Statistical Relational AI
Exploiting Symmetries

Tanya Braun
University of Münster, Germany

Marcel Gehrke
University of Lübeck, Germany

Marco Wilhelm
TU Dortmund University, Germany

KR 2023 Tutorial, Rhodes (Greece), 04/09/2023

Agenda

- Introduction [Tanya]
- Exploiting Symmetries
  — in Probabilistic Graphical Models [Marcel]

Focus Topic II

- Exploiting Symmetries
  — in Conditional Knowledge Bases [Marco]
  - Relational Probabilistic Conditionals
  - Principle of Maximum Entropy
  - Conditional Knowledge Compilation

- Summary [Tanya]

Relational Probabilistic Conditionals

Relational Background Theory

Definition (Function-free First-order Language)

$\mathcal{L}(\Sigma)$ is a function-free first-order language over a finite signature $\Sigma = (\mathrm{Pred}, \mathrm{Const})$.

- Sentences in $\mathcal{L}(\Sigma)$ interpreted based on Herbrand semantics.
  $\rightsquigarrow$ Interpretations $\widehat{=}$ truth assignment on ground atoms.
- Interpretations serve as possible worlds $\omega \in \Omega(\Sigma)$.
- Controlling domain size $k = |\mathrm{Const}|$ crucial for tractable reasoning.

$\mathcal{L}(\Sigma)$ $\widehat{=}$ relational language with quantifiers; quantifiers do not extend expressivity but convey symmetries.
$\rightsquigarrow$ Proper handling essential, esp. for efficient model counting.

Closed Conditionals

Definition (Closed Conditional)

$(B|A)[\xi]$ is a closed conditional if $A, B \in \mathcal{L}(\Sigma)$ are sentences and $\xi \in [0,1]$.

- Informal meaning:

  "If A holds, then B follows with probability ξ."

- Formal interpretation via conditional probabilities.

Definition (Probabilistic Model)

$\mathcal{P}\colon \Omega(\Sigma) \to [0,1]$ is a probabilistic model of a closed conditional $(B|A)[\xi]$, written

$$\mathcal{P} \models (B|A)[\xi] \quad \text{iff} \quad \mathcal{P}(A) > 0 \quad \text{and} \quad \mathcal{P}(B|A) = \xi.$$

$(\mathcal{P}(A) = \sum_{\omega \models A} \mathcal{P}(\omega))$

Open Conditionals

Definition (Open Conditional)

$(B|A)[\xi]$ is an (open) conditional if $A, B \in \mathcal{L}(\Sigma)$ are formulas and $\xi \in [0,1]$.

Example

"Patients who show symptom s suffer from disease d with probability 0.6."

- s, d as nullary predicates: closed conditional $(d|s)[0.6]$.
  (no linkage to individuals)
- s, d as unary predicates: open conditional $(d(X)|s(X))[0.6]$.

Problem Open conditionals need richer semantics than conditional probabilities provide.

Grounding Semantics

Idea Understand open conditionals as schemata / universally quantified.

Definition (Grounding Semantics)

$\mathcal{P}$ probability distribution, $r = (B|A)[\xi]$ (open) conditional, and $\mathrm{grnd}_{\mathrm{Const}}(r)$ set of proper groundings of r. Then,

$$\mathcal{P} \models_{\mathrm{gs}} r \quad \text{iff} \quad \forall r' \in \mathrm{grnd}_{\mathrm{Const}}(r) \colon \mathcal{P} \models r'.$$

Example

$(d(X)|s(X))[0.6]$ shortcut for $\{(d(p)|s(p))[0.6] \mid p \in \mathrm{Const}\}$.

Problem Not compatible with atypical individuals (e.g., if $\exists p \in \mathrm{Const} \colon (d(p)|(s(p)))[0.3]$).

Averaging Semantics

Idea Understand probabilities as arithmetic mean.

