Lifted First-Order Probabilistic Inference

Rodrigo de Salvo Braz
SRI International
based on work with
Eyal Amir and Dan Roth
Outline

- **Motivation**
  - Brief background
  - Lifted Inference example
- **First-order representations**
- **Lifted First-Order Probabilistic Inference**
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- **Conclusion**
**Bayesian Networks (directed)**

\[
P(\text{sick}_\text{john}, \text{sick}_\text{mary}, \text{sick}_\text{bob}, \text{epidemic})
\]

\[
= P(\text{sick}_\text{john} | \text{epidemic}) \times P(\text{sick}_\text{mary} | \text{epidemic})
\times P(\text{sick}_\text{bob} | \text{epidemic}) \times P(\text{epidemic})
\]
Factor Networks (undirected)

\[ P(\text{epi}_\text{france}, \text{epi}_\text{belgium}, \text{epi}_\text{uk}, \text{epi}_\text{germany}) \]

\[ / \quad \phi_1(\text{epi}_\text{france}, \text{epi}_\text{belgium}) \ast \phi_2(\text{epi}_\text{france}, \text{epi}_\text{uk}) \]

\[ \ast \phi_3(\text{epi}_\text{france}, \text{epi}_\text{germany}) \ast \phi_4(\text{epi}_\text{belgium}, \text{epi}_\text{germany}) \]
Bayesian Nets as Factor Networks

\[
P(sick\_john, sick\_mary, sick\_bob, epidemic) \\
/ \ P(sick\_john | epidemic) * P(sick\_mary | epidemic) \\
* P(sick\_bob | epidemic) * P(epidemic)
\]
Inference: Marginalization

\[
P(sick\_john) / \sum_{epidemic} \sum_{sick\_mary} \sum_{sick\_bob} P(sick\_john | epidemic) * P(sick\_mary | epidemic) * P(sick\_bob | epidemic) * P(epidemic)
\]
Inference: Variable Elimination (VE)

\[
\frac{P(sick\_john)}{\sum_{\text{epidemic}} P(sick\_john | \text{epidemic}) * P(\text{epidemic}) * \sum_{\text{sick\_mary}} P(sick\_mary | \text{epidemic}) * \sum_{\text{sick\_bob}} P(sick\_bob | \text{epidemic})}
\]
Inference: Variable Elimination (VE)

\[ P(\text{epidemic}) \]
\[ P(\text{sick}_\text{john} | \text{epidemic}) \]
\[ P(\text{sick}_\text{mary} | \text{epidemic}) \]
\[ P(\text{sick}_\text{john}) / \sum_{\text{epidemic}} P(\text{sick}_\text{john} | \text{epidemic}) \times P(\text{epidemic}) \times \sum_{\text{sick}_\text{mary}} P(\text{sick}_\text{mary} | \text{epidemic}) \times \phi_1(\text{epidemic}) \]
Inference: Variable Elimination (VE)

\[ P(sick\_john) / \sum_{epidemic} P(sick\_john \mid epidemic) \times P(epidemic) \times \phi_1(epidemic) \times \sum_{sick\_mary} P(sick\_mary \mid epidemic) \]
Inference: Variable Elimination (VE)

\[
P(\text{sick}_{\text{john}}) / \sum_{\text{epidemic}} P(\text{sick}_{\text{john}} | \text{epidemic}) \times P(\text{epidemic}) \\
\times \phi_1(\text{epidemic}) \\
\times \phi_2(\text{epidemic})
\]
Inference: Variable Elimination (VE)

\[ P(\text{sick\_john}) / \phi_3(\text{sick\_john}) \]
Outline

- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion
A Lifted inference example

- Dependencies between
  - epidemic(Disease) and sick(Person, Disease)
  - sick(Person, Disease) and hospital(Person)

- Applied to many people and many diseases.
Graphical model size can be large

epidemic(measles)  ...  epidemic(flu)

sick(mary,measles)  ...  sick(mary,flu)

sick(bob,measles)  ...  sick(bob,flu)

hospital(mary)  ...  hospital(bob)
Much redundancy

epidemic(measles) ... epidemic(flu)

sick(mary, measles) ... sick(mary, flu) ... sick(bob, measles) ... sick(bob, flu)

hospital(mary) ... ... hospital(bob)
Representing structure

Parfactors, for parameterized factors

sick(mary, measles)
epidemic(measles)
hospital(mary)
sick(mary, flu)
epidemic(flu)
...

sick(bob, measles)
epidemic(measles)
hospital(bob)
sick(bob, flu)
epidemic(flu)
...

sick(P, D)
epidemic(D)
hospital(P)
sick(P, D)
epidemic(D)
hospital(P)
Lifted Inference Example

- $P(\text{hospital(john)}) = ?$

Diagram:
- epidemic($D$)
- sick($P,D$)
- hospital($P$)
Lifted Inference Example

- \( P(\text{hospital}(\text{john})) = ? \)
Lifted Inference Example

P(hospital(john)) = ?

epidemic(flu)
epidemic(measles)
sick(mary,flu)
hospital(mary)
sick(bob,measles)
hospital(bob)
epidemic(D)
sick(P,D)
sick(john,D)
hospital(P)
hospital(john)
Lifted Inference Example

$P(\text{hospital(john)}) = ?$

$$\phi'(\text{e(D),h(P)}) = \sum_{s(P,D): P \neq \text{john}} \phi_1(s(P,D),\text{e(D)}) \phi_2(s(P,D),\text{h(P)})$$
Lifted Inference Example

\[ P(\text{hospital(john)}) = ? \]
Lifted Inference Example

- $P(\text{hospital(john)}) = ?$
Lifted Inference Example

\[ P(\text{hospital}(\text{john})) = ? \]

- \( \text{epidemic(D)} \)
- \( P \neq \text{john} \)
- \( \text{hospital}(P) \)
- \( \text{hospital}(\text{john}) \)
- \( \text{epidemic(flu)} \)
- \( \text{epidemic(measles)} \)
- \( \text{hospital(mary)} \)
- \( \text{hospital(bob)} \)
Lifted Inference Example

- $P(\text{hospital(john)}) = ?$
Lifted Inference Example

- \( P(\text{hospital(john)}) = ? \)
Outline

- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion
Representing structure

- More compact
- Represents structure explicitly
- Generalization of graphical models

Logical Variables parameterize random variables

Parfactors, for parameterized factors

Atoms represent a set of random variables
Semantics

\[ \phi(\text{sick}(\text{mary,measles}), \text{epidemic}(\text{measles})) \]
Making use of structure in Inference

- Task: calculate marginals and posteriors:
  \[ P(\text{sick(bob, measles)} \mid \text{sick(mary, measles)}) = ? \]

- Three approaches
  - plain propositionalization
  - dynamic construction ("smart" propositionalization)
  - lifted inference
Inference - Plain Propositionalization

- Instantiation of potential function for each instantiation
- Lots of redundant computation
- Lots of unnecessary random variables
Inference - Dynamic construction

- Instantiation of potential function for relevant parts
- $P(\text{hospital(mary)} \mid \text{no epidemics}) = ?$
Inference - Dynamic construction

- Instantiation of potential function for relevant parts
- $P(\text{hospital(mary)} \mid \text{no epidemics}) = ?$

Much redundancy still
Inference - Dynamic construction

Used to be the most common approach for exact First-order Probabilistic inference:

- Knowledge-based model construction (Breese, 1992)
- Probabilistic Logic Programming (Ng & Subrahmanian, 1992)
- Probabilistic Logic Programming (Ngo and Haddawy, 1995)
- Probabilistic Relational Models (Friedman et al., 1999)
- Relational Dependency Networks (Neville & Jensen, 2004)
- Relational Bayesian networks (Jaeger, 1997)
- Bayesian Logic Programs (Kersting & DeRaedt, 2001)
- MEBN (Laskey, 2004)
- Markov Logic Networks (Richardson & Domingos, 2004)
Inference - Lifted inference

- Inference on parameterized, or first-order, level;
- Performs certain inference steps once for a class of random variables;
- Poole (2003) describes a generalized Variable Elimination algorithm which we call Inversion Elimination.
- We formalized Inversion Elimination, showed its limitations and introduced Counting Elimination (IJCAI’05) and Partial Inversion (AAAI’06)
Outline

- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion
Inversion Elimination (IE)

\[
\text{sick}(P, D) \quad \text{epidemic}(D) \quad D \neq \text{measles}
\]

Eliminate
Inversion Elimination (IE)

\[ \text{epidemic}(D) \]

\[ D \neq \text{measles} \]

\[ \text{sick}(P, D) \]

