# Managing Uncertainty and Vagueness in Semantic Web Languages

# Tutorial at SWAP-2007

Umberto Straccia

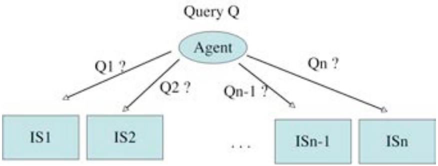
ISTI-CNR, Pisa, Italy
straccia@isti.cnr.it

# Sources of Uncertainty and Vagueness on the Web

- Information Retrieval:
  - To which degree is a Web site, a Web page, a text passage, an image region, a video segment, . . . relevant to my information need?
- Matchmaking
  - To which degree does an object match my requirements?
    - if I'm looking for a car and my budget is *about* 20.000 €, to which degree does a car's price of 20.500 € match my budget?

- Semantic annotation
  - To which degree does e.g., an image object represent a dog?
- Information extraction
  - To which degree am I'm sure that e.g., SW is an acronym of "Semantic Web"?
- Ontology alignment (schema mapping)
  - To which degree do two concepts of two ontologies represent the same, or are disjoint, or are overlapping?
- Representation of background knowledge
  - To some degree birds fly.
  - To some degree Jim is a blond and young.

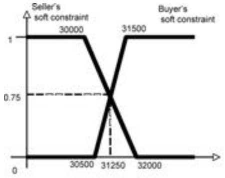
# Example (Distributed Information Retrieval) [7]




Then the agent has to perform automatically the following steps:

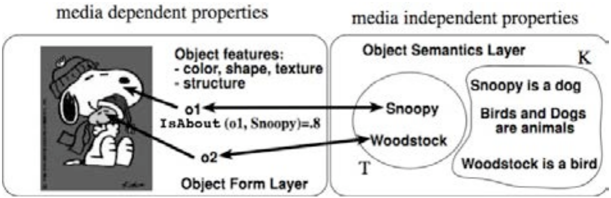
1. The agent has to select a subset of relevant resources $\mathscr{S}' \subseteq \mathscr{S}$, as it is not reasonable to assume to access to and query all resources (resource selection/resource discovery);
2. For every selected source $\mathcal{S}_i \in \mathscr{S}'$ the agent has to reformulate its information need $Q_A$ into the query language $\mathcal{L}_i$ provided by the resource (schema mapping/ontology alignment);
3. The results from the selected resources have to be merged together (data fusion/rank aggregation)

# Example (Negotiation) [2]




- A car seller sells an Audi TT for 31500 €, as from the catalog price.
- A buyer is looking for a sports-car, but wants to to pay not more than around 30000 €
- Classical DLs: the problem relies on the crisp conditions on price.
- More fine grained approach: to consider prices as vague constraints (fuzzy sets) (as usual in negotiation)
  - Seller would sell above 31500 €, but can go down to 30500 €
  - The buyer prefers to spend less than 30000 €, but can go up to 32000 €
  - Highest degree of matching is 0.75 . The car may be sold at 31250 €.

# Example (Logic-based information retrieval model)[1, 8]




| *IsAbout* | | |
|---|---|---|
| *ImageRegion* | *Object ID* | *degree* |
| *o*1 | *snoopy* | 0.8 |
| *o*2 | *woodstock* | 0.7 |
| ⋮ | ⋮ | |

"Find top-*k* image regions about animals"

$$Query(x) \leftarrow ImageRegion(x) \land isAbout(x, y) \land Animal(y)$$

# Example (Database query) [3, 4, 5, 6]

| *HotelID* | *hasLoc* |
|---|---|
| *h*1 | *hl*1 |
| *h*2 | *hl*2 |
| ⋮ | ⋮ |

| *ConferenceID* | *hasLoc* |
|---|---|
| *c*1 | *cl*1 |
| *c*2 | *cl*2 |
| ⋮ | ⋮ |

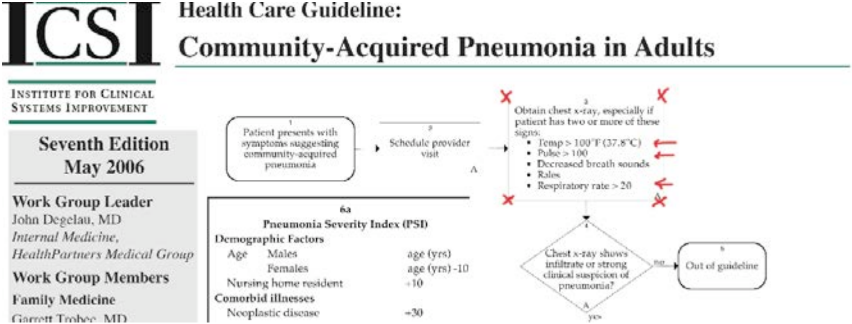
| *hasLoc* | *hasLoc* | *distance* |
|---|---|---|
| *hl*1 | *cl*1 | 300 |
| *hl*1 | *cl*2 | 500 |
| *hl*2 | *cl*1 | 750 |
| *hl*2 | *cl*2 | 800 |
| ⋮ | ⋮ | |

| *hasLoc* | *hasLoc* | *close* | *cheap* |
|---|---|---|---|
| *hl*1 | *cl*1 | 0.7 | 0.3 |
| *hl*1 | *cl*2 | 0.5 | 0.5 |
| *hl*2 | *cl*1 | 0.25 | 0.8 |
| *hl*2 | *cl*2 | 0.2 | 0.9 |
| ⋮ | ⋮ | ⋮ | |

“Find top-*k* cheapest hotels close to the train station”

$$q(h) \leftarrow hasLocation(h, hl) \wedge hasLocation(\mathtt{train}, cl) \wedge close(hl, cl) \wedge cheap(h)$$

# Example (Health-care: diagnosis of pneumonia)




- E.g., *Temp* = 37.5, *Pulse* = 98, *RespiratoryRate* = 18 are in the "danger zone" already
- Temperature, Pulse and Respiratory rate, . . . : these constraints are rather imprecise than crisp

# Uncertainty vs. Vagueness: a clarification

- What does the degree mean?
- There is often a misunderstanding between interpreting a degree as a measure of uncertainty or as a measure of vagueness
- The value 0.83 has a different interpretation in “Birds fly to degree 0.83” from that in “Hotel Verdi is close to the train station to degree 0.83”

# Uncertainty

- Uncertainty: statements are true or false. But, due to lack of knowledge we can only estimate to which probability/possibility/necessity degree they are true or false
  - For instance, a bird flies or does not fly. The probability/possibility/necessity degree that it flies is 0.83
- Usually we have a possible world semantics with a distribution over possible worlds:

$$W = \{I \text{ classical interpretation}\}, \quad I(\varphi) \in \{0,1\}$$

$$\mu\colon W \to [0,1], \quad \mu(I) \in [0,1]$$

$$Pr(\phi) = \sum_{I \models \phi} \mu(I)$$

$$Poss(\phi) = \sup_{I \models \phi} \mu(I)$$

$$Necc(\phi) = \inf_{I \not\models \phi} \mu(I) = 1 - Poss(\neg\phi)$$

# Vagueness

- Vagueness: statements involve concepts for which there is no exact definition, such as tall, small, close, far, cheap, expensive, isAbout, similarTo. Statements are true to some degree which is taken from a truth space.
    - E.g., "Hotel Verdi is close to the train station to degree 0.83"
- Truth space: set of truth values $L$ and an partial order $\leq$
- Many-valued Interpretation: a function $I$ mapping formulae into $L$, i.e. $I(\varphi) \in L$
- Fuzzy Logic: $L = [0, 1]$
- Uncertainty and Vagueness: "It is possible/probable to degree 0.83 that it will be hot tomorrow"
- The notion of imperfect information covers concepts such as uncertainty, vagueness, contradiction, incompleteness, imprecision.

C. Meghini, F. Sebastiani, and U. Straccia.
A model of multimedia information retrieval.
*Journal of the ACM*, 48(5):909–970, 2001.

A. Ragone, U. Straccia, T. D. Noia, E. D. Sciascio, and F. M. Donini.
Vague knowledge bases for matchmaking in p2p e-marketplaces.
In *4th European Semantic Web Conference (ESWC-07)*, number 4519 in Lecture Notes in Computer Science, pages 414—428. Springer Verlag, 2007.

U. Straccia.
Answering vague queries in fuzzy dl-lite.
In *Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, (IPMU-06)*, pages 2238–2245. E.D.K., Paris, 2006.

U. Straccia.
Towards top-k query answering in deductive databases.
In *Proceedings of the 2006 IEEE International Conference on Systems, Man and Cybernetics (SMC-06)*, pages 4873–4879. IEEE, 2006.

U. Straccia.
Towards top-k query answering in description logics: the case of DL-Lite.
In *Proceedings of the 10th European Conference on Logics in Artificial Intelligence (JELIA-06)*, number 4160 in Lecture Notes in Computer Science, pages 439–451, Liverpool, UK, 2006. Springer Verlag.

U. Straccia.
Towards vague query answering in logic programming for logic-based information retrieval.
In *World Congress of the International Fuzzy Systems Association (IFSA-07)*, Cancun, Mexico, 2007.

U. Straccia and R. Troncy.

Towards distributed information retrieval in the semantic web.
In *3rd European Semantic Web Conference (ESWC-06)*, number 4011 in Lecture Notes in Computer Science, pages 378–392. Springer Verlag, 2006.

U. Straccia and G. Visco.
DLMedia: an ontology mediated multimedia information retrieval system.
In *Proceedings of the International Workshop on Description Logics (DL-07)*, Innsbruck, Austria, 2007. CEUR.

# Web Ontology Languages

- Wide variety of languages for “Explicit Specification”
  - Graphical notations
    - Semantic networks
    - UML
    - RDF/RDFS
  - Logic based
    - Description Logics (e.g., OIL, DAML+OIL, OWL, OWL-DL, OWL-Lite)
    - Rules (e.g., RuleML, RIF, SWRL, LP/Prolog)
    - First Order Logic (e.g., KIF)
- RDF and OWL-DL are the major players (so far ...)

# RDF

- Statements are of the form

$$\langle \textit{subject}, \textit{predicate}, \textit{object} \rangle$$

  called triples: e.g.

$$\langle \textit{umberto}, \textit{plays}, \textit{soccer} \rangle$$

- can be represented graphically as:

$$\boxed{\text{umberto}} \xrightarrow{\textit{plays}} \boxed{\text{soccer}}$$

- Statements describe properties of resources
- A resource is any object that can be pointed to by a URI (Universal Resource Identifier):

# RDF Schema (RDFS)

- RDF Schema allows you to define vocabulary terms and the relations between those terms
- RDF Schema terms (just a few examples):
  - Class
  - Property
  - type
  - subClassOf
  - range
  - domain
- These terms are the RDF Schema building blocks (constructors) used to create vocabularies:

```
<Person,type, Class>
<hasColleague, type, Property>
<Professor, subClassOf,Person>
<Carole, type,Professor>
<hasColleague, range,Person>
<hasColleague, domain,Person>
```

# OWL [10]

- Three species of OWL
  - OWL full is union of OWL syntax and RDF (Undecidable)
  - OWL DL restricted to FOL fragment (decidable in NEXPTIME)
  - OWL Lite is “easier to implement” subset of OWL DL (decidable in EXPTIME)
- Semantic layering
  - OWL DL within Description Logic (DL) fragment
- OWL DL based on $\mathcal{SHOIN}(D_n)$ DL
- OWL Lite based on $\mathcal{SHIF}(D_n)$ DL

# Description Logics (DLs)

- The logics behind OWL-DL and OWL-Lite, `http://dl.kr.org/`.
- Concept/Class: names are equivalent to unary predicates
  - In general, concepts equiv to formulae with one free variable
- Role or attribute: names are equivalent to binary predicates
  - In general, roles equiv to formulae with two free variables
- Taxonomy: Concept and role hierarchies can be expressed
- Individual: names are equivalent to constants
- Operators: restricted so that:
  - Language is decidable and, if possible, of low complexity
  - No need for explicit use of variables
    - Restricted form of $\exists$ and $\forall$
  - Features such as counting can be succinctly expressed

# The DL Family

- A given DL is defined by set of concept and role forming operators
- Basic language: $\mathcal{ALC}$($\mathcal{A}$ttributive $\mathcal{L}$anguage with $\mathcal{C}$omplement)

| Syntax | | | | Semantics | Example |
|---|---|---|---|---|---|
| $C, D$ | $\rightarrow$ | $\top$ | \| | $\top(x)$ | |
| | | $\bot$ | \| | $\bot(x)$ | |
| | | $A$ | \| | $A(x)$ | *Human* |
| | | $C \sqcap D$ | \| | $C(x) \wedge D(x)$ | $Human \sqcap Male$ |
| | | $C \sqcup D$ | \| | $C(x) \vee D(x)$ | $Nice \sqcup Rich$ |
| | | $\neg C$ | \| | $\neg C(x)$ | $\neg Meat$ |
| | | $\exists R.C$ | \| | $\exists y.R(x,y) \wedge C(y)$ | $\exists has\_child.Blond$ |
| | | $\forall R.C$ | | $\forall y.R(x,y) \Rightarrow C(y)$ | $\forall has\_child.Human$ |
| $C \sqsubseteq D$ | | | | $\forall x.C(x) \Rightarrow D(x)$ | $Happy\_Father \sqsubseteq Man \sqcap \exists has\_child.Female$ |
| $a{:}C$ | | | | $C(a)$ | $John{:}Happy\_Father$ |

# Toy Example

$$
\begin{array}{rcl}
Sex & = & Male \sqcup Female \\
Male \sqcap Female & \sqsubseteq & \bot \\
Person & \sqsubseteq & Human \sqcap \exists hasSex.Sex \\
MalePerson & \sqsubseteq & Person \sqcap \exists hasSex.Male
\end{array}
$$

$$umberto{:}Person \sqcap \exists hasSex.\neg Female$$

$$KB \models umberto{:}MalePerson$$

# Note on DL Naming

- $\mathcal{AL}$: $C, D \longrightarrow \top \mid \bot \mid A \mid C \sqcap D \mid \neg A \mid \exists R.\top \mid \forall R.C$
- $\mathcal{C}$: Concept negation, $\neg C$. Thus, $\mathcal{ALC} = \mathcal{AL} + \mathcal{C}$
- $\mathcal{S}$: Used for $\mathcal{ALC}$ with transitive roles $\mathcal{R}_+$
- $\mathcal{U}$: Concept disjunction, $C_1 \sqcup C_2$
- $\mathcal{E}$: Existential quantification, $\exists R.C$
- $\mathcal{H}$: Role inclusion axioms, $R_1 \sqsubseteq R_2$, e.g. $is\_component\_of \sqsubseteq is\_part\_of$
- $\mathcal{N}$: Number restrictions, $(\geq n\ R)$ and $(\leq n\ R)$, e.g. $(\geq 3\ has\_Child)$ (has at least 3 children)
- $\mathcal{Q}$: Qualified number restrictions, $(\geq n\ R.C)$ and $(\leq n\ R.C)$, e.g. $(\leq 2\ has\_Child.Adult)$ (has at most 2 adult children)
- $\mathcal{O}$: Nominals (singleton class), $\{a\}$, e.g. $\exists has\_child.\{mary\}$.
  **Note**: $a{:}C$ equiv to $\{a\} \sqsubseteq C$ and $(a, b){:}R$ equiv to $\{a\} \sqsubseteq \exists R.\{b\}$
- $\mathcal{I}$: Inverse role, $R^-$, e.g. $isPartOf = hasPart^-$
- $\mathcal{F}$: Functional role, $f$, e.g. $functional(hasAge)$
- $\mathcal{R}_+$: transitive role, e.g. $transitive(isPartOf)$

For instance,

$$\mathcal{SHIF} = \mathcal{S}+\mathcal{H}+\mathcal{I}+\mathcal{F} = \mathcal{ALCR_+HIF}$$ OWL-Lite (EXPTIME)

$$\mathcal{SHOIN} = \mathcal{S}+\mathcal{H}+\mathcal{O}+\mathcal{I}+\mathcal{N} = \mathcal{ALCR_+HOIN}$$ OWL-DL (NEXPTIME)

# Concrete Domains

- Concrete domains: reals, integers, strings, ...

  $(tim, 14)$:*hasAge*
  $(sf,$ "*SoftComputing*"$)$:*hasAcronym*
  $(source1,$ "*ComputerScience*"$)$:*isAbout*
  $(service2,$ "*InformationRetrievalTool*"$)$:*Matches*
  $Minor = Person \sqcap \exists hasAge. \leq_{18}$

- Semantics: a clean separation between "object" classes and concrete domains
  - $D = \langle \Delta_D, \Phi_D \rangle$
  - $\Delta_D$ is an interpretation domain
  - $\Phi_D$ is the set of concrete domain predicates $d$ with a predefined arity $n$ and fixed interpretation $d^D \subseteq \Delta_D^n$
  - Concrete properties: $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta_D$
- Notation: $(D)$. E.g., $\mathcal{ALC}(D)$ is $\mathcal{ALC}$ + concrete domains

# LPs Basics (for ease, without default negation) [6]

- Predicates are *n*-ary
- Terms are variables or constants
- Rules are of the form

$$P(\mathbf{x}) \leftarrow \varphi(\mathbf{x}, \mathbf{y})$$

where $\varphi(\mathbf{x}, \mathbf{y})$ is a formula built from atoms of the form $B(\mathbf{z})$ and connectors $\wedge, \vee$
For instance,

$$has\_father(x, y) \quad \leftarrow \quad has\_parent(x, y) \wedge Male(y)$$

- Facts are rules with empty body
For instance,

$$has\_parent(mary, jo)$$

# Toy Example

$$
\begin{array}{rcl}
Q(x) & \leftarrow & B(x) \\
Q(x) & \leftarrow & C(x) \\
B(a) & \leftarrow & \\
C(b) & \leftarrow &
\end{array}
$$

$$KB \models Q(a) \quad KB \models Q(b) \qquad answers(KB, Q) = \{a, b\}$$

where $answers(KB, Q) = \{\mathbf{c} \mid KB \models Q(\mathbf{c})\}$

# DLPs Basics

- Combine DLs with LPs:
  - DL atoms and roles may appear in rules

$$\begin{array}{lcl} buy(x) & \leftarrow & Electronics(x), offer(x) \\ Camera & \sqsubseteq & Electronics \end{array}$$

- Knowledge Base is a pair $KB = \langle \mathcal{P}, \Sigma \rangle$, where
  - $\mathcal{P}$ is a logic program
  - $\Sigma$ is a DL knowledge base (set of assertions and inclusion axioms)
- Many different approaches exists with different semantics

C. V. Damásio, J. Z. Pan, G. Stoilos, and U. Straccia.
An approach to representing uncertainty rules in ruleml.
In *Second International Conference on Rules and Rule Markup Languages for the Semantic Web (RuleML-06)*, pages 97–106. IEEE, 2006.

C. V. Damasio, J. Z. Pan, G. Stoilos, and U. Straccia.
Representing uncertainty rules in ruleml.
*Fundamenta Informaticae*, 2007.

T. Eiter, T. Lukasiewicz, R. Schindlauer, and H. Tompits.
Combining answer set programming with description logics for the Semantic Web.
In *Proceedings KR-2004*, pages 141–151. AAAI Press, 2004.

T. Eiter, T. Lukasiewicz, R. Schindlauer, and H. Tompits.
Well-founded semantics for description logic programs in the Semantic Web.
In *Proceedings RuleML-2004*, volume 3323 of *LNCS*, pages 81–97. Springer, 2004.

T. Eiter, G. Ianni, R. Schindlauer, and H. Tompits.
Effective integration of declarative rules with external evaluations for Semantic Web reasoning.
In *Proceedings ESWC-2006*, volume 4011 of *LNCS*, pages 273–287. Springer, 2006.

J. W. Lloyd.
*Foundations of Logic Programming*.
Springer, Heidelberg, RG, 1987.

T. Lukasiewicz.
A novel combination of answer set programming with description logics for the Semantic Web.
In *Proceedings ESWC-2007*, volume 4519 of *LNCS*, pages 384–398. Springer, 2007.

A. Ragone, U. Straccia, T. D. Noia, E. D. Sciascio, and F. M. Donini.
Vague knowledge bases for matchmaking in p2p e-marketplaces.
In *4th European Semantic Web Conference (ESWC-07)*, number 4519 in Lecture Notes in Computer Science, pages 414—428. Springer Verlag, 2007.

U. Straccia.
Towards top-k query answering in deductive databases.
In *Proceedings of the 2006 IEEE International Conference on Systems, Man and Cybernetics (SMC-06)*, pages 4873–4879. IEEE, 2006.

W3C.
OWL web ontology language overview, 2004.
W3C Recommendation (10 Feb. 2004). Available at
`http://www.w3.org/TR/2004/REC-owl-features-20040210/`.

# Probabilistic Logic

- Integration of (propositional) logic- and probability-based representation and reasoning formalisms.
- Reasoning from logical constraints and interval restrictions for conditional probabilities (also called *conditional constraints*).
- Reasoning from convex sets of probability distributions.
- Model-theoretic notion of logical entailment.

# Syntax of Probabilistic Knowledge Bases

- Finite nonempty set of basic events $\Phi = \{p_1, \ldots, p_n\}$.
- Event $\phi$: Boolean combination of basic events
- Logical constraint $\psi \Leftarrow \phi$: events $\psi$ and $\phi$: "$\phi$ implies $\psi$".
- Conditional constraint $(\psi|\phi)[l, u]$: events $\psi$ and $\phi$, and $l, u \in [0, 1]$: "conditional probability of $\psi$ given $\phi$ is in $[l, u]$".
- Probabilistic knowledge base $KB = (L, P)$:
  - finite set of logical constraints $L$,
  - finite set of conditional constraints $P$.

# Example

Probabilistic knowledge base $KB = (L, P)$:

- $L = \{bird \Leftarrow eagle\}$:

  "All eagles are birds".

- $P = \{(have\_legs \mid bird)[1, 1], (fly \mid bird)[0.95, 1]\}$:

  "All birds have legs".

  "Birds fly with a probability of at least 0.95".

# Semantics of Probabilistic Knowledge Bases

- World $I$: truth assignment to all basic events in $\Phi$.
- $\mathcal{I}_\Phi$: all worlds for $\Phi$.
- Probabilistic interpretation $Pr$: probability function on $\mathcal{I}_\Phi$.
- $Pr(\phi)$: sum of all $Pr(I)$ such that $I \in \mathcal{I}_\Phi$ and $I \models \phi$.
- $Pr(\psi|\phi)$: if $Pr(\phi) > 0$, then $Pr(\psi|\phi) = Pr(\psi \wedge \phi) \,/\, Pr(\phi)$.
- Truth under $Pr$:
  - $Pr \models \psi \Leftarrow \phi$ iff $Pr(\psi \wedge \phi) = Pr(\phi)$ (iff $Pr(\psi \Leftarrow \phi) = 1$).
  - $Pr \models (\psi|\phi)[l, u]$ iff $Pr(\psi \wedge \phi) \in [l, u] \cdot Pr(\phi)$ (iff either $Pr(\phi) = 0$ or $Pr(\psi|\phi) \in [l, u]$).

# Example

- Set of basic propositions $\Phi = \{bird, fly\}$.
- $\mathcal{I}_\Phi$ contains exactly the worlds $I_1$, $I_2$, $I_3$, and $I_4$ over $\Phi$:

| | *fly* | $\neg$*fly* |
|---|---|---|
| *bird* | $I_1$ | $I_2$ |
| $\neg$*bird* | $I_3$ | $I_4$ |

- Some probabilistic interpretations:

| $Pr_1$ | *fly* | $\neg$*fly* |
|---|---|---|
| *bird* | 19/40 | 1/40 |
| $\neg$*bird* | 10/40 | 10/40 |

| $Pr_2$ | *fly* | $\neg$*fly* |
|---|---|---|
| *bird* | 0 | 1/3 |
| $\neg$*bird* | 1/3 | 1/3 |

- $Pr_1(fly \wedge bird) = 19/40$ and $Pr_1(bird) = 20/40$.
- $Pr_2(fly \wedge bird) = 0$ and $Pr_2(bird) = 1/3$.
- $\neg fly \Leftarrow bird$ is false in $Pr_1$, but true in $Pr_2$.
- $(fly \mid bird)[.95, 1]$ is true in $Pr_1$, but false in $Pr_2$.

# Satisfiability and Logical Entailment

- $Pr$ is a model of $KB = (L, P)$ iff $Pr \models F$ for all $F \in L \cup P$.
- $KB$ is satisfiable iff a model of $KB$ exists.
- $KB \,\|\!\!= (\psi|\phi)[l, u]$: $(\psi|\phi)[l, u]$ is a logical consequence of $KB$ iff every model of $KB$ is also a model of $(\psi|\phi)[l, u]$.
- $KB \,\|\!\!=_{tight} (\psi|\phi)[l, u]$: $(\psi|\phi)[l, u]$ is a tight logical consequence of $KB$ iff $l$ (resp., $u$) is the infimum (resp., supremum) of $Pr(\psi|\phi)$ subject to all models $Pr$ of $KB$ with $Pr(\phi) > 0$.

# Example

- Probabilistic knowledge base:

$$KB = (\{bird \Leftarrow eagle\}, \{(have\_legs \mid bird)[1,1], (fly \mid bird)[0.95,1]\}).$$

- $KB$ is satisfiable, since

  $Pr$ with $Pr(bird \wedge eagle \wedge have\_legs \wedge fly) = 1$ is a model.

- Some conclusions under logical entailment:

  $KB \Vdash (have\_legs \mid bird)[0.3,1], \quad KB \Vdash (fly \mid bird)[0.6,1].$

- Tight conclusions under logical entailment:

  $KB \Vdash_{tight} (have\_legs \mid bird)[1,1], \quad KB \Vdash_{tight} (fly \mid bird)[0.95,1],$
  $KB \Vdash_{tight} (have\_legs \mid eagle)[1,1], \quad KB \Vdash_{tight} (fly \mid eagle)[0,1].$

# Literature


- G. Boole. *An Investigation of the Laws of Thought, on which are Founded the Mathematical Theories of Logic and Probabilities*. Walton and Maberley, London, 1854.
- N. J. Nilsson. Probabilistic logic. *Artif. Intell.*, 28:71–88, 1986.
- D. Dubois, H. Prade, and J.-M. Touscas. Inference with imprecise numerical quantifiers. In *Intelligent Systems*, 1990.
- R. Fagin, J. Y. Halpern, and N. Megiddo. A logic for reasoning about probabilities. *Inf. Comput.*, 87:78–128, 1990.
- A. M. Frisch and P. Haddawy. Anytime deduction for probabilistic logic. *Artif. Intell.*, 69:93–122, 1994.
- T. Lukasiewicz. Probabilistic deduction with conditional constraints over basic events. *JAIR*, 10:199–241, 1999.
- T. Lukasiewicz. Probabilistic logic programming with conditional constraints. *ACM TOCL* 2(3):289–339, 2001.

# Probabilistic Ontologies

Main types of encoded probabilistic knowledge:

- Terminological probabilistic knowledge about concepts and roles: “Birds fly with a probability of at least 0.95”.
- Assertional probabilistic knowledge about instances of concepts and roles: “Tweety is a bird with a probability of at least 0.9”.

Main types of reasoning problems:

- Satisfiability of the terminological probabilistic knowledge.
- Tight conclusions about generic objects (from the terminological probabilistic knowledge).
- Satisfiability of the assertional probabilistic knowledge.
- Tight conclusions about concrete objects (from both the terminological and the assertional probabilistic knowledge).

# Use of Probabilistic Ontologies

- Representation of terminological and assertional probabilistic knowledge (e.g., in the medical domain or at the stock exchange market).
- Information retrieval, for an increased recall (e.g., Udrea et al.: Probabilistic ontologies and relational databases. In *Proc. CoopIS/DOA/ODBASE-2005*).
- Ontology matching (e.g., Mitra et al.: OMEN: A probabilistic ontology mapping tool. In *Proc. ISWC-2005*).
- Probabilistic data integration, especially for handling ambiguous and controversial pieces of information.

# Probabilistic RDF

O. Udrea, V. S. Subrahmanian, and Z. Majkic. Probabilistic RDF. In *Proceedings IRI-2006*.

- probabilistic generalization of RDF
- terminological probabilistic knowledge about classes
- assertional probabilistic knowledge about properties of individuals
- assertional probabilistic inference for acyclic probabilistic RDF theories, which is based on logical entailment in probabilistic logic, coupled with a local probabilistic semantics

# Probabilistic DLs

R. Giugno, T. Lukasiewicz. P-$\mathcal{SHOQ}(\mathbf{D})$: A probabilistic extension of $\mathcal{SHOQ}(\mathbf{D})$ for probabilistic ontologies in the SW. In *Proc. JELIA-2002*.

- probabilistic generalization of the description logic $\mathcal{SHOQ}(\mathbf{D})$ (recently also extended to $\mathcal{SHIF}(\mathbf{D})$ and $\mathcal{SHOIN}(\mathbf{D})$)
- terminological probabilistic knowledge about concepts and roles
- assertional probabilistic knowledge about instances of concepts and roles
- terminological probabilistic inference based on lexicographic entailment in probabilistic logic (stronger than logical entailment)
- assertional probabilistic inference based on lexicographic entailment in probabilistic logic (for combining assertional and terminological probabilistic knowledge)
- terminological and assertional probabilistic inference problems reduced to sequences of linear optimization problems

- M. Jaeger. Probabilistic reasoning in terminological logics. In *Proceedings KR-1994*.
- D. Koller, A. Levy, and A. Pfeffer. P-CLASSIC: A tractable probabilistic description logic. In *Proceedings AAAI-1997*.
- P. C. G. da Costa. Bayesian Semantics for the Semantic Web. PhD thesis, George Mason University, Fairfax, VA, USA, 2005.
- P. C. G. da Costa and K. B. Laskey. PR-OWL: A framework for probabilistic ontologies. In *Proceedings FOIS-2006*.

# Possibilistic DLs

Generalization of DLs by possibilistic uncertainty, which is based on possibilistic interpretations rather than probabilistic interpretations.

Possibilistic interpretation: mapping $\pi : \mathcal{I}_\Phi \rightarrow [0, 1]$.

"$\pi(I)$ is the degree to which the world $I$ is possible."

*Poss*($\phi$): possibility of $\phi$ in $\pi$: $Poss(\phi) = \max\{\pi(I) \mid I \in \mathcal{I}_\Phi, I \models \phi\}$

- B. Hollunder. An alternative proof method for possibilistic logic and its application to terminological logics. *Int. J. Approx. Reasoning*, 12(2):85–109, 1995.
- D. Dubois, J. Mengin, and H. Prade. Possibilistic uncertainty and fuzzy features in description logic: A preliminary discussion. In E. Sanchez, editor, *Capturing Intelligence: Fuzzy Logic and the Semantic Web*, 2006.
- C.-J. Liau and Y. Y. Yao. Information retrieval by possibilistic reasoning. In *Proc. DEXA-2001*.

# Other Works


- Z. Ding and Y. Peng. A probabilistic extension to ontology language OWL. In *Proceedings HICSS-2004*.
- Y. Yang and J. Calmet. OntoBayes: An ontology-driven uncertainty model. In *Proceedings IAWTIC-2005*.
- Z. Ding, Y. Peng, and R. Pan. BayesOWL: Uncertainty modeling in Semantic Web ontologies. In Z. Ma, editor, *Soft Computing in Ontologies and Semantic Web*. Springer, 2006.
- H. Nottelmann and N. Fuhr. Adding probabilities and rules to OWL Lite subsets based on probabilistic Datalog. *IJUFKS*, 14(1):17–42, 2006.

# Probabilistic Logic Programs

Probabilistic generalizations of logic programs / rule-based systems / deductive databases / Datalog:

(1) Probabilistic generalizations of (annotated) logic programs based on probabilistic logic (no uncertainty degrees associated with rules):


- R. T. Ng and V. S. Subrahmanian. Probabilistic logic programming. *Inf. Comput.*, 101(2):150–201, 1992.
- R. T. Ng and V. S. Subrahmanian. A semantical framework for supporting subjective and conditional probabilities in deductive databases. *J. Autom. Reasoning*, 10(2):191–235, 1993.
- A. Dekhtyar and V. S. Subrahmanian. Hybrid probabilistic programs. *J. Log. Program.* 43(3):187–250, 2000.

(2) Probabilistic generalizations of logic programs based on Bayesian networks / causal models:


- D. Poole. Probabilistic Horn abduction and Bayesian networks. *Artif. Intell.*, 64:81–129, 1993.
- D. Poole. The independent choice logic for modeling multiple agents under uncertainty. *Artif. Intell.*, 94:7–56, 1997.
- K. Kersting and L. De Raedt. Bayesian logic programs. *CoRR*, cs.AI/0111058, 2001.
- C. Baral, M. Gelfond, and J. N. Rushton. Probabilistic reasoning with answer sets. In *Proceedings LPNMR-2004*.

(3) Relational Bayesian networks:


- M. Jaeger. Relational Bayesian networks. In *Proc. UAI-1997*.
- D. Koller and A. Pfeffer. Object-oriented Bayesian networks. In *Proceedings UAI-1997*.
- H. Pasula and S. J. Russell. Approximate inference for first-order probabilisti languages. In *Proceedings IJCAI-2001*.
- D. Poole. First-order probabilistic inference. In *Proc. IJCAI-2003*.

## (4) First-order generalization of probabilistic knowledge bases in probabilistic logic (based on logical entailment, lexicographic entailment, and maximum entropy entailment):


- T. Lukasiewicz. Probabilistic logic programming. In *Proceedings ECAI-1998*.
- T. Lukasiewicz. Probabilistic logic programming with conditional constraints. *ACM TOCL* 2(3):289–339, 2001.
- T. Lukasiewicz. Probabilistic logic programming under inheritance with overriding. In *Proceedings UAI-2001*.
- G. Kern-Isberner and T. Lukasiewicz. Combining probabilistic logic programming with the power of maximum entropy. *Artif. Intell.*, 157(1–2):139–202, 2004.

# Probabilistic Description Logic Programs

T. Lukasiewicz. Probabilistic description logic programs. *IJAR*, 2007.

- Probabilistic dl-programs generalize (loosely coupled) dl-programs by probabilistic uncertainty as in Poole's ICL.
- They properly generalize Poole's ICL.
- They consist of a dl-program along with a probability distribution $\mu$ over total choices *B*.
- They specify a set of distributions over first-order models: Every total choice *B* along with the dl-program specifies a set of first-order models of which the probabilities should sum up to $\mu(B)$.
- There are also tightly coupled probabilistic dl-programs.
- Important applications are data integration and ontology mapping under probabilistic uncertainty and inconsistency.

# Example

Description logic knowledge base $L$
of a probabilistic dl-program $KB = (L, P, C, \mu)$:

$PC \sqcup Camera \sqsubseteq Electronics$; $PC \sqcap Camera \sqsubseteq \bot$;
$Book \sqcup Electronics \sqsubseteq Product$; $Book \sqcap Electronics \sqsubseteq \bot$;
$Textbook \sqsubseteq Book$;

$Product \sqsubseteq \geq 1\ related$;
$\geq 1\ related \sqcup \geq 1\ related^{-} \sqsubseteq Product$;

$Textbook(tb\_ai)$; $Textbook(tb\_lp)$;
$PC(pc\_ibm)$; $PC(pc\_hp)$;

$related(tb\_ai, tb\_lp)$; $related(pc\_ibm, pc\_hp)$;
$provides(ibm, pc\_ibm)$; $provides(hp, pc\_hp)$.

Classical dl-rules in $P$
of a probabilistic dl-program $KB = (L, P, C, \mu)$:

- $pc(pc\_1)$; $pc(pc\_2)$; $pc(pc\_3)$;
- $brand\_new(pc\_1)$; $brand\_new(pc\_2)$;
- $vendor(dell, pc\_1)$; $vendor(dell, pc\_2)$; $vendor(dell, pc\_3)$;
- $provider(P) \leftarrow vendor(P, X), DL[PC \uplus pc; Product](X)$;
- $provider(P) \leftarrow DL[provides](P, X), DL[PC \uplus pc; Product](X)$;
- $similar(X, Y) \leftarrow DL[related](X, Y)$;
- $similar(X, Z) \leftarrow similar(X, Y), similar(Y, Z)$.

## Probabilistic dl-rules in $P$ along with the probability $\mu$ on the choice space $C$ of a probabilistic dl-program $KB = (L, P, C, \mu)$:

- $avoid(X) \leftarrow DL[Camera](X), \text{not}\ offer(X), avoid\_pos;$
- $offer(X) \leftarrow DL[PC \uplus pc; Electronics](X), \text{not}\ brand\_new(X), offer\_pos;$
- $buy(C, X) \leftarrow needs(C, X), view(X), \text{not}\ avoid(X), v\_buy\_pos;$
- $buy(C, X) \leftarrow needs(C, X), buy(C, Y), also\_buy(Y, X), a\_buy\_pos.$

$\mu$: $avoid\_pos, avoid\_neg \mapsto 0.9\,, 0.1;\quad offer\_pos, offer\_neg \mapsto 0.9\,, 0.1;$
$v\_buy\_pos, v\_buy\_neg \mapsto 0.7\,, 0.3;\quad a\_buy\_pos, a\_buy\_neg \mapsto 0.7\,, 0.3.$

$\{avoid\_pos, offer\_pos, v\_buy\_pos, a\_buy\_pos\} : 0.9 \times 0.9 \times 0.7 \times 0.7, \ldots$

Probabilistic query: $\exists (buy(c, x) \mid needs(c, x) \wedge buy(c, y) \wedge also\_buy(y, x) \wedge view(x) \wedge \neg avoid(x))[L, U]$

# Example: Probabilistic Data Integration

Obtain a weather forecast by integrating the potentially different weather forecasts of three weather forecast institutes $A$, $B$, and $C$.

Our trust in the institutes $A$, $B$, and $C$ is expressed by the trust probabilities 0.6, 0.3, and 0.1, respectively.

Probabilistic integration of the source schemas of $A$, $B$, and $C$ to the global schema $G$ is specified by the following $KB_M = (\emptyset, P_M, C_M, \mu_M)$:

$$\begin{aligned}
P_M = \{ & \mathit{forecast_rome}(D, W, T, M) \leftarrow \mathit{forecast}(\mathit{rome}, D, W, T, M),\ \mathit{inst}_A; \\
& \mathit{forecast_rome}(D, W, T, M) \leftarrow \mathit{forecastRome}(D, W, T, M),\ \mathit{inst}_B; \\
& \mathit{forecast_rome}(D, W, T, M) \leftarrow \mathit{forecast_weather}(\mathit{rome}, D, W), \\
& \quad \mathit{forecast_temperature}(\mathit{rome}, D, T), \\
& \quad \mathit{forecast_wind}(\mathit{rome}, D, M),\ \mathit{inst}_C \}\,;
\end{aligned}$$

$$C_M = \{\{\mathit{inst}_A, \mathit{inst}_B, \mathit{inst}_C\}\}\,;$$

$$\mu_M : \ \mathit{inst}_A, \mathit{inst}_B, \mathit{inst}_C \ \mapsto \ 0.6, 0.3, 0.1\,.$$

# Example (Tightly Coupled): Ontology Mapping

The global schema contains the concept *logic_programming*, while the source schemas contain only the concepts *rule-based_systems* resp. *deductive_databases* in their ontologies.

A randomly chosen book from the area *rule-based_systems* (resp., *deductive_databases*) may belong to *logic_programming* with the probability 0.7 (resp., 0.8).

Probabilistic mapping from the two source schemas to the global schema expressed by the following $KB_M = (\emptyset, P_M, C_M, \mu_M)$:

$$
\begin{aligned}
P_M &= \{\mathit{logic\_programming}(X) \leftarrow \mathit{rule\text{-}based\_systems}(X),\ \mathit{choice}_1\,;\\
&\qquad \mathit{logic\_programming}(X) \leftarrow \mathit{deductive\_databases}(X),\ \mathit{choice}_2\}\,;\\
C_M &= \{\{\mathit{choice}_1, \mathit{not\_choice}_1\}, \{\mathit{choice}_2, \mathit{not\_choice}_2\}\}\,;\\
\mu_M &: \ \mathit{choice}_1, \mathit{not\_choice}_1, \mathit{choice}_2, \mathit{not\_choice}_2 \ \mapsto\ 0.7,\ 0.3,\ 0.8,\ 0.2\,.
\end{aligned}
$$

# Vagueness

- Vagueness: statements involve concepts for which there is no exact definition, such as tall, close, cheap, IsAbout, simialarTo . . .
- Statements are true to some degree which is taken from a truth space
  - E.g., "Hotel Verdi is close to the train station to degree 0.83"
  - "Find top-$k$ cheapest hotels close to the train station"

$$q(h) \leftarrow hasLocation(h, hl) \wedge hasLocation(\mathtt{train}, cl) \wedge close(hl, cl) \wedge cheap(h)$$

- Truth space: usually $[0, 1]$
- Interpretation: a function $I$ mapping atoms into $[0, 1]$, i.e. $I(A) \in [0, 1]$
- Problem: what is the interpretation of e.g. $close(verdi, train) \wedge cheap(200)$?
  - E.g., if $I(close(verdi, train)) = 0.83$ and $I(cheap(200)) = 0.2$, what is the result of $0.83 \wedge 0.2$?
- More generally, what is the result of $n \wedge m$, for $n, m \in [0, 1]$?
- The choice cannot be any arbitrary computable function, but has to reflect some basic properties that one expects to hold for a "conjunction"

# Propositional Fuzzy Logics Basics [5]

- Formulae: propositional formulae
- Truth space is $[0, 1]$
- Formulae have a a degree of truth in $[0, 1]$
- Interpretation: is a mapping $\mathcal{I} : \mathit{Atoms} \to [0, 1]$
- Interpretations are extended to formulae using norms to interpret connectives $\wedge, \vee, \neg, \to$

# Typical norms

| | Lukasiewicz Logic | Gödel Logic | Product Logic | Zadeh |
|---|---|---|---|---|
| $\neg x$ | $1 - x$ | if $x = 0$ then 1 else 0 | if $x = 0$ then 1 else 0 | $1 - x$ |
| $x \wedge y$ | $\max(x + y - 1, 0)$ | $\min(x, y)$ | $x \cdot y$ | $\min(x, y)$ |
| $x \vee y$ | $\min(x + y, 1)$ | $\max(x, y)$ | $x + y - x \cdot y$ | $\max(x, y)$ |
| $x \Rightarrow y$ | if $x \leq y$ then 1 else $1 - x + y$ | if $x \leq y$ then 1 else $y$ | if $x \leq y$ then 1 else $y/x$ | $\max(1 - x, y)$ |

Note: for Lukasiewicz Logic and Zadeh, $x \Rightarrow y \equiv \neg x \vee y$

$$
\begin{array}{rcl}
\mathcal{I}(\phi \wedge \psi) & = & \mathcal{I}(\phi) \wedge \mathcal{I}(\psi) \\
\mathcal{I}(\phi \vee \psi) & = & \mathcal{I}(\phi) \vee \mathcal{I}(\psi) \\
\\
\mathcal{I}(\phi \rightarrow \psi) & = & \mathcal{I}(\phi) \rightarrow \mathcal{I}(\psi) \\
\mathcal{I} \models \phi & \text{iff} & \mathcal{I}(\phi) = 1 \quad \text{iff } \phi \text{ satisfiable} \\
\mathcal{I} \models \mathcal{T} & \text{iff} & \mathcal{I} \models \phi \text{ for all } \phi \in \mathcal{T} \\
\models \phi & \text{iff} & \text{for all } \mathcal{I}\,.\mathcal{I} \models \phi \\
\mathcal{T} \models \phi & \text{iff} & \text{for all } \mathcal{I}.\text{ if } \mathcal{I} \models \mathcal{T} \text{ then } \mathcal{I} \models \phi
\end{array}
$$

- Note:

$$
\begin{array}{rcl}
\neg\phi & \text{is} & \phi \rightarrow 0 \\
\phi \bar{\wedge} \psi & \text{defined as} & \phi \wedge (\phi \rightarrow \psi) \\
\phi \bar{\vee} \psi & \text{defined as} & ((\phi \rightarrow \psi) \rightarrow \psi) \bar{\wedge} ((\psi \rightarrow \phi) \rightarrow \phi) \\
\mathcal{I}(\phi \bar{\wedge} \psi) & = & \min(\mathcal{I}(\phi), \mathcal{I}(\psi)) \\
\mathcal{I}(\phi \bar{\vee} \psi) & = & \max(\mathcal{I}(\phi), \mathcal{I}(\psi))
\end{array}
$$

- Zadeh semantics: not interesting for fuzzy logicians: its a sub-logic of Łukasiewicz and, thus, rarely considered by fuzzy logicians

$$
\begin{array}{rcl}
\neg_Z \phi & = & \neg_Ł \phi \\
\phi \wedge_Z \psi & = & \phi \wedge_Ł (\phi \rightarrow_Ł \psi) \\
\phi \rightarrow_Z \psi & = & \neg_Ł \phi \vee_Ł \psi
\end{array}
$$

Some additional properties of t-norms, s-norms, implication functions, and negation functions of various fuzzy logics.

| Property | Łukasiewicz Logic | Gödel Logic | Product Logic | Zadeh Logic |
|---|---|---|---|---|
| $x \wedge \neg x = 0$ | • | • | • | |
| $x \vee \neg x = 1$ | • | | | |
| $x \wedge x = x$ | | • | | • |
| $x \vee x = x$ | | • | | • |
| $\neg\neg x = x$ | • | | | • |
| $x \rightarrow y = \neg x \vee y$ | • | | | • |
| $\neg(x \rightarrow y) = x \wedge \neg y$ | • | | | • |
| $\neg(x \wedge y) = \neg x \vee \neg y$ | • | • | • | • |
| $\neg(x \vee y) = \neg x \wedge \neg y$ | • | • | • | • |

# Predicate Fuzzy Logics Basics [5]

- Formulae: First-Order Logic formulae, *terms* are either variables or constants
  - we may introduce functions symbols as well, with crisp semantics (but uninteresting), or we need to discuss also fuzzy equality (which we leave out here)
- Truth space is [0, 1]
- Formulae have a a degree of truth in [0, 1]
- Interpretation: is a mapping $\mathcal{I} : Atoms \rightarrow [0,1]$
- Interpretations are extended to formulae as follows:

$$\begin{array}{rcl}\mathcal{I}(\neg\phi) & = & \mathcal{I}(\phi) \rightarrow 0 \\ \mathcal{I}(\phi \wedge \psi) & = & \mathcal{I}(\phi) \wedge \mathcal{I}(\psi) \\ \mathcal{I}(\phi \rightarrow \psi) & = & \mathcal{I}(\phi) \rightarrow \mathcal{I}(\psi) \\ \mathcal{I}(\exists x\phi) & = & \sup_{c \in \Delta^{\mathcal{I}}} \mathcal{I}_x^c(\phi) \\ \mathcal{I}(\forall x\phi) & = & \inf_{c \in \Delta^{\mathcal{I}}} \mathcal{I}_x^c(\phi)\end{array}$$

  where $\mathcal{I}_x^c$ is as $\mathcal{I}$, except that variable $x$ is mapped into individual $c$
- Definitions of $\mathcal{I} \models \phi$, $\mathcal{I} \models \mathcal{T}$, $\models \phi$, $\mathcal{T} \models \phi$, $||\phi||_{\mathcal{T}}$ and $|\phi|_{\mathcal{T}}$ are as for the propositional case

# Fuzzy RDF (we generalize [15, 16, 34])

- Statement (triples) may have attached a degree in $[0, 1]$: for $n \in [0, 1]$

$$\langle (subject, predicate, object), n \rangle$$

- Meaning: the degree of truth of the statement is at least $n$
- For instance,

$$\langle (o1, IsAbout, snoopy), 0.8 \rangle$$

# Inferences in Fuzzy RDFS

Some inferences in fuzzy RDFS (set is not complete). Recall Rational Pavelka Logic ($\rightarrow$ is r-implication)

$$\frac{\langle (a, sp, b), n\rangle, \langle (b, sp, c), m\rangle}{(\langle (a, sp, c), n \wedge m\rangle} \qquad \frac{\langle (a, sp, b), n\rangle, \langle (x, a, y), m\rangle}{\langle (x, b, y), n \wedge m\rangle}$$

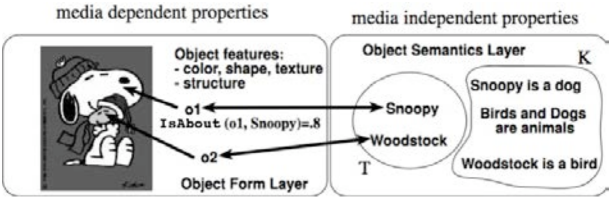
$$\frac{\langle (a, sc, b), n\rangle, \langle (b, sc, c), m\rangle}{\langle (a, sc, c), n \wedge m\rangle} \qquad \frac{\langle (a, sc, b), n\rangle, \langle (x, type, a), m\rangle}{\langle (x, type, b), n \wedge m\rangle}$$

$$\frac{\langle (a, dom, b), n\rangle, \langle (x, a, y), m\rangle}{\langle (x, type, b), n \wedge m\rangle} \qquad \frac{\langle (a, range, b), n\rangle, \langle (x, a, y), m\rangle}{\langle (y, type, b), n \wedge m\rangle}$$

$$\frac{\langle (a, dom, b), n\rangle, \langle (c, sp, a), m\rangle, \langle (x, c, y), k\rangle}{\langle (x, type, b), n \wedge m \wedge k\rangle} \qquad \frac{\langle (a, range, b), n\rangle, \langle (c, sp, a), m\rangle, \langle (x, c, y), k\rangle}{\langle (y, type, b), n \wedge m \wedge k\rangle}$$

sp = “subPropertyOf”, sc = “subClassOf”

# Example




- Fuzzy RDF representation

$$\begin{array}{l} \langle(o1, IsAbout, snoopy), 0.8\rangle \\ \langle(snoopy, type, dog), 1.0\rangle \\ \langle(woodstock, type, bird), 1.0\rangle \\ \langle(dog, subClassOf, Animal), 1.0\rangle \\ \langle(bird, subClassOf, Animal), 1.0\rangle \end{array}$$

- then

$$KB \models \langle \exists x.(o1, IsAbout, x) \wedge (x, type, Animal), 0.8\rangle$$

# Fuzzy DLs Basics [26]

The semantics is an immediate consequence of the First-Order-Logic translation of DLs expressions

Interpretation:

$$\mathcal{I} = \Delta^{\mathcal{I}}$$
$$C^{\mathcal{I}} : \Delta^{\mathcal{I}} \to [0,1]$$
$$R^{\mathcal{I}} : \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \to [0,1]$$

$\wedge$ = t-norm
$\vee$ = s-norm
$\neg$ = negation
$\rightarrow$ = implication

Concepts:

| Syntax | | | Semantics | | |
|---|---|---|---|---|---|
| $C, D$ | $\longrightarrow$ | $\top$ \| | $\top^{\mathcal{I}}(x)$ | $=$ | $1$ |
| | | $\bot$ \| | $\bot^{\mathcal{I}}(x)$ | $=$ | $0$ |
| | | $A$ \| | $A^{\mathcal{I}}(x)$ | $\in$ | $[0,1]$ |
| | | $C \sqcap D$ \| | $(C_1 \sqcap C_2)^{\mathcal{I}}(x)$ | $=$ | $C_1{}^{\mathcal{I}}(x) \wedge C_2{}^{\mathcal{I}}(x)$ |
| | | $C \sqcup D$ \| | $(C_1 \sqcup C_2)^{\mathcal{I}}(x)$ | $=$ | $C_1{}^{\mathcal{I}}(x) \vee C_2{}^{\mathcal{I}}(x)$ |
| | | $\neg C$ \| | $(\neg C)^{\mathcal{I}}(x)$ | $=$ | $\neg C^{\mathcal{I}}(x)$ |
| | | $\exists R.C$ \| | $(\exists R.C)^{\mathcal{I}}(x)$ | $=$ | $\sup_{y \in \Delta^{\mathcal{I}}} R^{\mathcal{I}}(x,y) \wedge C^{\mathcal{I}}(y)$ |
| | | $\forall R.C$ | $(\forall R.C)^{\mathcal{I}}(u)$ | $=$ | $\inf_{y \in \Delta^{\mathcal{I}}} R^{\mathcal{I}}(x,y) \rightarrow C^{\mathcal{I}}(y)\}$ |

Assertions: $\langle a{:}C, r\rangle, \mathcal{I} \models \langle a{:}C, r\rangle$ iff $C^{\mathcal{I}}(a^{\mathcal{I}}) \geq r$ (similarly for roles)

- individual $a$ is instance of concept $C$ at least to degree $r$, $r \in [0,1] \cap \mathbb{Q}$

Inclusion axioms: $C \sqsubseteq D$,

- $\mathcal{I} \models C \sqsubseteq D$ iff $\forall x \in \Delta^{\mathcal{I}}.C^{\mathcal{I}}(x) \leq D^{\mathcal{I}}(x)$
- this is equivalent to, $\forall x \in \Delta^{\mathcal{I}}.(C^{\mathcal{I}}(x) \rightarrow D^{\mathcal{I}}(x)) = 1$, if $\rightarrow$ is an r-implication

# Basic Inference Problems

Consistency: Check if knowledge is meaningful

- Is *KB* consistent, i.e. satisfiable?

Subsumption: structure knowledge, compute taxonomy

- $KB \models C \sqsubseteq D$ ?

Equivalence: check if two fuzzy concepts are the same

- $KB \models C = D$ ?

Graded instantiation: Check if individual *a* instance of class *C* to degree at least *r*

- $KB \models \langle a{:}C, r \rangle$ ?

BTVB: Best Truth Value Bound problem

- $|a{:}C|_{KB} = \sup\{r \mid KB \models \langle a{:}C, r \rangle\}$ ?

Top-k retrieval: Retrieve the top-k individuals that instantiate *C* w.r.t. best truth value bound

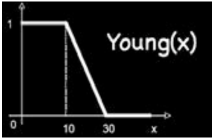
- $ans_{top-k}(KB, C) = Top_k\{\langle a, v \rangle \mid v = |a{:}C)|_{KB}\}$

# Towards fuzzy OWL Lite and OWL DL

- Recall that OWL Lite and OWL DL relate to $\mathcal{SHIF}(D)$ and $\mathcal{SHOIN}(D)$, respectively
- We need to extend the semantics of fuzzy $\mathcal{ALC}$ to fuzzy $\mathcal{SHOIN}(D) = \mathcal{ALCHOINR}_+(D)$
- Additionally, we add
  - modifiers (e.g., *very*)
  - concrete fuzzy concepts (e.g., *Young*)
  - both additions have explicit membership functions

# Concrete fuzzy concepts

- E.g., *Small*, *Young*, *High*, *etc*. with explicit membership function
- Use the idea of concrete domains:
  - $D = \langle \Delta_D, \Phi_D \rangle$
  - $\Delta_D$ is an interpretation domain
  - $\Phi_D$ is the set of concrete fuzzy domain predicates $d$ with a predefined arity $n = 1, 2$ and fixed interpretation $d^D \colon \Delta_D^n \to [0, 1]$
  - For instance,




$$\begin{array}{rcl} Minor & = & Person \sqcap \exists hasAge. \leq_{18} \\ YoungPerson & = & Person \sqcap \exists hasAge. Young \\ & & functional(hasAge) \end{array}$$

# Modifiers

- *Very*, *moreOrLess*, *slightly*, etc.
- Apply to fuzzy sets to change their membership function
  - $very(x) = x^2$
  - $slightly(x) = \sqrt{x}$
- For instance,


$$SportsCar = Car \sqcap \exists speed.very(High)$$

# Fuzzy $\mathcal{SHOIN}(D)$

Concepts:

| | | Syntax | Semantics |
|---|---|---|---|
| $C, D$ | $\longrightarrow$ | $\top$ \| | $\top(x)$ |
| | | $\bot$ \| | $\bot\ (x)$ |
| | | $A$ \| | $A(x)$ |
| | | $(C \sqcap D)$ \| | $C_1(x) \wedge C_2(x)$ |
| | | $(C \sqcup D)$ \| | $C_1(x) \vee C_2(x)$ |
| | | $(\neg C)$ \| | $\neg C(x)$ |
| | | $(\exists R.C)$ \| | $\exists x\ R(x,y) \wedge C(y)$ |
| | | $(\forall R.C)$ \| | $\forall x\ R(x,y) \rightarrow C(y)$ |
| | | $\{a\}$ \| | $x = a$ |
| | | $(\geq n\ R)$ \| | $\exists y_1, \ldots, y_n. \bigwedge_{i=1}^{n} R(x, y_i) \wedge \bigwedge_{1 \leq i < j \leq n} y_i \neq y_j$ |
| | | $(\leq n\ R)$ \| | $\forall y_1, \ldots, y_{n+1}. \bigwedge_{i=1}^{n+1} R(x, y_i) \rightarrow \bigvee_{1 \leq i < j \leq n+1} y_i = y_j$ |
| | | $FCC$ \| | $\mu_{FCC}(x)$ |
| | | $M(C)$ \| | $\mu_M(C(x))$ |
| $R$ | $\longrightarrow$ | $P$ \| | $P(x,y)$ |
| | | $P^-$ \| | $P(y,x)$ |

Assertions:

| | | Syntax | Semantics |
|---|---|---|---|
| $\alpha$ | $\longrightarrow$ | $\langle a{:}C, r\rangle$ \| | $C(a) \geq r$ |
| | | $\langle (a,b){:}R, r\rangle$ | $R(a,b) \geq r$ |

Axioms:

| | | Syntax | Semantics |
|---|---|---|---|
| $\tau$ | $\longrightarrow$ | $\langle C \sqsubseteq D, r\rangle$ \| | $\forall x\ (C(x) \rightarrow D(x)) \geq r,$ |
| | | $fun(R)$ \| | $\forall x \forall y \forall z\ R(x,y) \wedge R(x,z) \rightarrow y = z$ |
| | | $trans(R)$ | $(\exists z\ R(x,z) \wedge R(z,y)) \rightarrow R(x,y)$ |

# Example (Graded Entailment)

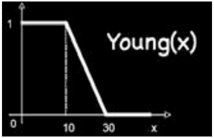
| | | |
|---|---|---|
|  |  |  |
| *audi_tt* | *mg* | *ferrari_enzo* |

| *Car* | *speed* |
|---|---|
| *audi_tt* | 243 |
| *mg* | $\leq$ 170 |
| *ferrari_enzo* | $\geq$ 350 |

$$SportsCar \quad = \quad Car \sqcap \exists hasSpeed.very(High)$$

$$KB \models \langle ferrari\_enzo{:}SportsCar, 1\rangle$$
$$KB \models \langle audi\_tt{:}SportsCar, 0.92\rangle$$
$$KB \models \langle mg{:}\neg SportsCar, 0.72\rangle$$

# Example (Graded Subsumption)

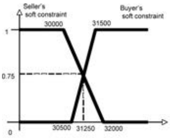



$$\begin{array}{rcl} \mathit{Minor} & = & \mathit{Person} \sqcap \exists \mathit{hasAge}. \leq_{18} \\ \mathit{YoungPerson} & = & \mathit{Person} \sqcap \exists \mathit{hasAge}.\mathit{Young} \end{array}$$

$$\mathit{KB} \models \langle \mathit{Minor} \sqsubseteq \mathit{YoungPerson}, 0.2 \rangle$$

Note: without an explicit membership function of *Young*, this inference cannot be drawn

# Example (Simplified Negotiation)

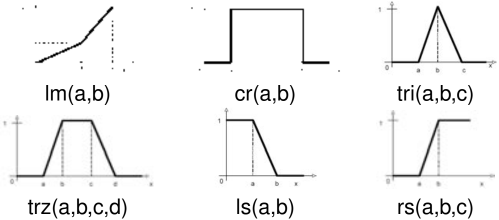



- a car seller sells an Audi TT for 31500 €, as from the catalog price.
- a buyer is looking for a sports-car, but wants to to pay not more than around 30000 €
- classical DLs: the problem relies on the crisp conditions on price
- more fine grained approach: to consider prices as fuzzy sets (as usual in negotiation)
  - seller may consider optimal to sell above 31500 €, but can go down to 30500 €
  - the buyer prefers to spend less than 30000 €, but can go up to 32000 €

$$\begin{array}{lcl} AudiTT & = & SportsCar \sqcap \exists hasPrice.rs(30500, 31500) \\ Query & = & SportsCar \sqcap \exists hasPrice.ls(30000, 32000) \end{array}$$

  - highest degree to which the concept
    $C = AudiTT \sqcap Query$
    is satisfiable is 0.75 (the possibility that the Audi TT and the query matches is 0.75)
  - the car may be sold at 31250 €

- Modifiers are definable as linear in-equations over $\mathbb{Q}, \mathbb{Z}$ (e.g., linear hedges), for instance, linear hedges, $lm(a, b)$, e.g. $very = lm(0.7, 0.49)$
- Fuzzy concrete concepts are definable as linear in-equations over $\mathbb{Q}, \mathbb{Z}$ (e.g., crisp, triangular, trapezoidal, left shoulder and right shoulder membership functions)



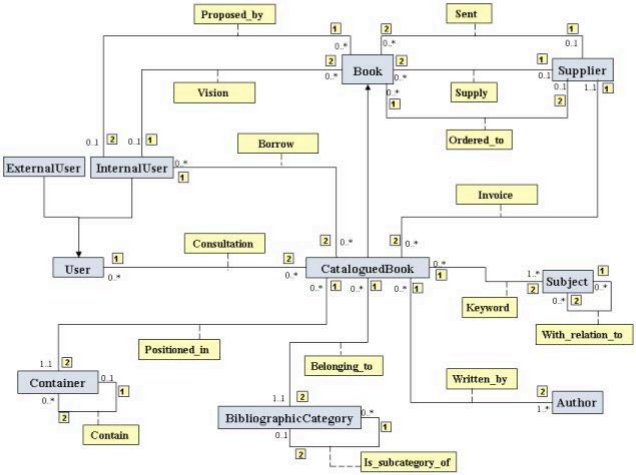
# Implementation issues

- Several options exists:
  - Try to map fuzzy DLs to classical DLs
    - difficult to work with modifiers and concrete fuzzy concepts
  - Try to map fuzzy DLs to some fuzzy logic programming framework
    - A lot of work exists about mappings among classical DLs and LPs
    - But, needs a theorem prover for fuzzy LPs
  - Build an ad-hoc theorem prover for fuzzy DLs, using e.g., MILP
- A theorem prover for fuzzy $\mathcal{SHIF}$ + linear hedges + concrete fuzzy concepts + linear equational constraints + datatypes, under classical, Zadeh, Lukasiewicz and Product t-norm semantics has been implemented (`http://gaia.isti.cnr.it/~straccia`)
- FIRE: a fuzzy DL theorem prover for fuzzy $\mathcal{SHIN}$ under Zadeh semantics (`http://www.image.ece.ntua.gr/~nsimou/`)

# Top-*k* retrieval in tractable DLs: the case of DL-Lite/DLR-Lite [25, 30]

- DL-Lite/DLR-Lite [3]: a simple, but interesting DLs
- Captures important subset of UML/ER diagrams
- Computationally tractable DL to query large databases
- Sub-linear, i.e. LOGSpace in data complexity
  - (same cost as for SQL)
- Good for very large database tables, with limited declarative schema design

- Knowledge base: $KB = \langle \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{T}$ and $\mathcal{A}$ are finite sets of axioms and assertions
- Axiom: $Cl \sqsubseteq Cr$ (inclusion axiom)
- Note for inclusion axioms: the language for left hand side is different from the one for right hand side
- DL-Lite$_{core}$:
  - Concepts:
    $$\begin{array}{rcl} Cl & \rightarrow & A \mid \exists R \\ Cr & \rightarrow & A \mid \exists R \mid \neg A \mid \neg \exists R \\ R & \rightarrow & P \mid P^{-} \end{array}$$
  - Assertion: $a{:}A$, $(a, b){:}P$
- DLR-Lite$_{core}$: ($n$-ary roles)
  - Concepts:
    $$\begin{array}{rcl} Cl & \rightarrow & A \mid \exists P[i] \\ Cr & \rightarrow & A \mid \exists P[i] \mid \neg A \mid \neg \exists P[i] \end{array}$$
  - $\exists P[i]$ is the projection on $i$-th column
  - Assertion: $a{:}A$, $\langle a_1, \ldots, a_n \rangle{:}P$
- Assertions are stored in relational tables
- Conjunctive query: $q(\mathbf{x}) \leftarrow \exists \mathbf{y}.conj(\mathbf{x}, \mathbf{y})$
  $conj$ is an aggregation of expressions of the form $B(z)$ or $P(z_1, z_2)$,


Proposed_by
Sent
Book
Supplier
Vision
Supply
Ordered_to
Borrow
ExternalUser
InternalUser
Invoice
User
Consultation
CataloguedBook
Subject
Keyword
With_relation_to
Positioned_in
Belonging_to
Written_by
Container
Author
BibliographicCategory
Contain
Is_subcategory_of

- Examples:

| | |
|---|---|
| *isa* | $CatalogueBook \sqsubseteq Book$ |
| *disjointness* | $Book \sqsubseteq \neg Author$ |
| *constraints* | $CatalogueBook \sqsubseteq \exists positioned\_In$ |
| *role − typing* | $\exists positioned\_In \sqsubseteq Container$ |
| *functional* | $fun(positioned\_In)$ |
| *constraints* | $Author \sqsubseteq \exists written\_By^-$ |
| | $\exists written\_By \sqsubseteq CatalogueBook$ |
| *assertion* | $Romeo\_and\_Juliet{:}CatalogueBook$ |
| | $(Romeo\_and\_Juliet, Shakespeare){:}written\_By$ |
| *query* | $q(x, y) \leftarrow CataloguedBook(x), Ordered\_to(x, y)$ |

- Consistency check is linear time in the size of the KB
- Query answering in linear in in the size of the number of assertions

# Top-$k$ retrieval in DL-Lite/DLR-Lite

- We extend the query formalism: conjunctive queries, where fuzzy predicates may appear
- conjunctive query

$$q(\mathbf{x}, s) \leftarrow \exists \mathbf{y}.conj(\mathbf{x}, \mathbf{y}), s = f(p_1(\mathbf{z}_1), \ldots, p_n(\mathbf{z}_n))$$

1. $\mathbf{x}$ are the *distinguished variables*;
2. $s$ is the *score variable*, taking values in $[0, 1]$;
3. $\mathbf{y}$ are existentially quantified variables, called *non-distinguished variables*;
4. $conj(\mathbf{x}, \mathbf{y})$ is a conjunction of DL-Lite/DLR-Lite atoms $R(\mathbf{z})$ in $KB$;
5. $\mathbf{z}$ are tuples of constants in $KB$ or variables in $\mathbf{x}$ or $\mathbf{y}$;
6. $\mathbf{z_i}$ are tuples of constants in $KB$ or variables in $\mathbf{x}$ or $\mathbf{y}$;
7. $p_i$ is an $n_i$-ary *fuzzy predicate* assigning to each $n_i$-ary tuple $\mathbf{c}_i$ the *score* $p_i(\mathbf{c}_i) \in [0, 1]$;
8. $f$ is a monotone *scoring* function $f\colon [0, 1]^n \rightarrow [0, 1]$, which combines the scores of the $n$ fuzzy predicates $p_i(\mathbf{c}_i)$

# Example:

$$\begin{array}{rcl} Hotel & \sqsubseteq & \exists HasHLoc \\ Hotel & \sqsubseteq & \exists HasHPrice \\ Conference & \sqsubseteq & \exists HasCLoc \\ Hotel & \sqsubseteq & \neg Conference \end{array}$$

| *HasHLoc* | | *HasCLoc* | | *HasHPrice* | |
|---|---|---|---|---|---|
| *HotelID* | *HasLoc* | *ConfID* | *HasLoc* | *HotelID* | *Price* |
| *h*1 | *hl*1 | *c*1 | *cl*1 | *h*1 | 150 |
| *h*2 | *hl*2 | *c*2 | *cl*2 | *h*2 | 200 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

$$q(h, s) \leftarrow HasHLoc(h, hl), HasHPrice(h, p), Distance(hl, cl, d)$$
$$HasCLoc(c1, cl), s = cheap(p) \cdot close(d) .$$

where the fuzzy predicates *cheap* and *close* are defined as

$$\begin{array}{rcl} close(d) & = & ls(0, 2km, d) \\ cheap(p) & = & ls(0, 300, p) \end{array}$$

## Tool exists and implemented in the DLMedia system

http://gaia.isti.cnr.it/~straccia

# DLMedia: a Multimedia Information Retrieval System [33]

- Based on fuzzy DLR-Lite with similarity predicates
  - Axioms: $Rl_1 \sqcap \ldots \sqcap Rl_m \sqsubseteq Rr$

$$
\begin{array}{lcl}
Rr & \longrightarrow & A \mid \exists[i_1, \ldots, i_k]R \\
Rl & \longrightarrow & A \mid \exists[i_1, \ldots, i_k]R \mid \exists[i_1, \ldots, i_k]R.(Cond_1 \sqcap \ldots \sqcap Cond_l) \\
Cond & \longrightarrow & ([i] \leq v) \mid ([i] < v) \mid ([i] \geq v) \mid ([i] > v) \mid ([i] = v) \mid ([i] \neq v) \mid \\
 & & ([i]\ simTxt\ 'k_1, \ldots, k_n') \mid ([i]\ simImg\ URN)
\end{array}
$$

  - $\exists[i_1, \ldots, i_k]R$ is the projection of the relation $R$ on the columns $i_1, \ldots, i_k$
  - $\exists[i_1, \ldots, i_k]R.(Cond_1 \sqcap \ldots \sqcap Cond_l)$ further restricts the projection $\exists[i_1, \ldots, i_k]R$ according to the conditions specified in $Cond_i$
  - $([i]\ simTxt\ 'k_1 \ldots k_n')$ evaluates the degree of being the text of the $i$-th column similar to the list of keywords $k_1 \ldots k_n$
  - $([i]\ simImg\ URN)$ returns the system's degree of being the image identified by the $i$-th column similar to the image identified by the $URN$
  - Facts: $\langle R(c_1, \ldots, c_n), s \rangle$

- Example axioms

$\exists[1,2]\mathit{Person} \sqsubseteq \exists[1,2]\mathit{hasAge}$
// constrains relation *hasAge*(*name*, *age*)
$\exists[3,1]\mathit{Person} \sqsubseteq \exists[1,2]\mathit{hasChild}$
// constrains relation *hasChild*(*father_name*, *name*)
$\exists[4,1]\mathit{Person} \sqsubseteq \exists[1,2]\mathit{hasChild}$
// constrains relation *hasChild*(*mother_name*, *name*)
$\exists[3,1]\mathit{Person}.(([2] \geq 18) \sqcap ([5] =' \mathit{female}')) \sqsubseteq \exists[1,2]\mathit{hasAdultDaughter}$
// constrains relation *hasAdultDaughter*(*father_name*, *name*)

- On the other hand examples axioms involving similarity predicates are,

$$\exists[1]\mathit{ImageDescr}.([2]\ \mathit{simImg}\ \mathit{urn}1) \sqsubseteq \mathit{Child} \quad (1)$$

$$\exists[1]\mathit{Title}.([2]\ \mathit{simTxt}\ '\mathit{lion}') \sqsubseteq \mathit{Lion} \quad (2)$$

where *urn*1 identifies the image

- Example queries

$q(x) \leftarrow Child(x)$
// find objects about a child (strictly speaking, find instances of *Child*)

$q(x) \leftarrow CreatorName(x, y) \wedge (y =' paolo'), Title(x, z), (z\ simTxt\ 'tour')$
// find images made by Paolo whose title is about 'tour'

$q(x) \leftarrow ImageDescr(x, y) \wedge (y\ simImg\ urn2)$
// find images similar to a given image identified by *urn*2

$q(x) \leftarrow ImageObject(x) \wedge isAbout(x, y_1) \wedge Car(y_1) \wedge isAbout(x, y_2) \wedge Racing(y_2)$
// find image objects about cars racing

# Fuzzy LPs Basics [4, 6, 7, 22, 23, 29, 35]

- Many Logic Programming (LP) frameworks have been proposed to manage uncertain and imprecise information. They differ in:
  - The underlying notion of uncertainty and vagueness: probability, possibility, many-valued, fuzzy logics
  - How values, associated to rules and facts, are managed
- We consider fuzzy LPs, where
  - Truth space is $[0, 1]$
  - Interpretation is a mapping $I : B_{\mathcal{P}} \to [0, 1]$
  - Generalized LP rules are of the form

$$R(\mathbf{x}) \leftarrow \exists \mathbf{y}.f(R_1(\mathbf{z}_1), \ldots, R_l(\mathbf{z}_l), p_1(\mathbf{z}'_1), \ldots, p_h(\mathbf{z}'_h)) ,$$

    - Meaning of rules: “take the truth-values of all $R_i(\mathbf{z}_i), p_j(\mathbf{z}'_j)$, combine them using the truth combination function $f$, and assign the result to $R(\mathbf{x})$”

- Same meaning as for fuzzy DLR-Lite queries

$$R(\mathbf{x}, s) \leftarrow \exists \mathbf{y}.conj(\mathbf{x}, \mathbf{y}), s = f(p_1(\mathbf{z}_1), \ldots, p_{l+h}(\mathbf{z}_{l+h}))$$

1. $\mathbf{x}$ are the *distinguished variables*;
2. $s$ is the *score variable*, taking values in $[0, 1]$;
3. $\mathbf{y}$ are existentially quantified variables, called *non-distinguished variables*;
4. $conj(\mathbf{x}, \mathbf{y})$ is a list of atoms $R_i(\mathbf{z})$ in $KB$;
5. $\mathbf{z}$ are tuples of constants in $KB$ or variables in $\mathbf{x}$ or $\mathbf{y}$;
6. $\mathbf{z_i}$ are tuples of constants in $KB$ or variables in $\mathbf{x}$ or $\mathbf{y}$;
7. $p_i$ is an $n_i$-ary *fuzzy predicate* assigning to each $n_i$-ary tuple $\mathbf{c}_i$ the *score* $p_i(\mathbf{c}_i) \in [0, 1]$;
8. $f$ is a monotone *scoring* function $f\colon [0, 1]^{l+h} \to [0, 1]$, which combines the scores of the $n$ fuzzy predicates $p_i(\mathbf{c}_i)$

# Example: Soft shopping agent

- I may represent my preferences in Logic Programming with the rules

$$
\begin{array}{rcl}
Pref_1(x, p, s) & \leftarrow & HasPrice(x, p), LS(10000, 14000, p, s) \\
Pref_2(x, s) & \leftarrow & HasKM(x, k), LS(13000, 17000, k, s) \\
Buy(x, p, s) & \leftarrow & Pref_1(x, p, s_1), Pref_2(x, s_2), s = 0.7 \cdot s_1 + 0.3 \cdot s_2
\end{array}
$$

| ID | MODEL | PRICE | KM |
|---|---|---|---|
| 455 | MAZDA 3 | 12500 | 10000 |
| 34 | ALFA 156 | 12000 | 15000 |
| 1812 | FORD FOCUS | 11000 | 16000 |
| ⋮ | ⋮ | ⋮ | ⋮ |

- Problem: All tuples of the database have a score:
  - We cannot compute the score of all tuples, then rank them. Brute force approach not feasible.
- Top-$k$ problem: Determine efficiently just the top-$k$ ranked tuples, without evaluating the score of all tuples. E.g. top-3 tuples

| ID | PRICE | SCORE |
|---|---|---|
| 1812 | 11000 | 0.6 |
| 455 | 12500 | 0.56 |
| 34 | 12000 | 0.50 |

# Top-*k* retrieval in LPs

- If the database contains a huge amount of facts, a brute force approach fails:
    - one cannot anymore compute the score of all tuples, rank all of them and only then return the top-*k*
- Better solutions exists for restricted fuzzy LP languages: Datalog + restriction on the score combination functions appearing in the body [29, 32]

# Fuzzy DLPs Basics [10, 11, 27, 31]

- Combine fuzzy DLs with fuzzy LPs:
  - Like fuzzy LPs, but DL atoms and roles may appear in rules

$$\begin{array}{lcl} LowCarPrice(z) & \leftarrow & \min(made\_by(x,y), DL[ChineseCarCompany](y) \\ & & price(x,z)) \cdot DL[Low](z) \\ & & \\ Low & = & LS(5.000, 15.000) \\ ChineseCarCompany & \sqsubseteq & \exists has\_location.China \end{array}$$

- Knowledge Base is a pair $KB = \langle \mathcal{P}, \Sigma \rangle$, where
  - $\mathcal{P}$ is a fuzzy logic program
  - $\Sigma$ is a fuzzy DL knowledge base (set of assertions and inclusion axioms)

# Fuzzy DLPs Semantics

- Semantics: several approaches
- In principle, for each classical semantics based integration between DLs and LPs, there is be a fuzzy analogue
  - Pay attention, the fuzzy variant may add further technical and computational complications

1. Axiomatic approach: fuzzy DL atoms and roles are managed uniformely
2. Loosely Coupled approach: fuzzy DL atoms and roles are like “procedural attachments” (procedural calls to a fuzzy DL theorem prover)
3. Tightly coupled approach: The DL component restricts the models to be considered for the LP component

F. Bobillo and U. Straccia.
A fuzzy description logic with product t-norm.
In *Proceedings of the IEEE International Conference on Fuzzy Systems (Fuzz-IEEE-07)*. IEEE Computer Society, 2007.

F. Bobillo and U. Straccia.
Mixed integer programming, general concept inclusions and fuzzy description logics.
In *Proceedings of the 5th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-07)*, 2007.

D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, and R. Rosati.
Data complexity of query answering in description logics.
In *Proceedings of the Tenth International Conference on Principles of Knowledge Representation and Reasoning (KR-06)*, pages 260–270, 2006.

C. Damásio and M. Medina, J. Ojeda-Aciego.
A tabulation procedure for first-order residuated logic programs.
In *Proceedings of the 11th International Conference on Information Processing and Managment of Uncertainty in Knowledge-Based Systems, (IPMU-06)*, 2006.

P. Hájek.
*Metamathematics of Fuzzy Logic*.
Kluwer, 1998.

M. Kifer and V. Subrahmanian.
Theory of generalized annotated logic programming and its applications.
*Journal of Logic Programming*, 12:335–367, 1992.

L. V. Lakshmanan and N. Shiri.

A parametric approach to deductive databases with uncertainty.
*IEEE Transactions on Knowledge and Data Engineering*, 13(4):554–570, 2001.

C. Li, M. Soliman, K. Chang, and I. Ilyas.
RankSQL: Supporting ranking queries in relational database management systems, 2005.

Y. Loyer and U. Straccia.
Any-world assumptions in logic programming.
*Theoretical Computer Science*, 342(2-3):351–381, 2005.

T. Lukasiewicz.
Fuzzy description logic programs under the answer set semantics for the semanticweb.
In *Second International Conference on Rules and Rule Markup Languages for the Semantic Web (RuleML-06)*. IEEE, 2006.

T. Lukasiewicz and U. Straccia.
Tightly integrated fuzzy description logic programs under the answer semantics for the semantic web.
In *Proceedings of the First International Conference on Web Reasoning and Rule Systems (RR-07)*, number 4524 in Lecture Notes in Computer Science, pages 289—298. Springer Verlag, 2007.

T. Lukasiewicz and U. Straccia.
Tightly integrated fuzzy description logic programs under the answer semantics for the semantic web.
INFSYS RESEARCH REPORT 1843-07-03, INSTITUT FüR INFORMATIONSSYSTEME ARBEITSBEREICH WISSENSBASIERTE SYSTEME, Technische Universität Wien, 2007.

C. Mateis.
Extending disjunctive logic programming by t-norms.
In *Proceedings of the 5th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR-99)*, number 1730 in Lecture Notes in Computer Science, pages 290–304. Springer-Verlag, 1999.

C. Mateis.
Quantitative disjunctive logic programming: Semantics and computation.
*AI Communications*, 13:225–248, 2000.

M. Mazzieri.
A fuzzy rdf semantics to represent trust metadata.
In *Proceedings of the 1st Italian Semantic Web Workshop: Semantic Web Applications and Perspectives (SWAP 2004)*, 2004.

M. Mazzieri and A. F. Dragoni.
A fuzzy semantics for semantic web languages.
In *Proceedings of the Proceedings of the ISWC Workshop on Uncertainty Reasoning for the Semantic Web (URSW-05)*. CEUR Workshop Proceedings, 2005.

G. Stoilos, G. Stamou, and J. Z. Pan.
D2.5.6: Fuzzy reasoning extensions.
Knowledge Web Technical Report, 2007.

G. Stoilos, U. Straccia, G. Stamou, and J. Pan.
General concept inclusions in fuzzy description logics.
In *Proceedings of the 17th Eureopean Conference on Artificial Intelligence (ECAI-06)*, pages 457–461. IOS Press, 2006.

U. Straccia.
Reasoning within fuzzy description logics.
*Journal of Artificial Intelligence Research*, 14:137–166, 2001.

U. Straccia.
Description logics with fuzzy concrete domains.

In F. Bachus and T. Jaakkola, editors, *21st Conference on Uncertainty in Artificial Intelligence (UAI-05)*, pages 559–567, Edinburgh, Scotland, 2005. AUAI Press.

U. Straccia.
Fuzzy alc with fuzzy concrete domains.
In *Proceeedings of the International Workshop on Description Logics (DL-05)*, pages 96–103, Edinburgh, Scotland, 2005. CEUR.

U. Straccia.
Query answering in normal logic programs under uncertainty.
In *8th European Conferences on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU-05)*, number 3571 in Lecture Notes in Computer Science, pages 687–700, Barcelona, Spain, 2005. Springer Verlag.

U. Straccia.
Uncertainty management in logic programming: Simple and effective top-down query answering.
In R. Khosla, R. J. Howlett, and L. C. Jain, editors, *9th International Conference on Knowledge-Based & Intelligent Information & Engineering Systems (KES-05), Part II*, number 3682 in Lecture Notes in Computer Science, pages 753–760, Melbourne, Australia, 2005. Springer Verlag.

U. Straccia.
Annotated answer set programming.
In *Proceedings of the 11th International Conference on Information Processing and Managment of Uncertainty in Knowledge-Based Systems, (IPMU-06)*, pages 1212–1219. E.D.K., Paris, 2006.

U. Straccia.
Answering vague queries in fuzzy dl-lite.
In *Proceedings of the 11th International Conference on Information Processing and Managment of Uncertainty in Knowledge-Based Systems, (IPMU-06)*, pages 2238–2245. E.D.K., Paris, 2006.

U. Straccia.
A fuzzy description logic for the semantic web.
In E. Sanchez, editor, *Fuzzy Logic and the Semantic Web*, Capturing Intelligence, chapter 4, pages 73–90. Elsevier, 2006.

U. Straccia.
Fuzzy description logic programs.
In *Proceedings of the 11th International Conference on Information Processing and Managment of Uncertainty in Knowledge-Based Systems, (IPMU-06)*, pages 1818–1825. E.D.K., Paris, 2006.

U. Straccia.
Query answering under the any-world assumption for normal logic programs.
In *Proceedings of the 10th International Conference on Principles of Knowledge Representation (KR-06)*, pages 329–339. AAAI Press, 2006.

U. Straccia.
Towards top-k query answering in deductive databases.
In *Proceedings of the 2006 IEEE International Conference on Systems, Man and Cybernetics (SMC-06)*, pages 4873–4879. IEEE, 2006.

U. Straccia.
Towards top-k query answering in description logics: the case of DL-Lite.
In *Proceedings of the 10th European Conference on Logics in Artificial Intelligence (JELIA-06)*, number 4160 in Lecture Notes in Computer Science, pages 439–451, Liverpool, UK, 2006. Springer Verlag.

U. Straccia.
Uncertainty and description logic programs over lattices.
In E. Sanchez, editor, *Fuzzy Logic and the Semantic Web*, Capturing Intelligence, chapter 7, pages 115–133. Elsevier, 2006.

U. Straccia.
Towards vague query answering in logic programming for logic-based information retrieval.
In *World Congress of the International Fuzzy Systems Association (IFSA-07)*, Cancun, Mexico, 2007.

U. Straccia and G. Visco.
DLMedia: an ontology mediated multimedia information retrieval system.
In *Proceeedings of the International Workshop on Description Logics (DL-07)*, Insbruck, Austria, 2007. CEUR.

V. V., B. J., G. P., and H. T.
Fuzzy rdf in the semantic web: Deduction and induction.
In *Proceedings of Workshop on Data Analysis (WDA 2005)*, 2005.

P. Vojtáš.
Fuzzy logic programming.
*Fuzzy Sets and Systems*, 124:361–370, 2001.

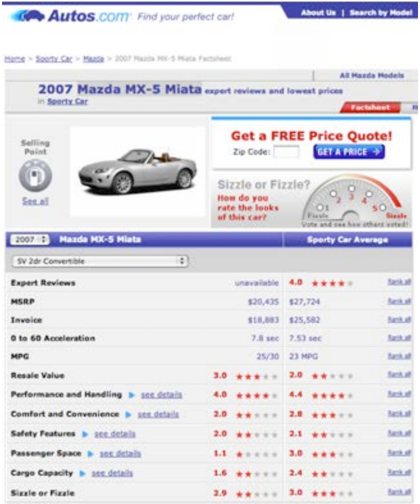
- Description logic programs that allow for dealing with probabilistic uncertainty and fuzzy vagueness.
- Semantically, probabilistic uncertainty can be used for data integration and ontology mapping, and fuzzy vagueness can be used for expressing vague concepts.
- Technically, allows for defining different rankings on ground atoms using fuzzy vagueness, and then for a probabilistic merging of these rankings using probabilistic uncertainty.
- Query processing based on fixpoint iterations.

Suppose a person would like to buy “a sports car that costs at most about 22 000 euro and that has a power of around 150 HP”.

In todays Web, the buyer has to *manually*

- search for car selling web sites, e.g., using Google;
- select the most promising sites;
- browse through them, query them to see the cars that each site sells, and match the cars with the requirements;
- select the offers in each web site that match the requirements; and
- eventually merge all the best offers from each site and select the best ones.


Autos.com Find your perfect car!
About Us | Search by Model | Search by Class
Best in Class
Thousands of cars & trucks ranked by industry professionals
Get Rankings, Reviews, Pricing & more...
Get the lowest price on your new car...
FREE Quote
no-obligation
GET A FREE QUOTE
Search for Used Cars
Passenger Cars
Best New Mid-Size Car
Compact Cars
Mid-Size Cars
Sporty Cars
2007 Volkswagen Passat
Luxury Cars
Best New Near-Luxury Car
Near-Luxury Cars
Mid-Luxury Cars
Ultra-Luxury Cars
2006 Acura TSX
Trucks
Best New Full-Size Truck
Compact Trucks
Full-Size Trucks
2006 GMC Sierra 1500HD
SUVs
Best New Mid-Size SUV
Compact SUVs
Mid-Size SUVs
Full-Size SUVs
Luxury SUVs
2006 Volkswagen Touareg
Vans
Best New Minivan
Minivans
Full-Size Vans
2006 Toyota Sienna
Compare Cars
See how your choices stack up
Compare NOW
Next Generation Nissan Altima
Find premium used cars near you!
Over 500,000 cars
Top dealers
SEE LISTINGS
Sizzle or Fizzle?


Autos.com Find your perfect car!
About Us | Search by Model
All Mazda Models
2007 Mazda MX-5 Miata expert reviews and lowest prices
in Sporty Car
Factsheet
Selling Point
See all
Get a FREE Price Quote!
Zip Code:
GET A PRICE
Sizzle or Fizzle?
How do you rate the looks of this car?
Vote and see how others voted!
2007
Mazda MX-5 Miata
Sporty Car Average
SV 2dr Convertible
Expert Reviews unavailable 4.0
MSRP $20,435 $27,724
Invoice $18,883 $25,582
0 to 60 Acceleration 7.8 sec 7.53 sec
MPG 25/30 23 MPG
Resale Value 3.0 2.0
Performance and Handling see details 4.0 4.4
Comfort and Convenience see details 2.0 2.8
Safety Features see details 2.0 2.1
Passenger Space see details 1.1 3.0
Cargo Capacity see details 1.6 2.4
Sizzle or Fizzle 2.9 3.0

A *shopping agent* may support us, *automatizing* the whole process once it receives the request/query *q* from the buyer:

- The agent selects some sites/resources *S* that it considers as *relevant* to *q* (represented by probabilistic rules).
- For the top-*k* selected sites, the agent has to reformulate *q* using the terminology/ontology of the specific car selling site (which is done using probabilistic rules).
- The query *q* may contain many *vague/fuzzy* concepts such as “the price is around 22 000 euro or less”, and so a car may *match* *q* to a *degree*. So, a resource returns a ranked list of cars, where the ranks depend on the degrees to which the cars match *q*.
- Eventually, the agent integrates the ranked lists (using probabilities) and shows the top-*n* items to the buyer.

$$Cars \sqcup Trucks \sqcup Vans \sqcup SUVs \sqsubseteq Vehicles$$
$$PassengerCars \sqcup LuxuryCars \sqsubseteq Cars$$
$$CompactCars \sqcup MidSizeCars \sqcup SportyCars \sqsubseteq PassengerCars$$

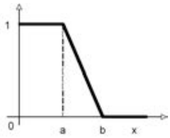
$$\begin{aligned} Cars \sqsubseteq\ & (\exists hasReview.Integer) \sqcap (\exists hasInvoice.Integer) \\ & \sqcap (\exists hasResellValue.Integer) \sqcap (\exists hasMaxSpeed.Integer) \\ & \sqcap (\exists hasHorsePower.Integer) \sqcap \ldots \end{aligned}$$

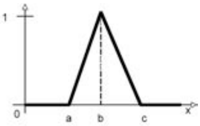
$$\begin{aligned} MazdaMX5Miata :\ & SportyCar \sqcap (\exists hasInvoice.18883) \\ & \sqcap (\exists hasHorsePower.166) \sqcap \ldots \\ MitsubishiEclipseSpyder :\ & SportyCar \sqcap (\exists hasInvoice.24029) \\ & \sqcap (\exists hasHorsePower.162) \sqcap \ldots \end{aligned}$$

We may now encode “costs at most about 22 000 euro ” and “has a power of around 150 HP” in the buyer’s request through the following concepts $C$ and $D$, respectively:

$$C = \exists hasInvoice.LeqAbout22000 \text{ and}$$
$$D = \exists hasHorsePower.Around150HP,$$

where $LeqAbout22000 = ls(22000, 25000)$ and $Around150HP = tri(125, 150, 175)$.



The following fuzzy dl-rule encodes the buyer's request "a sports car that costs at most about 22 000 euro and that has a power of around 150 HP".

$$
\begin{aligned}
query(x) \quad \leftarrow_{\otimes} \quad & SportyCar(x) \wedge_{\otimes} \\
& hasInvoice(x, y_1) \wedge_{\otimes} \\
& DL[LeqAbout22000](y_1) \wedge_{\otimes} \\
& hasHorsePower(x, y_2) \wedge_{\otimes} \\
& DL[Around150HP](y_2) \geq 1 \,.
\end{aligned}
$$

Here, $\otimes$ is the Gödel t-norm (that is, $x \otimes y = \min(x, y)$).

The buyer's request, but in a "different" terminology:

$$\begin{aligned} query(x) \leftarrow_\otimes\ & SportsCar(x) \wedge_\otimes hasPrice(x, y_1) \wedge_\otimes hasPower(x, y_2) \wedge_\otimes \\ & DL[LeqAbout22000](y_1) \wedge_\otimes DL[Around150HP](y_2) \geq 1 \end{aligned}$$

Ontology alignment mapping rules:

$$\begin{aligned} SportsCar(x) &\leftarrow_\otimes DL[SportyCar](x) \wedge_\otimes sc_{pos} \geq 0.9 \\ hasPrice(x) &\leftarrow_\otimes DL[hasInvoice](x) \wedge_\otimes hi_{pos} \geq 0.8 \\ hasPower(x) &\leftarrow_\otimes DL[hasHorsePower](x) \wedge_\otimes hhp_{pos} \geq 0.8\,, \end{aligned}$$

Probability distribution $\mu$:

$$\begin{array}{ll} \mu(sc_{pos}) = 0.91 & \mu(sc_{neg}) = 0.09 \\ \mu(hi_{pos}) = 0.78 & \mu(hi_{neg}) = 0.22 \\ \mu(hhp_{pos}) = 0.83 & \mu(hhp_{neg}) = 0.17\,. \end{array}$$

The following are some tight consequences:

$$KB \;\Vdash\!\sim_{tight}\; (\mathbf{E}[query((MazdaMX5Miata)])[0.21, 0.21]$$
$$KB \;\Vdash\!\sim_{tight}\; (\mathbf{E}[query((MitsubishiEclipseSpyder)])[0.19, 0.19]\,.$$

Informally, the expected degree to which *MazdaMX5Miata* matches the query $q$ is 0.21, while the expected degree to which *MitsubishiEclipseSpyder* matches the query $q$ is 0.19,

Thus, the shopping agent ranks the retrieved items as follows:

| rank | item | degree |
|---|---|---|
| 1. | *MazdaMX5Miata* | 0.21 |
| 2. | *MitsubishiEclipseSpyder* | 0.19 |