(1 - \text{difficult}(X)))(1 - \text{intelligent}(Y)) \)
\[ \Pr(\text{difficult}(1)) = \Pr(\text{difficult}(2)) = 0.3 \]
\[ \Pr(\text{intelligent}(1)) = \Pr(\text{intelligent}(2)) = 0.6 \]
\[ \Pr(\text{passes}(1, 2)|\text{difficult}(1), \text{intelligent}(2)) = \\
(0.4\text{difficult}(1) + 0.95(1 - \text{difficult}(1)))\text{intelligent}(2) + \\
(0.1\text{difficult}(1) + 0.5(1 - \text{difficult}(1)))(1 - \text{intelligent}(2)) \]
PROBABILITY FORMULA

- A rational \( q \in [0, 1] \) or an atom \( r(X_1, \ldots, X|_r|) \)
- Convex combination \( F_1 F_2 + (1 - F_1) F_3 \)
- Combination expression:

\[
\text{comb}\{F_1, \ldots, F_k|Y_1, \ldots, Y_m; \alpha\},
\]

where \( \alpha \) is an equality constraint such as

\[
X = Y \text{ or } (\neg X = Y) \lor (Y = Z \land X = Z)
\]

- Summarize information about individuals

Example:

\[
F(X) = \text{mean}\{0.6 \cdot r(X) + \\
0.7 \cdot \max\{1 - s(X, Y)|X; X = X\}|Y, Z; Y \neq X \land Z \neq X\}
\]
**SEMANTICS**

**INTERPRETATION**

- Set of constants $\mathcal{D}$ (domain)
- Map $\mu$:
  - relation symbol $r$ into relation $r^\mu \subseteq \mathcal{D}^{|r|}$
  - equality constraint into standard semantics
  - probability formula $F$ into function $F^\mu : \mathcal{D}^n \rightarrow [0, 1]$, where $n$ is the number of free variables in $F$
Probability Formula

\[
q \xrightarrow{\mu} q
\]

\[
r(X_1, \ldots, X_{|r|}) \xrightarrow{\mu} F^\mu(a) = \begin{cases} 
1 & \text{if } a \in r^\mu \\
0 & \text{if } a \notin r^\mu
\end{cases}
\]

\[
F_1 F_2 + (1 - F_1) F_3 \xrightarrow{\mu} F_1^\mu(a) F_2^\mu(a) + (1 - F_1^\mu(a)) F_3^\mu(a)
\]
**Probability Formula**

If \( F = \text{comb}\{F_1, \ldots, F_k | Y_1, \ldots, Y_m; \alpha\} \), then \( F^\mu(a) = \text{comb}(Q) \)

where

- \( \text{comb} \) is the combination function
- \( Q \) is the multiset containing a number \( F_i(a, b) \) for every \( (a, b) \in \alpha^\mu \)

**Example:**

- \( \mathcal{D} = \{1, 2\}; \ \max\{\frac{1}{3}, \frac{2}{3}, 1 | Y; Y = Y\} \xrightarrow{\mu} \max\{\frac{1}{3}, \frac{2}{3}, 1, \frac{1}{3}, \frac{2}{3}, 1\} \)
Given a domain $\mathcal{D}$, an RBN with graph $G = (V, A)$ induces a probability distribution over interpretations $\mu$ by

$$
\Pr(\mu|\mathcal{D}) = \prod_{r \in V} \prod_{a \in r^{\mu}} F_r(a) \prod_{a \notin r^{\mu}} (1 - F_r(a)) \cdot \Pr(r^{\mu}|\{s^{\mu}: s \in \text{pa}(r)\}, \mathcal{D})
$$

where occurrences of $s \in \text{pa}(r)$ in probability formula $F_r(a)$ are interpreted according to $s^{\mu}$
$F_{\text{passes}(1, 2)} =$

$(0.4 \text{difficult}(1) + 0.95(1 - \text{difficult}(1))) \text{intelligent}(2) +$

$(0.1 \text{difficult}(1) + 0.5(1 - \text{difficult}(1)))(1 - \text{intelligent}(2))$
COMPLEXITY OF MARGINAL INFERENCE

Given:

- Relational Bayesian Network (graph and probability formulas)
- domain $\mathcal{D}$ specified as a list of elements
- ground atom $r(a)$

Compute:

$$\Pr(r(a) = 1) = \sum_{\mu : a \in r^\mu} \Pr(\mu | \mathcal{D})$$

Note: $\Pr(r_1(a_1) = \alpha_1, \ldots, r_n(a_n) = \alpha_n)$ can be computed as $\Pr(r(a) = 1)$ by defining $F_r = [1-]r_1(X_1) \cdots [1-]r_n(n)$; conditional probability can be computed with two such calls.
COUNTING TURING MACHINE

Non-deterministic Turing machine that writes on a separate tape and in binary notation the number of accepting paths

```
1 0 1 1 1 0 ...
```

Input/Working Tape

```
0 1 0
```

Counting Tape

#P

integer-valued functions computed by counting Turing machine with polynomial steps

#EXP

integer-valued functions computed by counting Turing machine with exponential steps
ORACLE TURING MACHINE