- \[ \text{epidemic}(\text{flu}) \]
  - \[ \text{sick}(\text{john}, \text{flu}) \]
- \[ \text{epidemic}(\text{flu}) \]
  - \[ \text{sick}(\text{mary}, \text{flu}) \]
- \[ \ldots \]
  - \[ \text{sick}(\text{mary}, \text{rubella}) \]

- \[ \text{epidemic}(\text{rubella}) \]
Inversion Elimination (IE)

\[ \sum \text{sick}(P,D) \phi(\text{epidemic}(D),\text{sick}(P,D)) \]

Time complexity is independent from domain size
Inversion Elimination (IE)

Grounding

\[ \sum_{\text{sick}(P,D)} \phi(\text{epidemic}(D), \text{sick}(P,D)) \]

VE

\[ \text{epidemic}(flu) \]
\[ \text{sick}(john, flu) \]

\[ \text{epidemic}(rubella) \]
\[ \text{sick}(mary, rubella) \]

IE

\[ \text{epidemic}(D) \]
\[ D \neq \text{measles} \]
[...]

Abstraction

\[ \text{epidemic}(D) \]
\[ D \neq \text{measles} \]
[...]

\[ \text{sick}(P,D) \]
\[ \text{epidemic}(D) \]

\[ \text{sick}(P,D) \]

\[ D \neq \text{measles} \]
[...]
Formalization of IE

- Joint distribution
  \[ \prod_D \phi_1(e(D)) \prod_{P,D} \phi_2(e(D),s(P,D)) \]

- Marginalization by eliminating class \( s(P,D) \):
  \[
  \sum_{s(\cdot,\cdot)} \prod_D \phi_1(e(D)) \prod_{P,D} \phi_2(e(D),s(P,D)) = \prod_D \phi_1(e(D)) \sum_{s(\cdot,\cdot)} \prod_{P,D} \phi_2(e(D),s(P,D))
  \]
Formalization of IE

\[ \sum_{s(\cdot, \cdot)} \prod_{P,D} \phi_2(e(D), s(P,D)) \]

\[ = \prod_{P,D} \sum_{s(P,D)} \phi_2(e(D), s(P,D)) \]

\[ = \prod_{P,D} \phi_3(e(D)) \]

\[ = \prod_{D} \phi_3^{\mid P \mid}(e(D)) \]

\[ = \prod_{D} \phi_4(e(D)) \]
Formalization of IE

\[ \sum_{s(\cdot, \cdot)} \prod_{P,D} \phi_2(e(D), s(P,D)) \]
\[ = \sum_{s(p_1,d_1)} \sum_{s(p_1,d_2)} \cdots \sum_{s(p_n,d_m)} \phi_2(e(d_1), s(p_1,d_1)) \phi_2(e(d_1), s(p_1,d_2)) \cdots \phi_2(e(d_n), s(p_n,d_m)) \]
\[ = (\sum_{s(p_1,d_1)} \phi_2(e(d_1), s(p_1,d_1))) \cdots (\sum_{s(p_1,d_2)} \phi_2(e(d_1), s(p_1,d_2))) \cdots (\sum_{s(p_n,d_m)} \phi_2(e(d_n), s(p_n,d_m))) \]
\[ = \prod_{P,D} \sum_{s(P,D)} \phi_2(e(D), s(P,D)) \]
Formalization of IE

- Inversion Elimination requires certain conditions for its application:
  - Eliminated atom must contain all logical variables in parfactor involved;
  - Eliminated atom instances must not occur together in the same instance of parfactor.
- These are both conditions for parallelism.
Inversion Elimination - Limitations - I

- Eliminated atom must contain all logical variables in parfactors involved.

\[
\text{epidemic}(D) \rightarrow \text{sick}(P,D)
\]
Inversion Elimination - Limitations - I

- Eliminated atom must contain all logical variables in parfactors involved.

```
eliminated atom:
sick(P,D)
epidemic(D)
```

Ok, contains both P and D
Inversion Elimination - Limitations - I

- Eliminated atom must contain all logical variables in parfactors involved.

```
sick(P,D)
epidemic(D)
```

Not Ok, missing P
Eliminated atom must contain all logical variables in parfactors involved.