Definition (Averaging Semantics [Kern-Isberner, Thimm 2010])

$\mathcal{P}$, r, $\mathrm{grnd}_{\mathrm{Const}}(r)$ as before. Then,

$$\mathcal{P} \models_{\mathrm{as}} r \quad \text{iff} \quad \frac{\sum_{(B'|A')[\xi] \in \mathrm{grnd}_{\mathrm{Const}}(r)} \mathcal{P}(B'|A')}{|\mathrm{grnd}_{\mathrm{Const}}(r)|} = \xi.$$

Example

$$\mathcal{P} \models_{\mathrm{as}} ((d(X))|s(X))[0.6] \quad \text{iff} \quad \frac{\sum_{p \in \mathrm{Const}} \mathcal{P}(d(p)|s(p))}{|\mathrm{Const}|} = 0.6.$$

Problem In general, set of models not convex.
$\rightsquigarrow$ Maximum entropy model not unique.

Aggregating Semantics (I/IV)

Idea Statistic probabilities from a subjective point of view.

Definition (Aggregating Semantics [Kern-Isberner, Thimm 2010])

$\mathcal{P}$, r, $\text{grnd}_{\text{Const}}(r)$ as before. Then,

$$\mathcal{P} \models r \quad \text{iff} \quad \frac{\sum_{(B'|A')[\xi] \in \text{grnd}_{\text{Const}}(r)} \mathcal{P}(A' \wedge B')}{\sum_{(B'|A')[\xi] \in \text{grnd}_{\text{Const}}(r)} \mathcal{P}(A')} = \xi.$$

Example

$$\mathcal{P} \models ((d(X))|s(X))[0.6] \quad \text{iff} \quad \frac{\sum_{p \in \text{Const}} \mathcal{P}(s(p) \wedge d(p))}{\sum_{p \in \text{Const}} \mathcal{P}(s(p))} = 0.6.$$

possible worlds included multiple times $\rightsquigarrow$ double sums

Aggregating Semantics (II/IV)

Definition (Aggregating Semantics [Kern-Isberner, Thimm 2010])

$\mathcal{P}$, r, $\mathrm{grnd}_{\mathrm{Const}}(r)$ as before. Then,

$$\mathcal{P} \models r \quad \text{iff} \quad \frac{\sum_{(B'|A')[\xi] \in \mathrm{grnd}_{\mathrm{Const}}(r)} \sum_{\omega \models A' \wedge B'} \mathcal{P}(\omega)}{\sum_{(B'|A')[\xi] \in \mathrm{grnd}_{\mathrm{Const}}(r)} \sum_{\omega \models A'} \mathcal{P}(\omega)} = \xi.$$

For each ground instance, WFOMC tasks with weights that have to be calculated. (here, $\mathcal{P}$ will be inferred from a knowledge base)

Problem **Can this be done efficiently?**

$\rightsquigarrow$ Only if WFOMC tasks are symmetric.

(to some extent)

Aggregating Semantics (III/IV)

Definition (Conditional Structure (I/II) [Kern-Isberner 2004])

$r = (B|A)[\xi]$ conditional, ω possible world. Then,

$$\mathrm{ver}_r(\omega) = |\{(B'|A')[\xi] \in \mathrm{grnd}_{\mathrm{Const}}(r) \mid \omega \models A' \wedge B'\}|,$$
$$\mathrm{fal}_r(\omega) = |\{(B'|A')[\xi] \in \mathrm{grnd}_{\mathrm{Const}}(r) \mid \omega \models A' \wedge \neg B'\}|.$$

Definition (Aggregating Semantics [Kern-Isberner, Thimm 2010])

Let $\mathcal{P}$, r, $\mathrm{grnd}_{\mathrm{Const}}(r)$ as before. Then,

$$\mathcal{P} \models r \quad \text{iff} \quad \frac{\sum_{\omega \in \Omega(\Sigma)} \mathrm{ver}_r(\omega) \cdot \mathcal{P}(\omega)}{\sum_{\omega \in \Omega(\Sigma)} (\mathrm{ver}_r(\omega) + \mathrm{fal}_r(\omega)) \cdot \mathcal{P}(\omega)} = \xi.$$

Justification of Aggregating Semantics

Statistic probabilities in extreme cases,
otherwise statistics weighted by subjective beliefs.

