Fuzzy & Annotated Semantic Web Languages

Umberto Straccia

ISTI - CNR, Pisa, Italy

umberto.straccia@isti.cnr.it
www.umbertostraccia.it

February 2021
About Fuzziness
   On the Existence of Fuzzy Concepts
   On the Existence of Fuzzy Objects
   Fuzzy Statements
   Sources of Fuzziness
   Uncertainty vs Fuzziness: a clarification

From Fuzzy Sets to Mathematical Fuzzy Logic
   Fuzzy Sets Basics
   Mathematical Fuzzy Logics Basics

Fuzzy Semantic Web Languages
   Introduction
   The case of Fuzzy RDFS
   The case of Fuzzy Description Logics
   The case of Fuzzy Logic Programs

Conclusions
About Fuzziness
On the Existence of Fuzzy Concepts

What are fuzzy concepts and do they exist?
Try to answer: What is this picture about?

(Registan Square, Samarkand, Uzbekistan, Umberto)
Fuzzy concept: no unambiguous definition, e.g.
  - What is a picture or piece of text about?
  - What is a tall person?
  - What is a high temperature?
  - What is nice weather?
  - What is an adventurous trip?

Fuzzy concepts:
  - Are abundant in everyday speech and almost inevitable
  - Their meaning is typically subjective and context dependent
What are fuzzy objects and do they exist?
Are there fuzzy objects in the pictures?

(Erg Chebbi, pre-Sahara dunes, Merzouga, Morocco)
(The Sun)
Fuzzy object: its identity is lacking in clarity
  - Cloud
  - Dunes
  - Sun

Fuzzy objects:
  - Are not identical to anything, except to themselves (reflexivity)
  - Are characterised by a fuzzy identity relation (e.g. a similarity relation)
A statement is fuzzy whenever it involves fuzzy concepts or fuzzy objects.

The truth of a fuzzy statement is a matter of degree,

- it is intrinsically difficult to establish whether the statement is entirely true or false (can be e.g. almost true)

The weather temperature is $33^\circ$C. Is it hot?
“Find top-$k$ image regions about animals”

$\text{Query}(x) \leftarrow \text{ImageRegion}(x) \land \text{isAbout}(x, y) \land \text{Animal}(y)$
Sources of Fuzziness: Lifezone mapping

To which degree do certain areas have a specific bioclimate?

Holdridge life zones of USA
Sources of Fuzziness: ARPAT, Air quality in the province of Lucca

Sintesi dei dati rilevati dalle ore 0 alle ore 24 del giorno domenica 14/02/2010

<table>
<thead>
<tr>
<th>Stazione</th>
<th>Tipo stazione</th>
<th>SO\textsubscript{2} µg/m\textsuperscript{3} (media su 24h)</th>
<th>NO\textsubscript{2} µg/m\textsuperscript{3} (max oraria)</th>
<th>CO mg/m\textsuperscript{3} (max oraria)</th>
<th>O\textsubscript{3} µg/m\textsuperscript{3} (max oraria)</th>
<th>PM\textsubscript{10} µg/m\textsuperscript{3} (media su 24h)</th>
<th>Giudizio di qualità dell'aria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucca</td>
<td>P.za San Michele (***)</td>
<td>urbana - traffico</td>
<td>75</td>
<td>---</td>
<td>---</td>
<td>56</td>
<td>Scadente</td>
</tr>
<tr>
<td>Lucca</td>
<td>V.le Carducci</td>
<td>urbana - traffico</td>
<td>---</td>
<td>2</td>
<td>---</td>
<td>75</td>
<td>Pessima</td>
</tr>
<tr>
<td>Lucca</td>
<td>Carignano (***)</td>
<td>rurale - fondo</td>
<td>---</td>
<td>---</td>
<td>67 (h.18\textsuperscript{*})</td>
<td>---</td>
<td>Buona</td>
</tr>
<tr>
<td>Viareggio</td>
<td>Largo Risorgimento</td>
<td>urbana - traffico</td>
<td>---</td>
<td>1,7</td>
<td>---</td>
<td>n.d.</td>
<td>Buona</td>
</tr>
<tr>
<td>Viareggio</td>
<td>Via Maroncelli (***)</td>
<td>urbana - fondo</td>
<td>121</td>
<td>---</td>
<td>60 (h.17\textsuperscript{*})</td>
<td>45</td>
<td>Accettabile</td>
</tr>
<tr>
<td>Capannori</td>
<td>V. di Piaggia (***)</td>
<td>urbana - fondo</td>
<td>79</td>
<td>2</td>
<td>---</td>
<td>59</td>
<td>Accettabile</td>
</tr>
<tr>
<td>Porcari</td>
<td>V. Carrara (***)</td>
<td>periferica - fondo</td>
<td>72</td>
<td>---</td>
<td>82 (h.16\textsuperscript{*})</td>
<td>63</td>
<td>Scadente</td>
</tr>
</tbody>
</table>