No atom can be eliminated
Inversion Elimination - Limitations - I

Marginalization by eliminating class e(D):

\[ \sum_{e(.)} \prod_{D,P} \phi_2(e(D),s(P,D)) \]

\[ = \sum_{e(d_1)} \ldots \sum_{e(d_n)} \prod_{P} \phi_2(e(d_1),s(P,d_1)) \ldots \]

\[ \prod_{P} \phi_2(e(d_n),s(P,d_n)) \]

\[ = (\sum_{e(d_1)} \prod_{P} \phi_2(e(d_1),s(P,d_1))) \ldots \]

\[ (\sum_{e(d_n)} \prod_{P} \phi_2(e(d_n),s(P,d_n))) \]

\[ = \prod_{D} \sum_{e(D)} \prod_{P} \phi_2(e(D),s(P,D)) \]
Inversion Elimination - Limitations - II

- Requires eliminated RVs to occur in separate instances of parfactor.

Inversion Elimination
Ok
Inversion Elimination - Limitations - II

- Requires eliminated RVs to occur in separate instances of parfactor

Inversion Elimination
Not Ok

D1 ≠ D2

epidemic(D1)

epidemic(D2)

epidemic(measles)

epidemic(flu)

epidemic(rubella)

...
Inversion Elimination - Limitations - II

\[ \sum_{e(\cdot)} \prod_{D_1 \neq D_2} \phi(e(D_1), e(D_2)) \]

\[= \sum_{e(d_1)} \cdots \sum_{e(d_n)} \phi(e(d_1), e(d_2)) \cdots \phi(e(d_{n-1}), e(d_n)) \]

\[= \sum_{e(d_1)} \phi(e(d_1), e(d_2)) \cdots \sum_{e(d_n)} \phi(e(d_{n-1}), e(d_n)) \]
Outline

- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion
Counting Elimination

\[ \sum_{e(\cdot)} \prod_{D_1 \neq D_2} \phi(\cdot(D_1), e(D_2)) = \sum_{e(\cdot)} \prod_{v} \phi(v) \#v \text{ in } e(\cdot), D_1 \neq D_2 \]

Does depend on domain size, but not exponentially

\[
\sum_{e(\cdot)} \prod_{v} \phi(v) \#v \text{ in } e(\cdot), D_1 \neq D_2 (\text{from } i)
\]

\[\#(0,1) = 70 \times 30\]

\[\#(1,1) = 30 \times (30 - 1)\]
Counting Elimination Conditions

- No shared logical variables between atoms, so counting can be done independently.

\( \phi(\text{epidemic}(D_1, \text{Region}), \text{epidemic}(D_2, \text{Region})) \) is not suitable for Counting elimination.
Outline

- Motivation
  - Brief background
  - Lifted Inference example

- First-order representations

- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion

- Conclusion
Some things not covered by both Inversion and Counting

- Eliminating epidemic from
  \[ \phi( \text{epidemic}(\text{Disease1}, \text{Region}), \text{epidemic}(\text{Disease2}, \text{Region}), \text{other}_\text{RVs}) \]
- No atom with all logical variables
  ) no Inversion Elimination.
- Shared logical variables
  ) no Counting Elimination.
Partial Inversion, graphically

Each instance a counting elimination problem
Partial Inversion

\[ \sum_{e(.,.)} \prod_{D_1 \neq D_2, R} \phi(e(D_1, R), e(D_2, R), o) = \prod_{R} \sum_{e(.,R)} \prod_{D_1 \neq D_2} \phi(e(D_1, R), e(D_2, R), o) = \prod_{R} \phi'(o) = \phi'(o)^{|R|} - \phi''(o) \]

Generalizes both Inversion and Counting Elimination

R is bound

) no shared logical variables

) Counting Elimination
Partial Inversion Conditions

\[ \phi( \text{friends}(X,Y), \text{friends}(Y,X), \text{sick}(X,D), \text{sick}(Y,D)) \]

- Cannot partially invert on \( X,Y \) because \text{friends}(bob,mary) appears in more than one grounding of parfactor.
- But with Partial Inversion we do not need all logical variables in order to invert.
Outline

- Motivation
  - Brief background
  - Lifted Inference example
- First-order representations
- Lifted First-Order Probabilistic Inference
  - Inversion Elimination
  - Counting Elimination
  - Partial Inversion
- Conclusion
Conclusion

- Expressive probabilistic representations used during inference as well.
- Lifted inference equivalent to grounded inference; much faster, yet yielding the same exact answer.
Thanks!