Turing machine which can query an “oracle”, which then reads/writes content from/to the oracle tape in one step.

\[
\begin{array}{cccccc}
1 & 0 & 1 & 1 & 1 & 0 \\
\end{array}
\quad \cdots \quad \text{Input/Working Tape}
\]

\[
\begin{array}{cccc}
0 & 1 & 0 \\
\end{array}
\quad \cdots \quad \text{Oracle Tape}
\]

ORACLE COMPLEXITY CLASSES:

For complexity classes $A$ and $B$, say that a problem is in $A^B$ if it can be computed by an $A$-complete problem with access to an oracle that is complete for $B$. Examples: $P^{NP}$, $NP^{PSPACE}$. 
COUNTING HIERARCHY

\[ \#^\sum_p^k = \#^{\text{P}^\text{NP}^\text{NP}^\cdots}^k \text{ times} \]

integer-valued functions computed by counting Turing machine with oracle \( \sum_p^k \) (\( k \) “stacks” of \( \text{NP} \) machines) with polynomial steps

\[ \#\text{PH} = \#^{\text{P}^\text{PH}} \]

integer-valued functions computed by counting Turing machine with oracle \( \text{PH} \) (arbitrary stacking of \( \text{NP} \) machines) with polynomial steps

\[ \text{FP} \subseteq \text{NP} \subseteq \#\text{P} \subseteq \#^\sum_1^p \cdots \subseteq \#\text{PH} \subseteq \text{FPSPACE} \subseteq \#\text{EXP} \subseteq \text{FEXP} = \#\text{PSPACE} \]
**COMPLETENESS**

A problem is said complete for a class if it is the hardest problem in that class.

We cannot establish e.g. \( \#P \)-completeness of inference as it returns rationals (and \( \#P \) does not seem to be closed under division).

**EQUIVALENCE**

A problem \( A \) is \( X \)-equivalent for counting class \( X \) if it is \( X \)-hard (via parsimonious reduction) up to a polynomial scaling and can be solved with \( FP^X \).

Marginal inference in Bayesian networks specified by CPTs is \( \#P \)-equivalent.
**THEOREM**

*Inference in RBNs without combination functions is \( \#P \)-equivalent, even if the domain is specified solely by its size in binary notation*

**Proof:** RBNs without combination functions encode plate models, which are \( \#P \)-equivalent (Cozman & Mauá, 2015)
**Theorem**

Inference is \(\#\text{EXP}\)-equivalent when the only combination function is \(\text{max}\).

**Proof:** RBNs with max combination functions encode enhanced plate models, which are \(\#\text{EXP}\)-equivalent by reduction from domino tiling (Cozman & Mauá, 2015)
Exponential behavior due to unbounded arity (which allows us to specify an exponential number of relevant individuals)

**Theorem**

*Inference is FPSPACE-complete when the arity is bounded and the only combination function is max.*

**Proof:** Hardness is by reduction from counting QBF solutions:

- **Input:** A formula $\varphi(X_1, \ldots, X_n) = Q_1 Y_1 Q_2 Y_2 \ldots Q_m Y_m \psi(X_1, \ldots, X_n, Y_1, \ldots, Y_m)$, where each $Q_i$ is either $\exists$ or $\forall$, and $\psi$ is a 3-CNF formula over variables $X_1, \ldots, X_n, Y_1, \ldots, Y_m$.

- **Output:** The number of assignments to the variables $X_1, \ldots, X_n$ that satisfy $\varphi$. 
NESTING LEVEL

- Nesting level of $F = q$ or $F = r(X_1, \ldots, X_n)$ is zero
- Nesting level of $F = F_1F_2 + (1 - F_2)F_3$ is the highest nesting level of $F_1, F_2, F_3$
- Nesting level of $F = \text{comb}\{F_1, \ldots, F_k|Y_1, \ldots, Y_M; \alpha\}$ is the highest nesting level over all $F_i$, plus one.

**Theorem**

*Inference is $\#\Sigma_k^P$-equivalent when arity is bounded, nesting level is at most $k$, and the only combination function is max.*

**Proof:** Reduction from $\#\Pi_k\text{SAT}$ (counting satisfying assignments of CNF formulas with at most $k$ alternating quantifiers)
RESULTS

Other combination functions

**THEOREM**

*Inference is FEXP-complete when arity is bounded and combination functions are polynomial in their arguments*

**Proof:** Succinct specification of an exponential multiset allows exponential computation (as in succinct circuits).
Other combination functions

**THEOREM**

Inference is \(\#P\)-equivalent when arity is bounded and combination formulas are polynomial

**Proof:** Ground the model into a Bayesian network with polynomial effort.
CONCLUSION

- No combination functions, $\rightarrow$ inference is $\#P$-equivalent
- Only maximization as combination function
  - $\#\text{EXP}$-equivalent
  - $\text{FPSPACE}$-complete with bound on arity of relations
  - $\#\Sigma_k$-complete with bound on arity of relations and bound on number of nestings
- $\text{FEXP}$-complete if combination function is polynomial
- $\#P$-equivalent if probability formula is polynomial