Case 1 **Reasoner is certain about model of real world**

$\rightsquigarrow \mathcal{P}(\omega') = 1$ for a single possible world ω'. Then,

$$\mathcal{P} \models r \quad \text{iff} \quad \frac{\mathsf{ver}_r(\omega')}{\mathsf{ver}_r(\omega') + \mathsf{fal}_r(\omega')} = \xi.$$

Case 2 **Reasoner is maximally uncertain**

$\rightsquigarrow \mathcal{P}(\omega) = \frac{1}{|\Omega(\Sigma)|}$ for $\omega \in \Omega(\Sigma)$. Then,

$$\mathcal{P} \models r \quad \text{iff} \quad \frac{\sum_{\omega \in \Omega(\Sigma)} \mathsf{ver}_r(\omega)}{\sum_{\omega \in \Omega(\Sigma)} (\mathsf{ver}_r(\omega) + \mathsf{fal}_r(\omega))} = \xi.$$

Probabilistic Inference Task

Definition (Knowledge Base)

$\mathcal{R} = (\mathcal{F}, \mathcal{B})$ is a knowledge base if $\mathcal{F}$ is finite set of sentences and $\mathcal{B}$ finite set of non-deterministic conditionals ($\xi \notin \{0, 1\}$).

- Sentences ($\widehat{=}$ facts) force specific possible worlds to have probability zero / restrict probability space.
  $\rightsquigarrow \Omega_{\mathcal{F}}(\Sigma) = \{\omega \in \Omega(\Sigma) \mid \omega \models \mathcal{F}\}$.

Definition (Probabilistic Inference Task)

$\mathcal{R}$ consistent knowledge base. Then,

1. Calculate model $\mathcal{P}$ of $\mathcal{R}$. (model selection task)
2. a. Given conditional r, decide $\mathcal{R} \models_{\mathcal{P}} r$?
   b. Given formulas A, B, calculate ξ such that $\mathcal{R} \models_{\mathcal{P}} (B|A)[\xi]$.

Principle of Maximum Entropy

Maximum Entropy Model (I/III)

Model selection $\widehat{=}$ from knowledge base to belief state.

In principal not necessary. E.g., also possible:

$$\mathcal{R} \models_{\mathcal{P}} r \quad \text{if} \quad \mathcal{P} \models r \quad \text{for all models } \mathcal{P} \text{ of } \mathcal{R}.$$

Problem Leads to few / uninformative inferences.

Idea Select model which infers conditionals with most expected probability / which adds least information.

Definition (Maximum Entropy Model [cf. Paris 1994])

$\mathcal{R}$ consistent knowledge base. Then,

$$\mathrm{ME}_{\mathcal{R}} = \arg\max_{\mathcal{P} \models \mathcal{R}} - \sum_{\omega \in \Omega_{\mathcal{F}}(\Sigma)} \mathcal{P}(\omega) \cdot \log \mathcal{P}(\omega).$$

Maximum Entropy Model (II/III)

Proposition

$\mathcal{R}$ consistent knowledge base. Then,

$$\mathrm{ME}_{\mathcal{R}} \models \mathcal{R}. \qquad \text{(i.e., } \mathrm{ME}_{\mathcal{R}} \models r \; \forall r \in \mathcal{R}\text{)}$$

- $\mathrm{ME}_{\mathcal{R}}$ unique solution of convex optimization problem.
  $\rightsquigarrow$ Dual optimization problem [cf. Boyd, Vandenberghe 2004].

Product Representation

$r_i = (B_i|A_i)[\xi_i]$

$\mathcal{R}$ p-consistent*, $\mathcal{B} = \{r_1, \ldots, r_n\}$. Then, there is $\vec{\alpha} \in \mathbb{R}^n_{>0}$ s.t.

$$\mathrm{ME}_{\mathcal{R}}(\omega) = \alpha_0 \cdot \prod_{i=1}^{n} \alpha_i^{\mathrm{ver}_{r_i}(\omega) - \xi_i \cdot (\mathrm{ver}_{r_i}(\omega) + \mathrm{fal}_{r_i}(\omega))}, \quad \omega \in \Omega_{\mathcal{F}}(\Sigma).$$

(α_0 normalization constant)

*) has positive model on $\Omega_{\mathcal{F}}(\Sigma)$, i.e., 0-1-probabilities forced by sentences only

Maximum Entropy Model (III/III)

$$\mathrm{ME}_{\mathcal{R}}(\omega) = \alpha_0 \cdot \prod_{i=1}^{n} \alpha_i^{\mathrm{ver}_{r_i}(\omega) - \xi_i \cdot (\mathrm{ver}_{r_i}(\omega) + \mathrm{fal}_{r_i}(\omega))}$$

- Optimize for $\vec{\alpha}$ (n-many values; constant in k)
  instead of $\mathrm{ME}_{\mathcal{R}}$. ($|\Omega(\Sigma)|$-many values; exponential in k)

Definition (Conditional Structure (II/II) [Kern-Isberner 2004])

$\mathcal{R}$ knowledge base, $\omega, \omega' \in \Omega_{\mathcal{F}}(\Sigma)$. Then,

$$\sigma_{\mathcal{R}}(\omega) = ((\mathrm{ver}_{r_i}(\omega), \mathrm{fal}_{r_i}(\omega)))_{i=1}^{n}, \qquad \text{(conditional structure)}$$

$$\omega \sim_{\sigma_{\mathcal{R}}} \omega' \quad \text{iff} \quad \sigma_{\mathcal{R}}(\omega) = \sigma_{\mathcal{R}}(\omega). \qquad \text{(conditional equivalence)}$$

Principle of Conditional Indifference [Kern-Isberner 2004]

$$\omega \sim_{\sigma_{\mathcal{R}}} \omega' \implies \mathrm{ME}_{\mathcal{R}}(\omega) = \mathrm{ME}_{\mathcal{R}}(\omega'), \quad \omega, \omega' \in \Omega_{\mathcal{F}}(\Sigma).$$

Maximum Entropy Inference

Definition (Maximum Entropy Inference Relation)

$\mathcal{R}$ consistent knowledge base, r conditional. Then,

$$\mathcal{R} \mid\!\sim_{\mathsf{ME}} r \quad \text{iff} \quad \mathsf{ME}_{\mathcal{R}} \models r.$$

Task 1 Given $\mathcal{R}$, approximate $\vec{\alpha}$.

Task 2 Given approximation $\vec{\alpha}^*$ and r, decide $\mathsf{ME}^*_{\mathcal{R}} \models r$.

Evolution of Iterative Scaling Methods for Task 1

- Generalized iterative scaling (approx. $\mathsf{ME}_{\mathcal{R}}$) [Darroch, Ratcliff 1972]
- Improved iterative scaling (approx. $\vec{\alpha}$) [Berger, Della Pietra, Della Pietra 1996]
- Iterative scaling + conditional equivalence [Finthammer, Beierle 2014]
- Condensed iterative scaling [W, Kern-Isberner, Finthammer, Beierle 2019]

Excursion: Maximum Entropy and MLNs

Definition (Markov Logic Network [Richardson, Domingos 2006])

$\mathcal{N} = \{(F_i, \nu_i) \mid i = 1, \dots, m\}$ is a Markov logic network (MLN) if $F_i \in \mathcal{L}(\Sigma)$ are formulas and $\nu_i \in \mathcal{R} \cup \{\infty, -\infty\}$ are weights.
$\mathcal{N}$ specifies probability distribution ↖ for hard constraints

$$\mathcal{P}_{\mathcal{N}}(\omega) = \frac{1}{\eta} \cdot \exp\left(\sum_{i=1}^{m} \nu_i \cdot \mathsf{cnt}_i(\omega)\right), \quad \omega \in \Omega(\Sigma),$$

with $\mathsf{cnt}_i(\omega) = |\{F_i' \in \mathsf{grnd}_{\mathsf{Const}}(F_i) \mid \omega \models F_i'\}|$. ($\eta$ normalization constant)

Proposition [W, Kern-Isberner, Finthammer, Beierle 2019]

$\mathcal{R} = (\mathcal{F}, \mathcal{B})$ p-consistent knowledge base (with $r_i = (B_i|A_i)[\xi] \in \mathcal{B}$) and MLN

$$\mathcal{N} = \bigcup_{i=1}^{n} (\{(A_i \wedge B_i, (1 - \xi_i) \cdot \log \alpha_i)\} \cup \{(A_i \wedge \neg B_i, -\xi_i \cdot \log \alpha_i)\}) \cup \{(\neg F, -\infty) \mid F \in \mathcal{F}\}.$$

Then, $\mathrm{ME}_{\mathcal{R}} = \mathcal{P}_{\mathcal{N}}$.

Conditional Knowledge Compilation

Starting Point/Motivation: Condensed Iterative Scaling

Input: p-Consistent knowledge base $\mathcal{R}$
Output: Approximation $\vec{\alpha}^*$, normalization α_0^*

```
G = ∑_{i=1}^{n} |grnd_Const(r_i)|
k = 0
FOR i = 1, ..., n: α_i^k = 1
UNTIL ⟨termination condition⟩ DO :
    k ← k + 1
    FOR i = 1, ..., n: α_i^k = α_i^{k-1} · (1 + Φ_R^{i,k-1} / (ξ_i · |grnd_Const(r_i)| · Φ_R^{k-1}))^{-1/G}
    α_0^k = (Φ_R^k)^{-1}
RETURN (α_1^k, ..., α_n^k), α_0^k
```

- $\lim_{k\to\infty}(\alpha_1^k, \ldots, \alpha_n^k) = \vec{\alpha}, \quad \lim_{k\to\infty} \alpha_0^k = \alpha_0$
- No expensive iterations ($n = |\mathcal{B}|$ constant in k)

Problem Oracle needed for

$$\Phi_{\mathcal{R}}^{i,k} = \sum_{\omega\in\Omega_{\mathcal{F}}(\Sigma)} (\mathrm{ver}_{r_i}(\omega) + \xi_i \cdot (\mathrm{ver}_{r_i}(\omega) + \mathrm{fal}_{r_i}(\omega))) \cdot \Pi_{\mathcal{R}}^k$$

conditional knowledge compilation is needed to stay tractable

$$\Phi_{\mathcal{R}}^{k} = \sum_{\omega\in\Omega_{\mathcal{F}}(\Sigma)} \Pi_{\mathcal{R}}^k, \quad \Pi_{\mathcal{R}}^k = \prod_{j=1}^{n} (\alpha_j^k)^{\mathrm{ver}_{r_j}(\omega)} \cdot ((\alpha_j^k)^{-\xi_j})^{(\mathrm{ver}_{r_j}(\omega)+\mathrm{fal}_{r_j}(\omega))}$$

Conditional Knowledge Compilation (I/III)

- In this tutorial: Focus on

$$\Phi_{\mathcal{R}}^{k} = \sum_{\omega \in \Omega_{\mathcal{F}}(\Sigma)} \prod_{j=1}^{n} (\alpha_j^k)^{\mathsf{ver}_{r_j}(\omega)} \cdot ((\alpha_j^k)^{-\xi_j})^{(\mathsf{ver}_{r_j}(\omega)+\mathsf{fal}_{r_j}(\omega))}.$$

$\widehat{=}$ normalization (*the* expensive task in probabilistic reasoning)

- Calculating $\Phi_{\mathcal{R}}^{i,k}$ and answering queries work analogously.

[W, Kern-Isberner, Finthammer, Beierle 2019]

Task: Calculate polynomial and factorize whenever possible:

$(\sum \prod \rightarrow \prod \sum)$

$$\phi_{\mathcal{R}}(\vec{x}, \vec{y}) = \sum_{\omega \in \Omega_{\mathcal{F}}(\Sigma)} \prod_{j=1}^{n} (x_j)^{\mathsf{ver}_{r_j}(\omega)} \cdot (y_j)^{(\mathsf{ver}_{r_j}(\omega)+\mathsf{fal}_{r_j}(\omega))}$$

Conditional Knowledge Compilation (II/III)

Example

$\mathcal{R} = (\emptyset, \{(d(X)|s(X))[0.6]\})$

- Naïve approach:
  For each $\omega \in \Omega(\Sigma)$, calculate $\sigma_{\mathcal{R}}(\omega)$. ($4^{|\mathrm{Const}|}$-many possible worlds)
- Combinatorial arguments:
  $j = \#$ patients with symptom s, $m = \#$ patients with s and d:

$$\phi_{\mathcal{R}} = \sum_{j=0}^{|\mathrm{Const}|} \sum_{m=0}^{j} \binom{|\mathrm{Const}|}{j} \binom{j}{m} \cdot x_1^m \cdot y_1^j \cdot 2^{|\mathrm{Const}|-j}$$

