Relax, compensate and then integrate

Fast approximate probabilistic inference with logical and algebraic constraints

Zhe Zeng*
University of California, Los Angeles

Paolo Morettin*
University of Trento, Italy

Fanqi Yan*
AMSS, Chinese Academy of Sciences

Antonio Vergari
University of California, Los Angeles

Guy Van den Broeck
University of California, Los Angeles

September 14th, 2020 - DeCoDeML Workshop 2020 - ECML-PKDD
Relax, compensate and then integrate

Fast approximate probabilistic inference with logical and algebraic constraints

Zhe Zeng*
University of California, Los Angeles

Paolo Morettin*
University of Trento, Italy

Fanqi Yan*
AMSS, Chinese Academy of Sciences

Antonio Vergari
University of California, Los Angeles

Guy Van den Broeck
University of California, Los Angeles

September 14th, 2020 - DeCoDeML Workshop 2020 - ECML-PKDD
Relax, compensate and then integrate

Fast approximate probabilistic inference with logical and algebraic constraints

Zhe Zeng*
University of California, Los Angeles

Paolo Morettin*
University of Trento, Italy

Fanqi Yan*
AMSS, Chinese Academy of Sciences

Antonio Vergari
University of California, Los Angeles

Guy Van den Broeck
University of California, Los Angeles

September 14th, 2020 - DeCoDeML Workshop 2020 - ECML-PKDD
Relax, compensate and then integrate

Fast approximate probabilistic inference with logical and algebraic constraints

Zhe Zeng*
University of California, Los Angeles

Paolo Morettin*
University of Trento, Italy

Fanqi Yan*
AMSS, Chinese Academy of Sciences

Antonio Vergari
University of California, Los Angeles

Guy Van den Broeck
University of California, Los Angeles

September 14th, 2020 - DeCoDeML Workshop 2020 - ECML-PKDD
Skill matching system

Minka et al., “Trueskill 2: An improved bayesian skill rating system”, 2018
Skill matching system

Each player has a certain skill

Minka et al., “Trueskill 2: An improved bayesian skill rating system”, 2018
Skill matching system

Each player has a certain skill

continuous variables

Minka et al., “Trueskill 2: An improved bayesian skill rating system”, 2018
Skill matching system

Each player has a certain skill
Players can form teams

Minka et al., “Trueskill 2: An improved bayesian skill rating system”, 2018
Skill matching system

- Each player has a certain skill
- Players can form teams

Minka et al., “Trueskill 2: An improved bayesian skill rating system”, 2018
Skill matching system

- Each player has a certain skill
- Players can form teams
- Each team’s skill is bounded by its players’ skills

Minka et al., “Trueskill 2: An improved bayesian skill rating system”, 2018
Skill matching system

- Each player has a certain skill
- Players can form teams
- Each team’s skill is bounded by its players’ skills ⇒ complex constraints!

Minka et al., “Trueskill 2: An improved bayesian skill rating system”, 2018
Skill matching system

- Each player has a certain skill
- Players can form teams
- Each team’s skill is bounded by its players’ skills
- Good teams form a squad

Minka et al., “Trueskill 2: An improved bayesian skill rating system”, 2018
Skill matching system

Each player has a certain skill
Players can form teams
Each team’s skill is bounded by its players’ skills
Good teams form a squad

Minka et al., “Trueskill 2: An improved bayesian skill rating system”, 2018
“What is the probability of team $T_1$ to outperform team $T_2$, if $T_1$ is a squad but $T_2$ is not?”

Minka et al., “Trueskill 2: An improved bayesian skill rating system”, 2018
Continuous $+$ discrete $+$ constraints $=$ ?
Continuous + discrete + constraints = ?

Generative adversarial networks (GANs) [Goodfellow et al. 2014]
Variational Autoencoders (VAEs) [Kingma et al. 2013]
Continuous + discrete + constraints = ?

Generative adversarial networks (GANs) [Goodfellow et al. 2014]
Variational Autoencoders (VAEs) [Kingma et al. 2013]
⇒ limited inference capabilities, no constraints
Continuous + discrete + constraints = ?

Generative adversarial networks (GANs) [Goodfellow et al. 2014]
Variational Autoencoders (VAEs) [Kingma et al. 2013]

Hybrid Bayesian Networks (HBNs) [Heckerman et al. 1995; Shenoy et al. 2011]
Mixed Probabilistic Graphical Models (MPGMs) [Yang et al. 2014]
Continuous + discrete + constraints = ?

