Probabilistic Logics in Expert Systems

Gabriele Kern-Isberner[1]
Christoph Beierle[2], Marc Finthammer[2], Matthias Thimm[1]

[1]Dept. of Computer Science, TU Dortmund, Germany
[2]Dept. of Computer Science, FernUniversität in Hagen, Germany

DEXA 2011, Toulouse, September 1st, 2011

Probabilities in expert systems

Probabilities are of crucial importance for expert systems –

- Knowledge representation and processing
- Premium, powerful approach to handling uncertainty
- Excellent links to data mining and knowledge acquisition

Why probabilistic logic?

Modelling an expert agent requires elaborate knowledge representation and reasoning facilities.

Probabilistic logics provide both –

- syntax, semantics, and inference capabilities;
- explicitness and transparency;
- theoretical foundations of approaches, but also
- combination with heuristics, or reasoning strategies.

Expert knowledge – example

From the medical domain:

- Knowledge expressed by rules: "If a patient is administered medicament M, then he will become healthy with a probability of 80%."
- Reasoning: "If a patient has symptoms S and T, what is the probability that he suffers from disease D?"
- Incorporating new information → logical change management (revision): "A lab test shows that the patient is allergic to a certain antibiotic, how does this change the plans for therapy?"

Scope of this talk

This talk is about probabilistic modelling – and the relevance of logic for this:

- mainly on beliefs (probability $\notin \{0, 1\}$), not on knowledge (probability $\in \{0, 1\}$),
- but reasoning is important nevertheless.
- We will also talk about principles = quality criteria.

Logics . . .

. . . should help people to reason in a better, more rational way, but should not restrict the rational capacities of people.

Overview of this talk

- Motivation and scope of this talk
- Propositional probabilistic frameworks
  (featuring the MaxEnt principle)
- A new challenge: Relational probabilistic reasoning
- Relational approaches using maximum entropy
- Implementation: The KREATOR toolbox
- Evaluating relational probabilistic approaches from a logical viewpoint
- Conclusion: Summary and ongoing work

Overview of this talk

- Motivation and scope of this talk
- Propositional probabilistic frameworks
  (featuring the MaxEnt principle)
- A new challenge: Relational probabilistic reasoning
- Relational approaches using maximum entropy
- Implementation: The KReator toolbox
- Evaluating relational probabilistic approaches from a logical viewpoint
- Conclusion: Summary and ongoing work

Markov and Bayes networks

Efficient methodologies for propositional probabilistic reasoning:

- Markov nets: Undirected resp. symmetric (conditional) dependencies
- Bayes nets: Directed conditional dependencies, idea of causality

In both approaches:

- Links express and specify conditional dependencies
- Missing links express conditional independencies
- Network structure is crucial for efficient implementations by the use of graph-based algorithms

An alternative to BNs and MNs

What is less well known . . .

- Sets of probabilistic conditional dependencies (i.e. probabilistic rules) $(B|A)[x]$
  $\rightarrow$ conditional knowledge bases $\mathcal{R} = \{(B_1|A_1)[x_1], \ldots, (B_n|A_n)[x_n]\}$ (incomplete knowledge!)
- can be combined with information theory (instead of assumptions on conditional independencies!)
- for probabilistic reasoning $\rightarrow$

The Principle of Maximum Entropy (MaxEnt)

Principle of Maximum Entropy (MaxEnt)

Probabilistic-logical prerequisites:

$\mathcal{L}$	propositional language, base for probabilistic conditionals $(B\|A)[\alpha]$ $(A, B \in \mathcal{L}, \alpha \in [0,1])$
Ω	set of interpretations over $\mathcal{L}$ (elementary events, possible worlds)
P	probability distribution/function over Ω
$P \models (B\|A)[\alpha]$	iff $P(A) > 0$ and $\dfrac{P(AB)}{P(A)} = \alpha$; P is a model of $(B\|A)[\alpha]$.

Medical example

"If a patient is administered medicament A, then he will become healthy with a probability of 80%."

$$(\texttt{outcome} = \texttt{healthy} | \texttt{med\_A})[0.8]$$

Principle of Maximum Entropy (MaxEnt) (cont'd)

By maximizing the entropy

$$H(P) = -\sum_{\omega \in \Omega} P(\omega) \log P(\omega)$$

among the models P of a knowledge base $\mathcal{R}$,
a best[1] model $P_{\mathcal{R}}^{ME}$ of $\mathcal{R}$ can be computed

→ optimized probabilistic inference
→ optimized responses to queries

NB: Entropy might also be known from machine learning (decision trees, feature selection: "Choose the most informative attribute/feature").

[1] best = informative, but most cautious

MaxEnt reasoning and systems

The MaxEnt principle seems to work like a black box, but . . .

- $\longrightarrow$ excellent logical properties [LNAI 2001]
- $\longrightarrow$ optimized commonsense reasoning [Paris 1999]

$\longrightarrow$ most adequate for modelling human-like reasoning

Systems for MaxEnt reasoning:

- SPIRIT [Roedder et al., 2006],
- MECoRe [Finthammer et al., 2009].

MaxEnt example

Medical example – general setting:

- Treatment of a patient who suffers from a perilous bacterial infection
- The infection will possibly cause permanent neurological damage or even death if not treated appropriately.
- Two antibiotics, A and B, might be suited for ending the infection, provided that the bacteria are not resistant to the specific antibiotic.
- Each antibiotic might cause a life-threatening allergic reaction.
- The resistance of the bacteria to a specific antibiotic can be tested, but each test is very time-consuming.

MaxEnt example (cont'd)

Some rules from the knowledge base (see paper):

R_1 : $(\neg\texttt{effect\_A}|\neg\texttt{med\_A} \vee \texttt{resistance\_A})[1.00]$
R_3 : $(\texttt{effect\_A} \Leftrightarrow \texttt{med\_A}|\neg\texttt{resistance\_A})[1.00]$
R_5 : $(\texttt{allergic\_A})[0.10]$
R_7 : $(\texttt{resistance\_A})[0.01]$
R_9 : $(\texttt{med\_A} \wedge \texttt{med\_B})[0.00001]$
R_{10}: $(\texttt{outcome}=\texttt{letal}|\neg\texttt{med\_A} \wedge \neg\texttt{med\_B})[0.10]$
R_{11}: $(\texttt{outcome}=\texttt{healthy}|\neg\texttt{med\_A} \wedge \neg\texttt{med\_B})[0.10]$
R_{12}: $(\texttt{posResT\_A}|\texttt{resistance\_A})[0.97]$
R_{16}: $(\texttt{outcome}=\texttt{letal}|\texttt{med\_A} \wedge \texttt{allergic\_A})[0.99]$
R_{18}: $(\texttt{outcome}=\texttt{healthy}|\texttt{effect\_A})[0.8]$
R_{19}: $(\texttt{outcome}=\texttt{healthy}|\texttt{effect\_B})[0.7]$

MaxEnt example (cont'd)

Find an adequate therapy

The user wants to know which is the best therapy – administer either A, or B, or both.

The system returns the following probabilities:

	healthy	disabled	letal
no antibiotic	0.10	0.80	0.10
only A	0.79	0.06	0.15
only B	0.65	0.23	0.12
A and B	0.94	0.02	0.04

$\rightarrow$ The combined administration of both antibiotics seems to be the best treatment.

Overview of the talk

- Motivation and scope of this talk
- Propositional probabilistic frameworks
  (featuring the MaxEnt principle)
- A new challenge: Relational probabilistic reasoning
- Relational approaches using maximum entropy
- Implementation: The KReator toolbox
- Evaluating relational probabilistic approaches from a logical viewpoint
- Conclusion: Summary and ongoing work

Relational Probabilistic Logics

Propositional logic provides only a limited framework – often, one wants to express information on individuals, and on relationships between individuals (e.g., as in social networks).

What does relational probabilistic reasoning mean?

- Basically, probabilities are considered in a first-order framework,
- with quantification being encoded by probabilities,
- generic beliefs (valid in a population) and specific beliefs (about individuals) can be expressed,
- as well as relationships between individuals (i.e. arity of predicates can be > 1).
- Moreover, probabilistic conditionals should be supported (going beyond first-order!).

Relational example – tornado

Tornado

James is on the road and gets a call from his neighbor saying that the alarm of James' house is ringing. James has some uncertain beliefs about the relationships between burglaries, types of neighborhoods, natural disasters, and alarms. For example, he knows that if there is a tornado threatening his home place, then the probability of a tornado triggering the alarm of his house is 0.9. For James, a response to the query

"What is the probability of an actual burglary?"

would be a very useful information.

Syntactical encoding of knowledge

Tornado in Austin – RME

$$
\begin{aligned}
r_1 &= (alarm(x) \mid burglary(x))\ [0.9] \\
r_2 &= (alarm(x) \mid lives\_in(x, y), tornado(y))\ [0.9]\ \} \\
r_3 &= (burglary(x) \mid nhood(x, bad))\ [0.6] \\
r_4 &= (burglary(x) \mid nhood(x, average))\ [0.4] \\
r_5 &= (burglary(x) \mid nhood(x, good))\ [0.3] \\
r_6 &= (nhood(x, \mathsf{Z}) \mid nhood(x, y))\ [y \neq z]\ [0.0] \\
r_7 &= (lives\_in(x, \mathsf{Z}) \mid lives\_in(x, y))\ [y \neq z]\ [0.0]
\end{aligned}
$$

. . . with the following evidential information:

$$lives\_in(james, yorkshire), \ldots$$

Relational probabilistic reasoning

Basic idea

Make use of propositional techniques after grounding the relational knowledge base appropriately.

- Markov logic networks (MLN)
- Bayesian logic programs (BLP)
- Relational MaxEnt reasoning

Problem for all approaches: Grounding makes the knowledge base huge !

Tornado example: Imagine to ground r_1-r_7 for all inhabitants of Germany!

Markov logic networks (MLN)

A Markov logic network is a set of first-order logic formulas F_i, where each formula F_i is quantified by a real value w_i (a weight).

Note: MLNs do not support conditional probabilities !

An MLN defines a template for constructing ground Markov networks (over a specified universe).

MLN encoding of knowledge

Tornado in Austin – MLN

$$\begin{array}{rll} 2.2 & burglary(x) & => alarm(x) \\ 2.2 & lives\_in(x,y) \wedge tornado(y) & => alarm(x) \\ -0.8 & nhood(x, Good) & => burglary(x) \\ -0.4 & nhood(x, Average) & => burglary(x) \\ 0.4 & nhood(x, Bad) & => burglary(x) \end{array}$$

Bayesian logic programs (BLP)

A Bayesian logic program B is a tuple $B = (C, D, R)$ with

- a (finite) set of Bayesian (conditional) clauses $C = \{c_1, \ldots, c_n\}$,
- a set of conditional probability distributions $D = \{\mathsf{cpd}_{c_1}, \ldots, \mathsf{cpd}_{c_n}\}$,
- and a set of combining functions (one for each Bayesian predicate appearing in C) $R = \{\mathsf{cr}_{p_1}, \ldots, \mathsf{cr}_{p_m}\}$.

Semantics are given to BLPs via transformation into propositional Bayesian networks over a specified universe (dependent on the query!).

BLP encoding of knowledge

Tornado in Austin – BLP

$$
\begin{aligned}
c_1 &: \quad (alarm(x) \,|\, burglary(x)) \\
c_2 &: \quad (alarm(x) \,|\, lives\_in(x, y), tornado(y)) \\
c_3 &: \quad (burglary(x) \,|\, nhood(x))
\end{aligned}
$$

Conditional probability tables for

$$
\begin{aligned}
(alarm(x) = true \quad &| \quad burglary(x) = true) \\
(alarm(x) = true \quad &| \quad burglary(x) = false) \ldots
\end{aligned}
$$

Combination functions, e.g., for $alarm(x)$: Noisy-or (or the like)

Overview of this talk

- Motivation and scope of this talk
- Propositional probabilistic frameworks
  (featuring the MaxEnt principle)
- A new challenge: Relational probabilistic reasoning
- Relational approaches using maximum entropy
- Implementation: The KREATOR toolbox
- Evaluating relational probabilistic approaches from a logical viewpoint
- Conclusion: Summary and ongoing work

Relational MaxEnt techniques

Probabilistic-logical prerequisites:

$\mathcal{L}$ relational language
(i.e., first-order without quantifiers, functions,
and with a finite set of constants)

Ω set of Herbrand interpretations (over ground atoms)

P probability distribution over Ω

The (relational) MaxEnt principle

$$P_{\mathcal{R}}^{ME} = \arg \max_{P \models^{rel} \mathcal{R}} H(P) = -\sum_{\omega \in \Omega} P(\omega) \log P(\omega)$$

Relational probabilistic entailment

The MaxEnt principle can be applied whenever

- the logical notion of a model (semantical entailment relation) is properly defined: $P \models^{rel} \mathcal{R}$

(with respect to this constraint, the entropy can be maximized).

What should $P \models^{rel} (B(\vec{x})|A(\vec{x}))[\alpha]$ mean?

Relational probabilistic entailment (cont'd)

This problem is not solved by BLPs and MLNs –

- the probability distribution defined by a BLP (i.e., the model of the BLP) is not unique, but depends on the query;
- formulas in MLNs are not based on (declarative) probabilities, but on (relative) weights (unclear semantical meaning).

We need to do more logic here!

More motivation for logic

A clear logical interpretation of statements like $P \models^{rel} (B(\vec{x})|A(\vec{x}))[\alpha]$ is not only important for technical reasons, or for people who like logic –

it helps to avoid ambiguities, and may build a bridge between subjective (beliefs of agents) and objective (statistical) interpretations of probabilities.

Usually, in commonsense and expert reasoning, people do both!

Relational probabilistic reasoning – example

$$P(\ likes(clyde, fred) \mid elephant(clyde) \land keeper(fred)\ ) \ = \ 0.9$$

- statement on the degree of belief (as in propositional probabilistic conditional logic),
- thus saying that our subjective belief in $likes(clyde, fred)$ is 0.9

$$P(\ likes(x, fred) \mid elephant(x) \land keeper(fred)\ ) \ = \ 0.3$$

- interpretation is ambiguous as it ranges over a set of individuals:
  - Is our subjective belief for all elephants liking Fred 0.3, or
  - do $30\ \%$ of all elephants like Fred,
  - or something in between?

Relational probabilistic reasoning – example (cont'd)

Subjective and objective interpretations of probabilities can not be strictly separated – people use both:

The statements

$$
\begin{aligned}
P(\, likes(X,Y) \mid elephant(X) \wedge keeper(Y)\,) &= 0.6 \\
P(\, likes(X,fred) \mid elephant(X) \wedge keeper(fred)\,) &= 0.3 \\
P(\, likes(clyde,fred) \mid elephant(clyde) \wedge keeper(fred)\,) &= 0.9
\end{aligned}
$$

have an intuitive commonsense meaning, but their logical interpretation is far from being obvious.

Relational probabilistic entailment and MaxEnt

We make use of three different definitions of relational probabilistic entailment $P \models^{rel} \mathcal{R}$:

- simply, by using a ground version of the knowledge base: $\models^{\mathcal{G}}$, with $\mathcal{G}$ being a grounding operator (grounding semantics);
- by considering the average (conditional) probability of a conditional $(B(\vec{x})|A(\vec{x})$ over all instantiations for $\vec{x}$: $\models^{\varnothing}$ (averaging semantics);
- by aggregating probabilities to define $P(A(\vec{x}))$ and mimick conditional probabilities: $\models^{\odot}$ (aggregating semantics).

Relational MaxEnt principle

$$P_{\mathcal{R}}^{ME_{\circ}} = \arg \max_{P \models^{\circ} \mathcal{R}} \mathcal{H}(P)$$

with $\circ$ being one of $\mathcal{G}$, $\varnothing$, or $\odot$.

Grounding semantics

A grounding operator $\mathcal{G}$ maps relational probabilistic knowledge bases to ground (propositional) knowledge bases over a specified universe U, e.g.

$$\begin{aligned} & (alarm(x) \mid lives\_in(x,y), tornado(y))\ [0.9] \\ \mapsto\ & (alarm(james) \mid lives\_in(james, austin), tornado(austin))\ [0.9] \end{aligned}$$

with $james, austin \in U$.

Grounding semantics

A relational conditional q is $\mathrm{ME}_{\mathcal{G}}$-entailed by the knowledge base $\mathcal{R}$ under the grounding $\mathcal{G}$, in symbols

$$\mathcal{R} \models^{ME}_{\mathcal{G}} q \quad \text{iff} \quad P^{ME}_{\mathcal{G}(\mathcal{R})} \models \mathcal{G}(q),$$

i. e. iff for all groundings $q^* \in \mathcal{G}(q), P^{ME}_{\mathcal{G}(\mathcal{R})} \models q^*$.

Averaging semantics

Averaging semantics

Probabilistic semantics to relational conditionals by averaging subjective beliefs:

$P \models_D^{\varnothing} (\phi(\vec{x}) \,|\, \psi(\vec{x}))[\alpha]$ iff

$$\frac{\sum_{(\phi(\vec{c})|\psi(\vec{c}))\in \text{ground}_D((\phi(\vec{x})|\psi(\vec{x})))} P(\phi(\vec{c}) \mid \psi(\vec{c}))}{|\text{ground}_D{}^{a}(\phi(\vec{x}) \mid \psi(\vec{x}))|} = \alpha$$

[a] All groundings with constants from D

In the ground case:

$$P \models_D^{\varnothing} (\phi(\vec{c}) \,|\, \psi(\vec{c}))[\alpha] \quad \text{iff} \quad P((\phi(\vec{c}) \mid \psi(\vec{c}))) = \alpha$$

$P \models_D^{\varnothing} R$ iff $P \models_D^{\varnothing} r$ for all $r \in R$.

Aggregating semantics

Aggregating semantics

Probabilistic semantics to relational conditionals by aggregating subjective beliefs:

$P \models_D^{\odot} (\phi(\vec{x}) \mid \psi(\vec{x}))[\alpha]$ iff

$$\frac{\sum_{(\phi((\vec{c})|\psi(\vec{c}))\in \text{ground}_D((\phi(\vec{x})|\psi(\vec{x})))} P(\psi(\vec{c})\phi(\vec{c}))}{\sum_{(\phi(\vec{c})|\psi(\vec{c}))\in \text{ground}_D((\phi(\vec{x})|\psi(\vec{x})))} P(\psi(\vec{c}))} = \alpha$$

In the ground case:

$$P \models_D^{\odot} (\phi(\vec{c}) \mid \psi(\vec{c}))[\alpha] \quad \text{iff} \quad P((\phi(\vec{c}) \mid \psi(\vec{c}))) = \alpha$$

$P \models_D^{\varnothing} R$ iff $P \models_D^{\varnothing} r$ for all $r \in R$.

Comparing the three semantics

- Grounding semantics has to solve conflicts (e.g., in the elephant-keeper-example, naive grounding causes conflicts by including both the conflicting conditionals $(\, likes(clyde, fred) \mid elephant(clyde) \wedge keeper(fred) \,) \;=\; 0.9$ and $(\, likes(clyde, fred) \mid elephant(clyde) \wedge keeper(fred) \,) \;=\; 0.3$).
- Averaging and aggregating semantics are similar with subtle differences (average conditional probability vs. aggregated conditional probability); however, it is still an open problem whether the averaging semantics yields a unique MaxEnt distribution.
- If the knowledge base does not mention constants then all three semantical relations coincide.

Overview of the talk

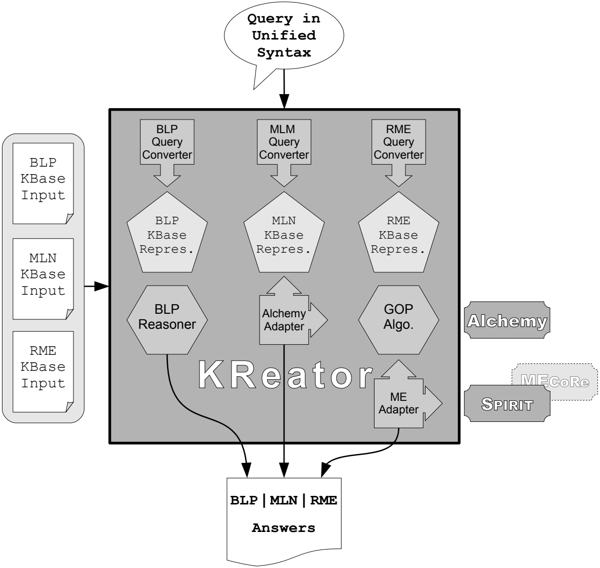
- Motivation and scope of this talk
- Propositional probabilistic frameworks
  (featuring the MaxEnt principle)
- A new challenge: Relational probabilistic reasoning
- Relational approaches using maximum entropy
- Implementation: The KReator toolbox
- Evaluating relational probabilistic approaches from a logical viewpoint
- Conclusion: Summary and ongoing work

The KReator toolbox

KReator

- provides an environment for representing, reasoning, and learning with relational probabilistic knowledge;
- currently supports using BLPs, MLNs, MaxEnt techniques and other approaches;
- allows for easy integration of new approaches via plug-ins;
- makes use of specialized reasoning tools like Alchemy (for MLNs) and SPIRIT (for (propositional) ME reasoning);
- offers a unified query syntax for querying MaxEnt, BLP and MLN knowledge bases.

The KREATOR system – Graphical overview



Comparing MaxEnt grounding with BLP and MLN

...in the tornado example with the following evidential information:

$$lives\_in(james, yorkshire), lives\_in(carl, austin),$$
$$burglary(james), tornado(austin),$$
$$nhood(james) = \text{average}, nhood(carl) = \text{good}$$

Some responses to queries:

queries	BLP	MLN	$\mathrm{ME}_{\mathcal{G}}$
$alarm(james)$	0.900	0.896	0.918
$alarm(carl)$	0.550	0.900	0.880
$burglary(carl)$	0.300	0.254	0.362

Overview of the talk

- Motivation and scope of this talk
- Propositional probabilistic frameworks
- A new challenge: Relational probabilistic reasoning
- Relational approaches using maximum entropy
- Implementation: The KREATOR toolbox
- Evaluating relational probabilistic approaches from a logical viewpoint
- Conclusion: Summary and ongoing work

Criteria for evaluation

In order to compare and evaluate the approaches in more detail, we developed catalogues of criteria:

- informal criteria based on common sense [KI 2011, current work];
- formal criteria based on logic [KR 2010, ECSQARU 2011].

Informal criteria

Interpretation and meaning: Do the elements in the knowledge base have a clear probabilistic and/or commonsense meaning?

Individuals: Is it possible to specify and infer information related to specific individuals?

Universes: Do the inferred probabilities depend on the number of elements in the universe? Is it possible to have an open universe whose elements are not a-priori known?

Compatibility with classical logics: What inferences can be drawn if the knowledge base contains only strict knowledge (i.e., probabilities of 0 or 1), or if all formulas in the knowledge base are ground? Are the inferences compatible with first-order resp. propositional logic in these cases?

Formal criteria

Prototypical indifference: If constants c_1, c_2 are "indistinguishable" with respect to the knowledge base, then the same probabilistic knowledge should be inferred for them.

Convergence: If the knowledge base $\mathcal{R}$ mentions only finitely many constants then for any tuple $\vec{c}$ of constants not mentioned in $\mathcal{R}$, and for any $(B(\vec{x})|A(\vec{x}))[\alpha] \in \mathcal{R}$, the conditional probability of $(B(\vec{x})|A(\vec{x}))$ should converge to α with growing universes.

System P: The usual logical properties for nonmonotonic reasoning (System P) should be satisfied.

Evaluation according to the criteria

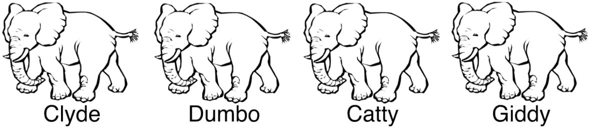
By and large, the following observations can be made:

- BLP and MLN are more robust and efficient, but less well behaved with respect to logical properties (some criteria can not even be applied to them!).
- The relational MaxEnt approaches are more flexible with respect to modelling and match human expectations better, but are also “more logical” (e.g., satisfy all properties of system P).
- We found that logical principles underlie most of the criteria and can help a lot to make them more precise.

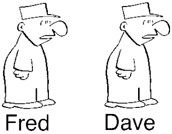
Averaging semantics: example 1/4

Let a knowledge base on elephants and keepers be given by

- $(elephant(clyde))[1] \in R$, ...



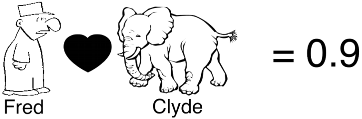
- $(keeper(fred))[1] \in R$, ...



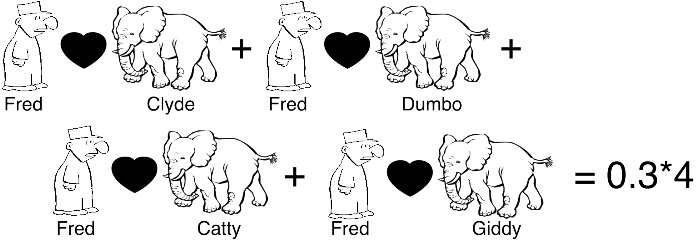
- $(likes(x, y) \mid elephant(x) \wedge keeper(y))[0.6] \in R$
  $(likes(x, fred) \mid elephant(x) \wedge keeper(fred))[0.3] \in R$
  $(likes(clyde, fred) \mid elephant(clyde) \wedge keeper(fred))[0.9] \in R$

Averaging MaxEnt-semantics: example 2/4

- $(likes(clyde, fred) \,|\, elephant(clyde) \wedge keeper(fred))[0.9] \in R$ directly yields the inference $(likes(clyde, fred))[0.9]$



- For $(likes(x, fred) \,|\, elephant(x) \wedge keeper(fred))[0.3]$ to be satisfied under $\models_D^{\varnothing}$ it must hold that



Averaging MaxEnt-semantics: example 3/4

Let $\mathcal{I}_{\varnothing}(R)$ denote the inference operation based on averaging MaxEnt-semantics.

- For the second conditional this yields via (Prototypical Indifference)

$$\begin{aligned}\mathcal{I}_{\varnothing}(R)(\text{likes(dumbo,fred)}) &= \mathcal{I}_{\varnothing}(R)(\text{likes(catty,fred)}) \\ &= \mathcal{I}_{\varnothing}(R)(\text{likes(giddy,fred)}) \\ &= 0.1\end{aligned}$$

as $\underbrace{0.1}_{(dumbo,fred)} + \underbrace{0.1}_{(catty,fred)} + \underbrace{0.1}_{(giddy,fred)} + \underbrace{0.9}_{(clyde,fred)} = 0.3 * 4.$

Averaging MaxEnt-semantics: example 4/4

- Due to the first conditional $(likes(X,Y) \,|\, elephant(X) \wedge keeper(Y))[0.6]$, we have

$$\begin{aligned} \mathcal{I}_{\varnothing}(R)(\text{likes(dumbo,dave)}) &= \mathcal{I}_{\varnothing}(R)(\text{likes(catty,dave)}) \\ &= \mathcal{I}_{\varnothing}(R)(\text{likes(giddy,dave)}) \\ &= \mathcal{I}_{\varnothing}(R)(\text{likes(clyde,dave)}) = 0.9 \end{aligned}$$

as $\underbrace{0.1}_{(dumbo,fred)} + \underbrace{0.1}_{(catty,fred)} + \underbrace{0.1}_{(giddy,fred)} + \underbrace{0.9}_{(clyde,fred)} +$

$\underbrace{0.9}_{(dumbo,dave)} + \underbrace{0.9}_{(catty,dave)} + \underbrace{0.9}_{(giddy,dave)} + \underbrace{0.9}_{(clyde,dave)} = 0.6 * 8.$

Overview of this talk

- Motivation and scope of this talk
- Propositional probabilistic frameworks
  (featuring the MaxEnt principle)
- A new challenge: Relational probabilistic reasoning
- Relational approaches using maximum entropy
- Implementation: The KREATOR toolbox
- Evaluating relational probabilistic approaches from a logical viewpoint
- Conclusion: Summary and ongoing work

Summary

- This talk presented well-known and novel approaches to modelling probabilistic knowledge for expert systems,
- emphasising the coherence between representation and learning of knowledge:
  - All presented approaches are to be used both for learning and knowledge representation.
  - In particular, the maximum entropy techniques play crucial roles for both areas.
- We focused on logical approaches using maximum entropy, thus featuring the combination of two powerful methodologies.

Ongoing work

- Relational probabilistic learning and reasoning is a new challenge, first steps have been taken, a lot of work is still to be done (in particular, with respect to efficient implementations!).
- The presented novel relational probabilistic logics and inferences still have to be evaluated in more details, and corresponding learning algorithms have to be developed.
- Actually, the newly developed relational probabilistic semantics have to be investigated more thoroughly.
- Most importantly: More efficient algorithms have to be developed, at least for special cases (by, e.g., using lifted inference techniques).