- Exploiting symmetries:
  Impact of all patients the same (treat independently):

$$\phi_{\mathcal{R}} = (x_1 \cdot y_1 + y_1 + 2)^{|\mathrm{Const}|}$$

Conditional Knowledge Compilation (III/III)

Proposition [W, Kern-Isberner, Ecke 2017]

$\mathcal{R} = (\emptyset, \mathcal{B})$ where conditionals are built upon boolean combinations of unary predicates (e.g., $(d(X)|s_1(X) \wedge \neg s_2(X))$), $c \in \text{Const}$. Then,

$$\mathcal{R} \mid\!\sim_{\text{ME}} (B|A)[\xi] \quad \text{iff} \quad \mathcal{R}\langle c\rangle \mid\!\sim_{\text{ME}} (B\langle c\rangle)|A\langle c\rangle)[\xi]$$

independent of $|\text{Const}|$! ($A\langle c\rangle = A$ grounded by constant c)

Is there a general framework?

Idea: Translate knowledge base into structured sentence, count typed models of structured sentence.

(types reflect conditional structures of models)

$\rightsquigarrow$ First-order typed model counting

[W, Finthammer, Kern-Isberner, Beierle 2017]

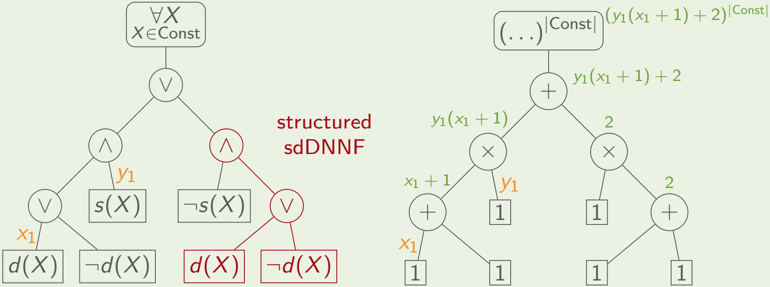
Typed Model Counting (I/III)

Example

$\mathcal{R} = (\emptyset, \{(d(X)|s(X))[0.6]\})$

$$\phi^s_{\mathcal{R}} = \forall X.[y_1 \circ s(X) \wedge (x_1 \circ d(X) \vee \neg d(X)) \vee \neg s(X)]$$

Typed first-order circuit **Algebraic circuit**



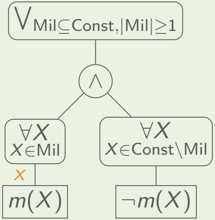
Typed Model Counting (II/III)

- Typed model counting works for arbitrary structured sentences as long as no structure elements are in scope of negations. (1)
- Same model counting techniques as in WFOMC of sentences in sdDNNF can be applied as long as (1) is guaranteed.

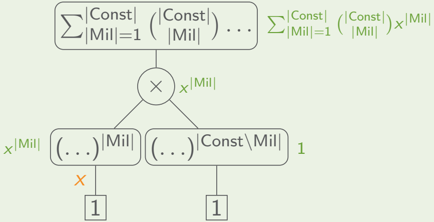
Example

$\phi^s = \exists X . x \circ m(X)$ (m = millionaire)

Typed first-order circuit



Algebraic circuit



Typed Model Counting (III/III)

Algebraic Model Counting [Kimmig, Van den Broeck, De Raedt 2017]

- (Symmetric) WFOMC of sentences.
- Ground atoms are interpreted by weights.
- Weights are elements of an algebraic semiring.

Typed Model Counting [W, Finthammer, Kern-Isberner, Beierle 2017]

- Elements of semiring are written directly into sentence.
- Ground atoms are interpreted by 1.
- Disadvantage: Structured background language necessary.
- Advantage: No external processing of weights.
  $\rightsquigarrow$ Symmetry between weights easy to handle.

Proposition

Typed model counting problems can be transformed into algebraic model counting problems.

Closing Remarks

Summary of Focus Topic II

- (Only) open conditionals make full use of relational language.
- Aggregating semantics gives open conditionals a sophisticated meaning. [Kern-Isberner, Thimm 2010; Thimm, Kern-Isberner 2012]
- Principle of maximum entropy provides a model of conditional knowledge bases which is preferable from an information theoretically point of view, [Jayes 1957; Shannon 1948]

  and also from a commonsense point of view.

  [Paris, Vencovská 1990; Paris 1994; Paris 1998]
- Concepts from symmetric WFOMC can be transferred to conditional maximum entropy reasoning via typed model counting / algebraic model counting.

  [Van den Broeck 2013; Kimmig, Van den Broeck, De Raedt 2017; W, Finthammer, Kern-Isberner, Beierle 2017]

  $\rightsquigarrow$ In specific cases, drawing lifted inferences possible.

Bibliography (I/III)

- **A. Berger, S. Della Pietra, V. Della Pietra** (1996): A maximum entropy approach to natural language processing. *Computational Linguistics 22(1)*
- **S. Boyd, L. Vandenberghe** (2004): Convex Optimization. *Cambridge University Press*
- **G. Van den Broeck** (2013): Lifted Inference and Learning in Statistical Relational Models (thesis). *KU Leuven, Belgium*
- **J. N. Darroch, D. Ratcliff** (1972): Generalized iterative scaling for log-linear models. *Ann. Math. Stat. 43(5)*
- **M. Finthammer, C. Beierle** (2014): A Two-Level Approach to Maximum Entropy Model Computation for Relational Probabilistic Logic Based on Weighted Conditional Impacts. *Proceedings of SUM 2014*
- **E. T. Jaynes** (1957): Information Theory and Statistical Mechanics (I + II). *Physical Review. Series II 106(4) + 108(2)*
- **A. Kimmig, G. Van den Broeck, L. De Raedt** (2017): Algebraic model counting. *J. Appl. Log. 22*

Bibliography (II/III)

- **G. Kern-Isberner** (2004): A Thorough Axiomatization of a Principle of Conditional Preservation in Belief Revision. *Ann. Math. Artif. Intell. 40(1-2)*
- **G. Kern-Isberner, M. Thimm** (2010): Novel Semantical Approaches to Relational Probabilistic Conditionals. *Proceedings of KR 2010*
- **J. B. Paris, A. Vencovská** (1990): A note on the inevitability of maximum entropy. *Int. J. Approx. Reason. 4(3)*
- **J. B. Paris** (1994): The Uncertain Reasoner's Companion: A Mathematical Perspective. *Cambridge University Press*
- **J. B. Paris** (1998): Common Sense and Maximum Entropy. *Synth. 117(1)*
- **M. Richardson, P. Domingos** (2006): Markov Logic Networks. *Machine Learning 62(1-2)*
- **C. E. Shannon** (1948): A Mathematical Theory of Communication. *Bell System Technical Journal 27(3) + 27(4)*

Bibliography (III/III)

- **M. Thimm, G. Kern-Isberner** (2012): On probabilistic inference in relational conditional logics. *Log. J. IGPL 20(5)*
- **M. Wilhelm, M. Finthammer, G. Kern-Isberner, C. Beierle** (2017): First-Order Typed Model Counting for Probabilistic Connditional Reasoning at Maximum Entropy. *Proceedings of SUM 2017*
- **M. Wilhelm, G. Kern-Isberner, A. Ecke** (2017): Basic Independence Results for Maximum Entropy Reasoning Based on Relational Conditionals. *Proceedings of GCAI 2017*
- **M. Wilhelm, G. Kern-Isberner, M. Finthammer, C. Beierle** (2019): Integrating Typed Model Counting into First-Order Maximum Entropy Computations and the Connection to Markov Logic Networks. *Proceedings of FLAIRS 2019*

Agenda

Statistical Relational AI – Exploiting Symmetries
Tanya Braun, Marcel Gehrke, Marco Wilhelm

- Introduction [Tanya]
- Exploiting Symmetries
  — in Probabilistic Graphical Models [Marcel]
- Exploiting Symmetries
  — in Conditional Knowledge Bases [Marco]
- **Discussion** [all]
- Summary [Tanya]