Generative adversarial networks (GANs) [Goodfellow et al. 2014]
Variational Autoencoders (VAEs) [Kingma et al. 2013]
Hybrid Bayesian Networks (HBNs) [Heckerman et al. 1995; Shenoy et al. 2011]
Mixed Probabilistic Graphical Models (MPGMs) [Yang et al. 2014]

⇒ strong distributional assumptions
Continuous + discrete + constraints = ?

Generative adversarial networks (GANs) [Goodfellow et al. 2014]
Variational Autoencoders (VAEs) [Kingma et al. 2013]
Hybrid Bayesian Networks (HBNs) [Heckerman et al. 1995; Shenoy et al. 2011]
Mixed Probabilistic Graphical Models (MPGMs) [Yang et al. 2014]
Tractable Probabilistic Circuits (PCs) [Molina et al. 2018; Vergari et al. 2019]
Continuous + discrete + constraints = ?

Generative adversarial networks (GANs) [Goodfellow et al. 2014]
Variational Autoencoders (VAEs) [Kingma et al. 2013]

Hybrid Bayesian Networks (HBNs) [Heckerman et al. 1995; Shenoy et al. 2011]
Mixed Probabilistic Graphical Models (MPGMs) [Yang et al. 2014]

Tractable Probabilistic Circuits (PCs) [Molina et al. 2018; Vergari et al. 2019]

⇒ cannot deal with complex algebraic constraints
Continuous + discrete + constraints = SMT

Satisfiability Modulo Theories
of the linear arithmetic over the reals (SMT(\(LRA\))) delivers all these ingredients by design!

Widely used as a representation language for robotics, verification
and planning [Barrett et al. 2010]

Barrett et al., “Satisfiability modulo theories”, 2018
Continuous + discrete + constraints = SMT

Each player has a certain skill

Barrett et al., “Satisfiability modulo theories”, 2018
Continuous + discrete + constraints = SMT

\[ 0 \leq X_{P_i} \leq 10 \]
for \( i = 1, \ldots, N \)

Barrett et al., “Satisfiability modulo theories”, 2018
Continuous + discrete + constraints = SMT

$0 \leq X_{P_i} \leq 10$

for $i = 1, \ldots, N$

Each team’s skill is bounded by its players’ skills

*Barrett et al., “Satisfiability modulo theories”, 2018*
Continuous + discrete + constraints = SMT

\[ 0 \leq X_{P_i} \leq 10 \]
for \( i = 1, \ldots, N \)

\[ |X_{T_j} - X_{P_i}| < 1 \]
for \( j = 1, \ldots, M, i = 1, \ldots, |T_j| \)

Barrett et al., “Satisfiability modulo theories”, 2018
Continuous + discrete + constraints = SMT

- $0 \leq X_{P_i} \leq 10$ for $i = 1, \ldots, N$
- $|X_{T_j} - X_{P_i}| < 1$ for $j = 1, \ldots, M, i = 1, \ldots, |T_j|$
- Good teams form a squad

Barrett et al., “Satisfiability modulo theories”, 2018
Continuous $+$ discrete $+$ constraints $=$ SMT

\[ 0 \leq X_{P_i} \leq 10 \quad \text{for } i = 1, \ldots, N \]

\[ |X_{T_j} - X_{P_i}| < 1 \quad \text{for } j = 1, \ldots, M, i = 1, \ldots, |T_j| \]

\[ B_{S_j} \Rightarrow X_{T_j} > 2 \quad \text{for } j = 1, \ldots, M, i = 1 \]