Giudizio di qualità

<table>
<thead>
<tr>
<th>Giudizio di qualità</th>
<th>SO\textsubscript{2} µg/m\textsuperscript{3} (media su 24h)</th>
<th>NO\textsubscript{2} µg/m\textsuperscript{3} (max oraria)</th>
<th>CO mg/m\textsuperscript{3} (max oraria)</th>
<th>O\textsubscript{3} µg/m\textsuperscript{3} (max oraria)</th>
<th>PM\textsubscript{10} µg/m\textsuperscript{3} (media su 24h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buona</td>
<td>0-50</td>
<td>0-50</td>
<td>0-2,5</td>
<td>0-120</td>
<td>0-25</td>
</tr>
<tr>
<td>Accettabile</td>
<td>51-125</td>
<td>51-200</td>
<td>2,6-15</td>
<td>121-180</td>
<td>26-50</td>
</tr>
<tr>
<td>Scadente</td>
<td>126-250</td>
<td>201-400</td>
<td>15,1-30</td>
<td>181-240</td>
<td>51-74</td>
</tr>
<tr>
<td>Pessima</td>
<td>&gt;250</td>
<td>&gt;400</td>
<td>&gt;30</td>
<td>&gt;240</td>
<td>&gt;74</td>
</tr>
</tbody>
</table>

http://www.arpat.toscana.it/
2,889 Reviews from our TripAdvisor Community

<table>
<thead>
<tr>
<th>Traveler rating</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>1,467</td>
</tr>
<tr>
<td>Very good</td>
<td>1,029</td>
</tr>
<tr>
<td>Average</td>
<td>271</td>
</tr>
<tr>
<td>Poor</td>
<td>86</td>
</tr>
<tr>
<td>Terrible</td>
<td>36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>See reviews for</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Families</td>
<td>402</td>
</tr>
<tr>
<td>Couples</td>
<td>1,154</td>
</tr>
<tr>
<td>Solo</td>
<td>243</td>
</tr>
<tr>
<td>Business</td>
<td>671</td>
</tr>
</tbody>
</table>

Rating summary:
- Location: 3 stars
- Sleep Quality: 4 stars
- Rooms: 4 stars
- Service: 3 stars
- Value: 2 stars
- Cleanliness: 4 stars
Uncertainty vs Fuzziness: a clarification

- Initial difficulty:
  - Understand the conceptual differences between uncertainty and fuzziness
- Main problem:
  - Interpreting a degree as a measure of uncertainty rather than as a measure of fuzziness
Uncertain Statements

- A statement is **true or false** in any world/interpretation
  - We are “**uncertain**” about which world to consider as the actual one
  - We may have e.g. a probability/possibility distribution over possible worlds
- E.g., of uncertain statement: “it will rain tomorrow”
  - We cannot exactly establish whether it will rain tomorrow or not, due to our **incomplete** knowledge about our world
  - But, we may estimate to which **degree** this is e.g. **probable/possible**
Fuzzy Statements

- A statement is **fuzzy** if it involves fuzzy concepts/objects.
- A statement is **true** to some **degree**, which is taken from a truth space (usually $[0, 1]$).
- E.g. of fuzzy statement: “heavy rain”
  - is graded and the degree depends on the amount of rain is falling.
In weather forecasts one may find:

- **Rain.** Falling drops of water larger than 0.5 mm in diameter. “Rain” usually implies that the rain will fall steadily over a period of time;

- **Light rain.** Rain falls at the rate of 2.6 mm or less an hour;

- **Moderate rain.** Rain falls at the rate of 2.7 mm to 7.6 mm an hour;

- **Heavy rain.** Rain falls at the rate of 7.7 mm an hour or more.

- Quite harsh distinction: \( R = 7.7 \text{mm/h} \rightarrow \text{heavy rain} \)
  \( R = 7.6 \text{mm/h} \rightarrow \text{moderate rain} \)

- Unsatisfactory:
  - the more rain is falling, the more the sentence “heavy rain” is true
  - vice-versa, the less rain is falling the more the sentence “heavy rain” is false
- I.e., the sentence “heavy rain” is intrinsically graded.
- More fine grained approach:
  - Define the various types of rains as
    - Light rain, moderate rain and heavy rain are fuzzy concepts.
Are there sentences combining the two orthogonal concepts of uncertainty and fuzziness?

Yes, and we use them daily!

E.g. “There will be heavy rain tomorrow.”

This type of sentences are called **uncertain fuzzy sentences**

Essentially, there is

- **uncertainty** about the world we will have tomorrow
- **fuzziness** about the various types of rain
From Fuzzy Sets to Mathematical Fuzzy Logic
Fuzzy Sets Basics

From Crisp Sets to Fuzzy Sets.

- Let $X$ be a universal set of objects
- The crisp membership function of a set $A \subseteq X$:
  $$\chi_A: X \to \{0, 1\}$$
  where $\chi_A(x) = 1$ iff $x \in A$
- Fuzzy set $A$:
  $$\chi_A: X \to [0, 1]$$
  or simply $A: X \to [0, 1]$
- Example: the fuzzy set $C = \{x \mid x$ is a day with heavy precipitation rate $R\}$

is defined via the membership function

$$\chi_C(x) = \begin{cases} 
1 & \text{if } R \geq 7.5 \\
(x - 5)/2.5 & \text{if } R \in [5, 7.5) \\
0 & \text{otherwise}
\end{cases}$$
Fuzzy membership functions may depend on the **context** and may be **subjective**

**Shape** may be quite different

Usually, it is sufficient to consider functions:

(a) Trapezoidal $trz(a, b, c, d)$; (b) Triangular $tri(a, b, c)$; (c) left-shoulder $ls(a, b)$; (d) right-shoulder $rs(a, b)$
Fuzzy Sets Construction

- Simple and typically satisfactory method (numerical domain):
  - uniform partitioning into 5 fuzzy sets

Fuzzy sets construction using trapezoidal functions

Fuzzy sets construction using triangular functions
Another popular method is based on clustering.

Use Fuzzy C-Means to cluster data into 5 clusters.

- Fuzzy C-Means extends K-Means to accommodate graded membership.

From the clusters $c_1, \ldots, c_5$ take the centroids $\pi_1, \ldots, \pi_5$.

Build the fuzzy sets from the centroids.
Standard fuzzy set operations are not the only ones. Most notable ones are triangular norms:
- t-norm $\otimes$ for set intersection
- t-conorm $\oplus$ (also called s-norm) for set union
- negation $\ominus$ for set complementation
- implication $\Rightarrow$ for set inclusion

These functions satisfy some properties that one expects to hold.
One distinguishes three different sets of fuzzy set operations (called fuzzy logics)

- Łukasiewicz, Gödel, and Product logic
- Standard Fuzzy Logic (SFL) is a sublogic of Łukasiewicz

\[
\begin{align*}
\min(a, b) &= a \otimes (a \Rightarrow b), \quad \max(a, b) = 1 - \min(1 - a, 1 - b) \\
\end{align*}
\]

<table>
<thead>
<tr>
<th>Łukasiewicz Logic</th>
<th>Gödel Logic</th>
<th>Product Logic</th>
<th>SFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a \otimes b)</td>
<td>(\max(a + b - 1, 0))</td>
<td>(\min(a, b))</td>
<td>(a \cdot b)</td>
</tr>
<tr>
<td>(a \oplus b)</td>
<td>(\min(a + b, 1))</td>
<td>(\max(a, b))</td>
<td>(a + b - a \cdot b)</td>
</tr>
<tr>
<td>(a \Rightarrow b)</td>
<td>(\min(1 - a + b, 1))</td>
<td>(\begin{cases} 1 &amp; \text{if } a \leq b \ b &amp; \text{otherwise} \end{cases})</td>
<td>(\min(1, b/a))</td>
</tr>
<tr>
<td>(\ominus a)</td>
<td>(1 - a)</td>
<td>(\begin{cases} 1 &amp; \text{if } a = 0 \ 0 &amp; \text{otherwise} \end{cases})</td>
<td>(\begin{cases} 1 &amp; \text{if } a = 0 \ 0 &amp; \text{otherwise} \end{cases})</td>
</tr>
</tbody>
</table>

Mostert–Shields theorem: any continuous t-norm can be obtained as an ordinal sum of Ł, G and P.
OWL 2 is grounded on Mathematical Logic

Fuzzy OWL 2 is grounded on Mathematical Fuzzy Logic

A statement is graded

Truth space: set of truth values $L$

Given a statement $\phi$

Fuzzy Interpretation: a function $\mathcal{I}$ mapping $\phi$ into $L$, i.e.

$$\mathcal{I}(\varphi) \in L$$

Usually

\[
L = [0, 1] \\
L_n = \{0, \frac{1}{n}, \ldots, \frac{n-2}{n-1}, \ldots, 1\} \quad (n \geq 1)
\]
Fuzzy statement: for formula $\phi$ and $r \in [0, 1]$

$$\langle \phi, r \rangle$$

*The degree of truth of $\phi$ is equal or greater than $r$*
Fuzzy Semantic Web Languages
The Semantic Web Family of Languages

- Wide variety of languages
  - **RDFS**: Triple language, -Resource Description Framework
    - The logical counterpart is \( \rho_{df} \)
  - **RIF**: Rule language, -Rule Interchange Format,
    - Relate to the *Logic Programming* (LP) paradigm
  - **OWL 2**: Conceptual language, -Ontology Web Language
    - Relate to Description Logics (DLs)
RDFS: the triple language

\langle subject, predicate, object \rangle

e.g. \langle umberto, born, zurich \rangle
OWL 2 family: an object oriented language

class PERSON partial
restriction (hasName someValuesFrom String)
restriction (hasBirthPlace someValuesFrom GEOPLACE)
...

OWL 2 Profiles

OWL 2 EL
- Useful for large size of properties and/or classes
- The EL acronym refers to the $\mathcal{EL}$ family of DLs

OWL 2 QL
- Useful for very large volumes of instance data
- Conjunctive query answering via query rewriting and SQL
- OWL 2 QL relates to the DL family $DL$-$Lite$

OWL 2 RL
- Useful for scalable reasoning without sacrificing too much expressive power
- OWL 2 RL maps to Datalog
RIF/RuleML family: the rule language

\[\text{Forall } ?\text{Buyer} ?\text{Item} ?\text{Seller} \]
Important point: RDFS, OWL 2 and RIF/RuleML are logical languages

- RDFS: logic with intensional semantics
- OWL 2: relates to the *Description Logics* family
- RIF/RuleML: relates to the *Logic Programming* paradigm (e.g., Datalog, Datalog\(^\pm\))
- OWL 2 and RIF/RuleML have extensional semantics
The case of Fuzzy RDFS
Fuzzy RDFS

- Triples may have attached a degree $n$ in $L$ or $L_n$
  \[ \langle (\text{subject}, \text{predicate}, \text{object}), n \rangle \]
- Meaning: the degree of truth of the statement is at least $n$
- Example:
  \[ \langle (o1, \text{IsAbout}, \text{snoopy}), 0.8 \rangle \]
- How to represent fuzzy triples in RDFS?
  - Use reification method:
    \[ (s1, \text{hasObj}, o1), (s1, \text{hasRel}, \text{IsAbout}), (s1, \text{hasObj}, \text{snoopy}), (s1, \text{hasDeg0.8}) \]
- Unfortunately, RDFS is lacking the "annotation property" of triples
Fuzzy RDFS Query Answering

- **Conjunctive query**: extends a crisp RDF query and is of the form

  $$\langle q(x), s \rangle \leftarrow \exists y. \langle \tau_1, s_1 \rangle, \ldots, \langle \tau_n, s_n \rangle,$$

  $$s = f(s_1, \ldots, s_n, p_1(z_1), \ldots, p_h(z_h))$$

  where

  - $$\tau_i$$ triples involving literals and variables in $$x, y$$
  - $$z_i$$ are tuples of literals or variables in $$x$$ or $$y$$
  - $$p_j(t) \in [0, 1]$$
  - $$f$$ is a *scoring* function $$f: ([0, 1])^{n+h} \rightarrow [0, 1]$$

- **Example**:

  $$\langle q(x), s \rangle \leftarrow \langle (x, \text{type}, \text{SportCar}), s_1 \rangle, (x, \text{hasPrice}, y), s = s_1 \cdot \text{cheap}(y)$$

  where e.g. $$\text{cheap}(y) = \text{ls}(0, 10000, 12000)(y)$$, has intended meaning to “retrieve all cheap sports car”
Example

Consider the query

\[
\langle q(x), s \rangle \leftarrow \langle (x, \text{IsAbout}, y), s_1 \rangle, \langle (y, \text{type}, \text{Animal}), s_2 \rangle, s = s_1 \cdot s_2
\]

Then

\[
\text{ans}(G, q) = \{ \langle o1, 0.32 \rangle, \langle o2, 0.63 \rangle \}
\]
Annotation domains & RDFS

- Generalisation of fuzzy RDFS
  - a triple is annotated with a value taken from a so-called annotation domain, rather than with a value in [0,1]
  - allows to deal with several domains (such as, fuzzy, temporal, provenance) and their combination, in a uniform way
- Fuzzyness
  - $\langle (\text{HolidayInnHotel}, \text{closeTo}, \text{IEA17 Venue}), 0.7 \rangle$
  - true to some degree
- Time
  - $\langle (\text{umberto}, \text{workedFor}, \text{IEI}), [1992, 2001] \rangle$
  - true during 1992–2001
- Provenance
  - $\langle (\text{umberto}, \text{knows}, \text{salem}), \text{http://www.straccia.info/foaf.rdf} \rangle$
  - true in http://www.straccia.info/foaf.rdf
- Multiple Domains:
  - $\langle (\text{CountryXYZ}, \text{type}, \text{Dangerous}), \langle [1975, 1983], 0.8, 0.6 \rangle \rangle$

$Time \times Fuzzy \times Trust$
Annotation Domain: idempotent, commutative semi-ring

\[ D = \langle L, \oplus, \otimes, \bot, \top \rangle \]

where \( \oplus \) is \( \top \)-annihilating, i.e.

1. \( \oplus \) is idempotent, commutative, associative;
2. \( \otimes \) is commutative and associative;
3. \( \bot \oplus \lambda = \lambda, \top \otimes \lambda = \lambda, \bot \otimes \lambda = \bot, \) and \( \top \oplus \lambda = \top; \)
4. \( \otimes \) is distributive over \( \oplus \),
   i.e. \( \lambda_1 \otimes (\lambda_2 \oplus \lambda_3) = (\lambda_1 \otimes \lambda_2) \oplus (\lambda_1 \otimes \lambda_3); \)

Induced partial order:

\[ \lambda_1 \preceq \lambda_2 \iff \lambda_1 \oplus \lambda_2 = \lambda_2 \]

Annotated triple: for \( \lambda \in L \)

\[ \langle (s, p, o), \lambda \rangle \]
The case of Fuzzy Description Logics
For a degree $n$ in $L$ or $L_n$

- $\langle a: C, n \rangle$ states that $a$ is an instance of concept/class $C$ with degree at least $n$
- $\langle C_1 \sqsubseteq C_2, n \rangle$ states that class $C_1$ is subclass of $C_2$ to degree $n$
Towards Fuzzy OWL 2 and its Profiles

- Fuzzy OWL 2 added value:
  - fuzzy concrete domains (e.g., young)
  - modifiers (e.g., very young)
  - other extensions:
    - aggregation functions: weighted sum, OWA, fuzzy integrals
    - fuzzy rough sets
    - fuzzy spatial relations
    - fuzzy numbers, ...
E.g., Small, Young, High, etc. with explicit membership function

Representation of Young Person:

\[
\text{Minor} = \text{Person} \cap \exists \text{hasAge}. \leq 18 \\
\text{YoungPerson} = \text{Person} \cap \exists \text{hasAge}.ls(10, 30)
\]

Representation of Heavy Rain:

\[
\text{HeavyRain} = \text{Rain} \cap \exists \text{hasPrecipitationRate}.rs(5, 7.5)
\]
Fuzzy Modifiers

- Very, moreOrLess, slightly, etc.
- Representation of Sport Car
  \[
  \text{SportsCar} = \text{Car} \cap \exists \text{speed.very}(rs(80, 250))
  \]
- Representation of Very Heavy Rain
  \[
  \text{VeryHeavyRain} = \text{Rain} \cap \exists \text{hasPrecipitationRate.very}(rs(5, 7.5))
  \]
Aggregation Operators

- **Aggregation operators**: aggregate concepts, using functions such as the mean, median, weighted sum operators, etc.
- Allows to express the concept

\[ 0.3 \cdot \text{ExpensiveHotel} + 0.7 \cdot \text{LuxuriousHotel} \sqsubseteq \text{GoodHotel} \]

- A good hotel is the weighted sum of being an expensive and luxurious hotel.
- Aggregated concepts are popular in robotics:
  - to recognise complex objects from atomic ones
Fuzzy DLs Query Answering

- **Conjunctive query**: similar to fuzzy RDFS CQs:

\[
\langle q(x), s \rangle \leftarrow \exists y. \langle \tau_1, s_1 \rangle, \ldots, \langle \tau_n, s_n \rangle, \\
s = f(s_1, \ldots, s_n, p_1(z_1), \ldots, p_h(z_h))
\]

where

- \( \tau_1, \ldots, \tau_n \) are expressions \( A(z) \) or \( R(z, z') \), where \( A \) is a concept name, \( R \) is a role name, \( z, z' \) are individuals or variables in \( x \) or \( y \)

- **Example:**

\[
\langle q(x), s \rangle \leftarrow \langle \text{SportCar}(x), s_1 \rangle, \text{hasPrice}(x, y), s = s_1 \cdot \text{cheap}(y)
\]

where e.g. \( \text{cheap}(y) = \text{ls}(10000, 12000)(y) \), has intended meaning to retrieve all cheap sports car.
Some Applications

- (Multimedia) Information retrieval
- Recommendation systems
- Image interpretation
- Ambient intelligence
- Ontology merging
- Matchmaking
- Decision making
- Summarization
- Robotics perception
- Software design
- Machine learning
Consider the query

$$\langle q(x), s \rangle \leftarrow \langle \text{IsAbout}(x, y), s_1 \rangle, \langle \text{Animal}(y), s_2 \rangle, s = s_1 \cdot s_2$$

Then

$$\text{ans}(G, q) = \{\langle o1, 0.32 \rangle, \langle o2, 0.63 \rangle\}, \quad \text{ans}_1(G, q) = \{\langle o2, 0.63 \rangle\}$$
Example (Simplified Matchmaking)

- 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 pay not more than around 30000 €
- Classical sets: 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 €

\[
\begin{align*}
\text{AudiTT} &= \text{SportsCar} \sqcap \exists \text{hasPrice}.rs(30500, 31500) \\
\text{Query} &= \text{SportsCar} \sqcap \exists \text{hasPrice}.ls(30000, 32000)
\end{align*}
\]

- Highest degree to which the concept
  \[ C = \text{AudiTT} \sqcap \text{Query} \]
  is satisfiable is 0.75 (the degree to which the Audi TT and the query matches is 0.75)

- The car may be sold at 31250 €
Example: Learning fuzzy GCIs from OWL data

- Learning of fuzzy GCIs from crisp OWL data
- Use Case: What are Good hotels, using TripAdvisor data?
  - Given
    - OWL 2 Ontology about meaningful city entities and their descriptions
    - TripAdvisor data about hotels and user judgments
  - We have learnt that in e.g., Pisa, Italy

\[
\langle \exists \text{hasAmenity.Babysitting} \sqcap \exists \text{hasPrice.fair} \sqsubseteq \text{Good_Hotel}, 0.782 \rangle
\]

“A hotel having babysitting as amenity and a fair price is a good hotel (to degree 0.782)”

- Real valued price attribute hasPrice has been automatically fuzzyfied
Representing Fuzzy OWL Ontologies in OWL

- OWL 2 is W3C standard, with classical logic semantics
  - Hence, cannot support natively Fuzzy Logic
- However, **Fuzzy OWL 2**, has been defined using OWL 2
  - Uses the axiom annotation feature of OWL 2
- Any Fuzzy OWL 2 ontology is a legal OWL 2 ontology
- A java parser for Fuzzy OWL 2 exists
- Protégé plug-in exists to encode Fuzzy OWL ontologies
For OWL 2, it is like for RDFS, but annotation domain has to be a complete lattice.

Exception for OWL profiles OWL EL, OWL QL and OWL RL: annotation domains may be as for RDFS.
The case of Fuzzy Logic Programs
Truth space is $[0, 1]$ or $\{0, \frac{1}{n}, \ldots, \frac{n-2}{n-1}, \ldots, 1\}$ \ $(n \geq 1)$

Generalized LP rules are of the form

$$\langle R(x), s \rangle \leftarrow \exists y. \langle R_1(z_1), s_1 \rangle, \ldots, \langle R_k(z_k), s_k \rangle, \ s = f(s_1, \ldots, s_k, p_1(z'_1), \ldots, p_h(z'_h))$$

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

Facts: ground expressions of the form $\langle R(c), n \rangle$

Meaning of facts: “the degree of truth of $R(c)$ is at least $n$”

Fuzzy LP: a set of facts (extensional database) and a set of rules (intentional database). No extensional relation may occur in the head of a rule
Example: Soft shopping agent

User preferences:

\[ \langle \text{Pref}_1(x, p), s \rangle \leftarrow \text{HasPrice}(x, p), s = l_{s}(10000, 14000)(p) \]
\[ \langle \text{Pref}_2(x), s \rangle \leftarrow \text{HasKM}(x, k), s = l_{s}(13000, 17000)(k) \]
\[ \langle \text{Buy}(x, p), s \rangle \leftarrow \langle \text{Pref}_1(x, p), sp \rangle, \langle \text{Pref}_2(x, k), sk \rangle, s = 0.7 \cdot sp + 0.3 \cdot sk \]

<table>
<thead>
<tr>
<th>ID</th>
<th>MODEL</th>
<th>PRICE</th>
<th>KM</th>
</tr>
</thead>
<tbody>
<tr>
<td>455</td>
<td>MAZDA 3</td>
<td>12500</td>
<td>10000</td>
</tr>
<tr>
<td>34</td>
<td>ALFA 156</td>
<td>12000</td>
<td>15000</td>
</tr>
<tr>
<td>1812</td>
<td>FORD FOCUS</td>
<td>11000</td>
<td>16000</td>
</tr>
</tbody>
</table>

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 for very large databases

Top-k fuzzy LP problem: Determine efficiently just the top-k ranked tuples, without evaluating the score of all tuples. E.g. top-3 tuples

<table>
<thead>
<tr>
<th>ID</th>
<th>PRICE</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1812</td>
<td>11000</td>
<td>0.6</td>
</tr>
<tr>
<td>455</td>
<td>12500</td>
<td>0.56</td>
</tr>
<tr>
<td>34</td>
<td>12000</td>
<td>0.50</td>
</tr>
</tbody>
</table>
There are quite many LP/ASP systems (monotone/non-monotone)
- each with its own feature set
- some with SW interface
  - SWIProlog, DLV, ...

More SW related: various frameworks exist ...
- SWRL: rules with concept and role expressions as atoms
- Datalog\(^\pm\): Datalog with existential restriction on rule head
- RuleML: quite large range of features

The development of fuzzy LPs is essentially in parallel with that of classical LPs (since early ’80s)

A common problem with LP frameworks (incl. fuzzy)
- Lack of standardised language and semantics
- SWRL, RuleML are exceptions
For Datalog, it is like for RDFS
- The reasoning decision problems’ complexity is inherited from their fuzzy variants. Decidable if lattice and truth space are finite, else undecidable in general
Conclusions
Conclusions & Future work

- We’ve overviewed basic concepts related to Fuzzyness in Semantic Web Languages, such as
  - RDFS, OWL 2, Datalog
- Semantic Web Applications:
  - Robotics, Ontology Mappings, Multimedia Object annotation, Matchmaking, (Multimedia/Distributed) Information Retrieval, Recommender Systems, User Profiling, ...
Emerging Field for SWLs: Enhanced Fuzzy Queries

- Fuzzy Quantified queries may provide many opportunities to improve CQ query features for any SWL: e.g.

  Visualise roads in which many of the recent car incidents involved severely injured people

- Fuzzy quantified query schema:
  \[ Q \text{ of } B X \text{ are } A \]
  - \( Q \) is a fuzzy quantifier, e.g. many
  - \( B X \) is a reference fuzzy set over which \( Q \) quantifies, e.g. recent (\( B \)) car incidents (\( X \))
  - \( A \) is a fuzzy set imposing a condition to be satisfied, e.g. severely injured people

- Fuzzy Queries may be applied both to crisp ontologies as well as to fuzzy ontologies
That’s it !