Barrett et al., “Satisfiability modulo theories”, 2018
Continuous + discrete + constraints = SMT

\[ \Delta = \bigwedge_{i} 0 \leq X_{P_i} \leq 10 \bigwedge_{j} \bigwedge_{i \in T_j} |X_{T_j} - X_{P_i}| < 1 \bigwedge_{j} (B_{S_j} \Rightarrow X_{T_j} > 2) \]

a single CNF SMT(\(\mathcal{LRA}\)) formula \(\Delta\)...
Continuous + discrete + constraints = SMT

a single CNF SMT(\mathcal{LRA}) formula $\Delta$...and its \textit{primal graph}

\textit{Barrett et al., “Satisfiability modulo theories”, 2018}
\[
\begin{aligned}
&\bigwedge_{i} 0 \leq X_{P_i} \leq 10 \\
&\bigwedge_{j} \bigwedge_{i \in T_j} |X_{T_j} - X_{P_i}| < 1 \\
&\bigwedge_{j} (B_{S_j} \Rightarrow X_{T_j} > 2)
\end{aligned}
\]

\[
\begin{aligned}
&w(X_{P_i}), \\
&\text{if } 0 \leq X_{P_i} \leq 10 \\
&w(X_{T_j}, X_{P_i}), \\
&\text{if } |X_{T_j} - X_{P_i}| < 1 \\
&w(B_{S_j}, X_{T_j}), \\
&\text{if } B_{S_j} \Rightarrow X_{T_j} > 2
\end{aligned}
\]

**SMT formula** \(\triangle\) **weight functions** \(\triangledown\)

---

*Belle et al., “Probabilistic inference in hybrid domains by weighted model integration”, 2015*
$\textbf{SMT} + \textbf{weights} = \textbf{Weighted Model Integration}$

\[\bigwedge_i 0 \leq X_{P_i} \leq 10\]
\[\bigwedge_j \bigwedge_{i \in T_j} |X_{T_j} - X_{P_i}| < 1\]
\[\bigwedge_j (B_{S_j} \Rightarrow X_{T_j} > 2)\]

\[w(X_{P_i}), \text{ if } 0 \leq X_{P_i} \leq 10\]
\[w(X_{T_j}, X_{P_i}), \text{ if } |X_{T_j} - X_{P_i}| < 1\]
\[w(B_{S_j}, X_{T_j}), \text{ if } B_{S_j} \Rightarrow X_{T_j} > 2\]

complex support  \hspace{2cm} \textbf{densities}  \hspace{2cm} \text{(unnormalized)}

$\text{Pr}_\Delta(X, B)$

Belle et al., “Probabilistic inference in hybrid domains by weighted model integration”, 2015
\textbf{SMT} $+$ densities $=$ \textbf{Weighted Model Integration}

Given an SMT($\mathcal{LRA}$) formula $\Delta$ over continuous vars $X$ and discrete ones $B$, and weight function $\mathcal{W}$, the \textit{weighted model integral} (WMI) is

$$\text{WMI}(\Delta, \mathcal{W}; X, B) \triangleq \sum_{b \in B | B \models (x, b) \models \Delta} \int_{(x, b) \models \Delta} w(x, b) \, dx.$$ 

i.e., computing the \textit{partition function} of the unnormalized distribution $\Pr_{\Delta}$

$\Rightarrow$ i.e., integrating the \textit{weighted volumes of the feasible regions of} $\Delta$!

---

Belle et al., “Probabilistic inference in hybrid domains by weighted model integration”, 2015
“What is the probability of team $T_1$ to outperform team $T_2$, if $T_1$ is a squad but $T_2$ is not?”

Belle et al., “Probabilistic inference in hybrid domains by weighted model integration”, 2015
Advanced probabilistic reasoning

\( \Phi_S : (B_{S_1} = 1 \land B_{S_2} = 0) \implies T_1 \text{ is a squad, } T_2 \text{ is not} \)

\( \Phi_T : (X_{T_1} > X_{T_2}) \implies T_1 \text{ outperforms } T_2 \)

Belle et al., “Probabilistic inference in hybrid domains by weighted model integration”, 2015
\[ \Phi_S : (B_{S_1} = 1 \land B_{S_2} = 0) \quad \implies \quad T_1 \text{ is a squad, } T_2 \text{ is not} \]

\[ \Phi_T : (X_{T_1} > X_{T_2}) \quad \implies \quad T_1 \text{ outperforms } T_2 \]

\[
\Pr_{\Delta}(\Phi_T \mid \Phi_S) = \frac{\text{WMI}(\Delta \land \Phi_T \land \Phi_S, \mathcal{W})}{\text{WMI}(\Delta \land \Phi_S, \mathcal{W})} = \frac{4,206}{7,225} \approx 58.22\%
\]

\[ \implies \text{ conditional probabilities as a ratio of two weighted model integrals} \]

Belle et al., “Probabilistic inference in hybrid domains by weighted model integration”, 2015
Tractability of WMI

#P-hard in general
tree-shaped primal graph

tractable weight conditions

polytime WMI inference

\[ \begin{align*}
    w(X_{P_i}) &= X_{P_i}^2 \\
    w(X_{T_j}, X_{P_i}) &= X_{T_j} + X_{P_i} \\
    w(B_{S_j}) &= 100
\end{align*} \]

\[ \text{treeWMI} \]

[Zeng et al. 2020]

Zeng et al., “Scaling up Hybrid Probabilistic Inference with Logical and Arithmetic Constraints via Message Passing”, 2020
Tractability of WMI

WMI

#P-hard in general

Zeng et al., “Scaling up Hybrid Probabilistic Inference with Logical and Arithmetic Constraints via Message Passing”, 2020
Tractability of WMI

- #P-hard in general
- largest tractable class known so far
- can we expand it?

---

Zeng et al., “Scaling up Hybrid Probabilistic Inference with Logical and Arithmetic Constraints via Message Passing”, 2020
Tractability of WMI

- #P-hard in general
- largest tractable class known so far
- can we expand it? No!

Zeng et al., “Scaling up Hybrid Probabilistic Inference with Logical and Arithmetic Constraints via Message Passing”, 2020
Thm. Let $\text{WMI}(\text{tw})$ be the class of WMI problems with bounded diameter and treewidth $\text{tw}$. $\text{WMI}(\text{tw})$ is a tractable WMI class iff treewidth $\text{tw} = 1$. 

#P-hard in general

largest tractable class known so far

can we expand it? No!
ReCoIn

Approximate WMI Inference
Given a WMI problem with *loopy primal graph*
Given a WMI problem with **loopy primal graph**

Relax it by adding **copies** of literals
Given a WMI problem with **loopy primal graph**

- Relax it by adding *copies* of literals, then removing equality constraints
  - removing dependencies, breaking loops
Given a WMI problem with *loopy primal graph*

Relax it by adding *copies* of literals, then removing equality constraints

Compensate for the removed dependencies, by introducing certain literals and weights
ReCoIn
Approximate WMI Inference

\[ w_\ell \leftarrow f(\Pr_\Delta(\ell); w') \]

\[ w'_\ell \leftarrow f(\Pr_\Delta(\ell); w) \]

- Given a WMI problem with \textit{loopy primal graph}
- \textit{Relax} it by adding \textit{copies} of literals, then removing equality constraints
- \textit{Compensate} for the removed dependencies, by introducing certain literals and weights
- Optimize compensating weights iteratively by solving a series of exact \textit{Integration} problems
ReCoIn better scales to larger WMI problems than other approximate WMI solvers while still delivering accurate approximations.
Conclusions

Real-world data is noisy...
Conclusions

Real-world data is *noisy, complex*...
Real-world data is *noisy, complex* and *mixed continuous-discrete*...
Conclusions

Real-world data is noisy, complex and mixed continuous-discrete...

The WMI framework is very appealing for probabilistic inference in the real-world!
Conclusions

Real-world data is **noisy, complex** and **mixed continuous-discrete**...

**The WMI framework** is very appealing for probabilistic inference in the real-world!

Efficient approximations are not only useful, but **needed**
Conclusions

Real-world data is nois[y, complex and mixed continuous-discrete... The WMI framework is very appealing for probabilistic inference in the real-world! Efficient approximations are not only useful, but needed

⇒ ReCoIn delivers fast approximate inference
Real-world data is **noisy**, **complex** and **mixed continuous-discrete**...

*The WMI framework* is very appealing for probabilistic inference in the real-world! Efficient approximations are not only useful, but **needed**

⇒ *ReCoIn delivers fast approximate inference*

**Next**

Application to program verification, probabilistic (logic) programming, ...
Conclusions

Real-world data is *noisy, complex* and *mixed continuous-discrete*...

*The WMI framework* is very appealing for probabilistic inference in the real-world!

Efficient approximations are not only useful, but *needed*

⇒ *ReCoIn delivers fast approximate inference*

Next

Application to program verification, probabilistic (logic) programming, ...

Questions?
References


Minka, Tom, Ryan Cleven, and Yordan Zaykov (2018). “Trueskill 2: An improved bayesian skill rating system". In:
