Extending Datatype Restrictions in Fuzzy Description Logics

Fernando Bobillo[1] Umberto Straccia[2]

[1]Dept. of Computer Science and Systems Engineering, University of Zaragoza, Spain
fbobillo@unizar.es

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

www.straccia.info

Int. Conference on Intelligent Systems Design and Applications, 2009

Outline

Semantic Web

- In the last years the interest in ontologies has significantly grown
- An ontology is defined as an explicit and formal specification of a shared conceptualization
- Description Logics (DLs) are a family of logics that are the logical foundation of the standard W3C ontology language OWL [HPS04] (and OWL 2).

Fuzzy Logics and the Semantic Web

- It is widely agreed that "classical" ontology languages are not appropriate to deal with fuzzy/vague knowledge
- Fuzzy ontologies emerge as useful in several applications, such as multimedia information retrieval, image interpretation, ontology mapping, matchmaking and the Semantic Web [LS08]
- Several fuzzy extensions of DLs can be found in the literature (see the survey in [LS08])
- Some fuzzy DL reasoners have been implemented, such as FUZZYDL [BS08], DELOREAN [BDGR08] or FIRE [SSSK06].

Concrete Data Types

We deal with concrete data types, a major feature for fuzzy DLs

- Previous work allow simple datatype restrictions: *e.g.*,
  - $(\geqslant \mathit{hasAge}\ 18)$

    "people of age $\geqslant 18$"

- Extended data type restrictions: *e.g.*,
  - $\langle \mathit{john} : (= \mathit{hasAge}\ x)\rangle$ and $\langle \mathit{tom} : (= \mathit{hasAge}\ x)\rangle$

    "John and Tom have the same age"

  - $\langle \mathit{john} : (\geqslant \mathit{hasAge}\ \mathit{tri}(20, 40, 60))\rangle$

    "John's age is at least about 40 years, with about 40 being defined with a triangular function"

  - $\langle \mathit{john} : (= \mathit{hasGrossSalary}(0.8 \cdot \mathit{hasNetSalary} + 0.2 \cdot \mathit{paysTaxes})\rangle$

    "20% of John's gross salary is payed as taxes"

Basics of Fuzzy Logic [Háj98]

- Fuzzy statements: $\langle \phi, n \rangle$, where $n \in [0, 1]$ and ϕ is a statement
  - The degree of truth of ϕ is *at least n*
- Fuzzy interpretation: $\mathcal{I} : \mathit{Atoms} \to [0, 1]$ and is then extended inductively:

$$
\begin{array}{ll}
\mathcal{I}(\phi \wedge \psi) = \mathcal{I}(\phi) \otimes \mathcal{I}(\psi) & \mathcal{I}(\phi \vee \psi) = \mathcal{I}(\phi) \oplus \mathcal{I}(\psi), \\
\mathcal{I}(\phi \to \psi) = \mathcal{I}(\phi) \Rightarrow \mathcal{I}(\psi) & \mathcal{I}(\neg \phi) = \ominus \mathcal{I}(\phi), \\
\mathcal{I}(\exists x.\phi(x)) = \sup_{c \in \Delta^{\mathcal{I}}} \mathcal{I}(\phi(c)) & \mathcal{I}(\forall x.\phi(x)) = \inf_{c \in \Delta^{\mathcal{I}}} \mathcal{I}(\phi(c))
\end{array}
$$

$\otimes$, $\oplus$, $\Rightarrow$, and $\ominus$ are *truth combination functions*

	Łukasiewicz Logic	Gödel Logic	Product Logic	"Zadeh Logic"
$a \otimes b$	$\max(a + b - 1, 0)$	$\min(a, b)$	$a \cdot b$	$\min(a, b)$
$a \oplus b$	$\min(a + b, 1)$	$\max(a, b)$	$a + b - a \cdot b$	$\max(a, b)$
$a \Rightarrow b$	$\min(1 - a + b, 1)$	$\begin{cases} 1 & \text{if } a \leqslant b \\ b & \text{otherwise} \end{cases}$	$\min(1, b/a)$	$\max(1 - a, b)$
$\ominus a$	$1 - a$	$\begin{cases} 1 & \text{if } a = 0 \\ 0 & \text{otherwise} \end{cases}$	$\begin{cases} 1 & \text{if } a = 0 \\ 0 & \text{otherwise} \end{cases}$	$1 - a$

Entailment & Reasoning

- $\mathcal{I} \models \langle \phi, n \rangle$ iff $\mathcal{I}(\phi) \geqslant n$
- Best Entailment Degree (BED): $bed(\mathcal{K}, \phi) = \sup\{r \mid \mathcal{K} \models \langle \phi, r \rangle\}$
- BED can be computed as (where $\phi \leqslant x$ is $\langle \neg\phi, 1 - x \rangle$)

$$bed(\mathcal{K}, \phi) = \min x. \text{ such that } \mathcal{K} \cup \{\phi \leqslant x\} \text{ satisfiable}$$

- E.g., for Łukasiewicz logic, use Mixed Integer Linear Programming

$$\begin{array}{ll} bed(\mathcal{K}, \phi) = \min x. \text{ such that} & \\ & x \in [0,1], x_{\neg\phi} \geqslant 1 - x, \sigma(\neg\phi), \\ & \text{for all } \langle \phi', n \rangle \in \mathcal{K}, x_{\phi'} \geqslant n, \sigma(\phi'), \end{array}$$

$$\sigma(\phi) = \begin{cases} x_p \in [0,1] & \text{if} \quad \phi = p \\ x_{\phi'} = \ominus x_\phi, x_\phi \in [0,1] & \text{if} \quad \phi = \neg\phi' \\ \begin{array}{l} x_{\phi_1} \otimes x_{\phi_2} = x_\phi, \\ \sigma(\phi_1), \sigma(\phi_2), x_\phi \in [0,1] \end{array} & \text{if} \quad \phi = \phi_1 \wedge \phi_2 \\ x_{\phi_1} \oplus x_{\phi_2} = x_\phi & \text{if} \quad \phi = \phi_1 \vee \phi_2 \\ \sigma(\neg\phi_1 \vee \phi_2) & \text{if} \quad \phi = \phi_1 \rightarrow \phi_2 \, . \end{cases}$$

Basics of DLs

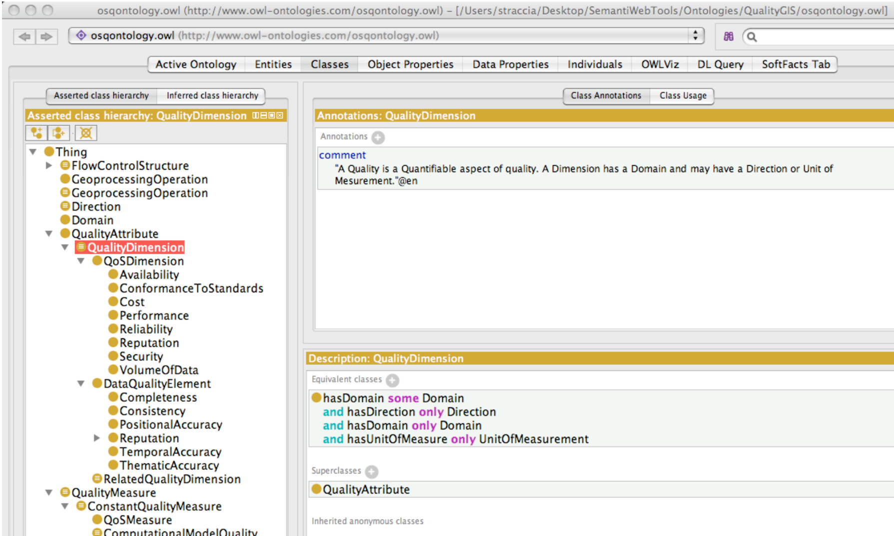
- The logics behind OWL 2, `http://dl.kr.org/`.
- Concept/Class: names are equivalent to unary predicates
  - In general, concepts equiv to formulae with one free variable
- Role or attribute: names are equivalent to binary predicates
  - In general, roles equiv to formulae with two free variables
- Taxonomy: Concept and role hierarchies can be expressed
- Individual: names are equivalent to constants

The Crisp 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	|	$A(x)$	$Human$
		$C \sqcap D$	|	$C(x) \wedge D(x)$	$Human \sqcap Male$
		$C \sqcup D$	|	$C(x) \vee D(x)$	$Nice \sqcup Rich$
		$\neg C$	|	$\neg C(x)$	$\neg Meat$
		$\exists R.C$	|	$\exists y.R(x, y) \wedge C(y)$	$\exists has\_child.Blond$
		$\forall R.C$		$\forall y.R(x, y) \rightarrow C(y)$	$\forall has\_child.Human$
$C \sqsubseteq D$				$\forall x.C(x) \rightarrow D(x)$	$Happy\_Father \sqsubseteq Man \sqcap \exists has\_child.Female$
$a{:}C$				$C(a)$	$John{:}Happy\_Father$

Example: GIS Quality Assessment Ontology [OWML08]



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, $(\geqslant n\ R)$ and $(\leqslant n\ R)$, *e.g.*, $(\geqslant 3\ has\_Child)$ (has at least 3 children)

$\mathcal{Q}$: Qualified number restrictions, $(\geqslant n\ R.C)$ and $(\leqslant n\ R.C)$, *e.g.*, $(\leqslant 2\ has\_Child.Adult)$ (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.*, $transitive(isPartOf)$

$\mathcal{R}$: role inclusions with composition, $R_1 \circ R_2 \sqsubseteq S$, *e.g.*, $isPartOf \circ isPartOf \sqsubseteq isPartOf$

For instance,

$$\mathcal{SHIF} = \mathcal{S} + \mathcal{H} + \mathcal{I} + \mathcal{F} = \mathcal{ALCR_+HIF} \quad \text{OWL-Lite}$$
$$\mathcal{SHOIN} = \mathcal{S} + \mathcal{H} + \mathcal{O} + \mathcal{I} + \mathcal{N} = \mathcal{ALCR_+HOIN} \quad \text{OWL-DL}$$
$$\mathcal{SROIQ} = \mathcal{S} + \mathcal{R} + \mathcal{O} + \mathcal{I} + \mathcal{Q} = \mathcal{ALCR_+ROIN} \quad \text{OWL 2}$$

Basics on Concrete Domains

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

  $(tim, 14):hasAge$
  $(sf, \text{``}SoftComputing\text{''}):hasAcronym$
  $(source1, \text{``}ComputerScience\text{''}):isAbout$
  $(service2, \text{``}InformationRetrievalTool''):Matches$
  $Minor = Person \sqcap \exists hasAge. \leqslant_{18}$

- Semantics: a clean separation between "object" classes and concrete domains
  - $D = \langle \Delta_D, \Phi_D \rangle$
  - Δ_D is an interpretation domain
  - Φ_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

Fuzzy DLs Basics

The semantics is an immediate consequence of applying mathematical fuzzy logic to the First-Order-Logic translation of DLs expressions

Interpretation:

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

$\otimes$ = t-norm
$\oplus$ = s-norm
$\ominus$ = negation
$\Rightarrow$ = implication

Concepts:

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

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

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

Inclusion axioms: $\langle C \sqsubseteq D, n\rangle$,

- $\mathcal{I} \models \langle C \sqsubseteq D, n\rangle$ iff $\inf_{x \in \Delta^{\mathcal{I}}} C^{\mathcal{I}}(x) \Rightarrow D^{\mathcal{I}}(x) \geqslant n$

Fuzzy DL: Specific Constructs

- Concrete data types
  - *e.g.*, $Sedan \sqcap (\geqslant price\ 22.000)$
- Fuzzy constraints
  - numerical features may be constrained by so-called fuzzy membership functions

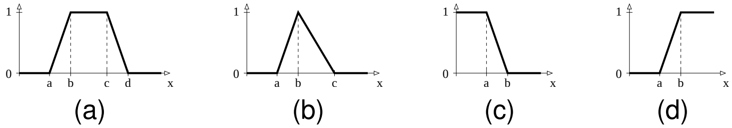


Figure: (a) Trapezoidal function $trz(a, b, c, d)$, (b) triangular function $tri(a, b, c)$, (c) left shoulder function $ls(a, b)$, and (d) right shoulder function $rs(a, b)$.

- For instance, *item*4's price is about 24000

$$item4 : \exists price.tri(22000, 24000, 26000)$$

Definition (Specific Concept Expressions)

$$\begin{array}{lcl} C & \rightarrow & \forall t.d \mid \exists t.d \quad \text{(fuzzy constraints)} \\ d & \rightarrow & ls(a,b) \mid rs(a,b) \mid tri(a,b,c) \mid trz(a,b,c,d) \end{array}$$

e.g.

$$Car \sqcap \exists price.tri(22000, 24000, 26000)$$

$$\begin{array}{lcl} C & \rightarrow & TC \quad \text{(threshold concept)} \\ TC & \rightarrow & C[\geqslant n] \mid C[\leqslant n] \end{array}$$

e.g. $(Sedan \sqcap Cheap \sqcap (\leqslant price\ 30.000))[\geqslant 0.8]$

$$\begin{array}{lcl} C & \rightarrow & WC \quad \text{(weighted sum concept)} \\ WC & \rightarrow & (w_1 \cdot C_1 + w_2 \cdot C_2 + \ldots + w_k \cdot C_k) \end{array}$$

e.g.,

$$NiceHotel \doteq 0.3 \cdot CheapHotel + 0.7 \cdot ConfortableHotel$$

$$C \rightarrow mod(C) \quad \text{(modified concept)}$$

where *mod* is a linear hedge. E.g.,

$$SportCar \sqsubseteq Car \sqcap \exists \text{hasSpeed}.very(High)$$

Language

$$
\begin{array}{lcl}
C & \rightarrow & \forall t.d \mid \exists t.d \quad \text{(fuzzy constraints)} \\
d & \rightarrow & ls(a,b) \mid rs(a,b) \mid tri(a,b,c) \mid trz(a,b,c,d)
\end{array}
$$

$$
\begin{array}{lcl}
C & \rightarrow & DR \quad \text{(datatype restriction)} \\
DR & \rightarrow & (\geqslant t\ val) \mid (\leqslant t\ val) \mid (= t\ val) \\
val & \rightarrow & string \mid rational \mid FN \mid AE \\
FN & \rightarrow & rational \mid fuzzynumber \mid FN_1 \star FN_2 \quad \star \in \{+,-,\cdot,\div\} \\
AE & \rightarrow & rational \mid t \mid n \cdot t \mid AE_1 + AE_2
\end{array}
$$

e.g.

$$audi234 : Sedan \sqcap (\leqslant price\ 26000)$$

$$
\begin{array}{l}
SoldItem \sqsubseteq (= totalPrice\ netprice + VAT) \\
SoldItem \sqsubseteq (= VAT\ 0.2 \cdot netprice)
\end{array}
$$

Reasoning Method

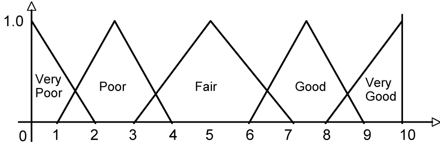
Depends on the semantics of the operators

- MILP based method: under Zadeh semantics, Łukasiewicz semantics, and classical semantics
  - Use of Logical Inference + Mixed integer linear Programming
- MIQP based method: under Product T-norm, use of
  - Use of Logical Inference + Mixed Integer Quadratically Constrained Programming optimization problem (MICQP)

Multi-Criteria Decision Making

- Assume that we have to chose among three offers for a GIS system that have been evaluated according to
  - Criteria: Cost, Delivery Time and Quality
- Assume the decision matrix and the definition of the vague performance values are

Offer	Cost 0.258	DeliveryTime 0.105	Quality 0.637
a_1	VeryPoor	Fair	Good
a_2	Good	VeryGood	Poor
a_3	Fair	Fair	Poor



- Fuzzy DL encoding:

$$\mathrm{VeryPoor} \doteq \mathrm{ls}(0,2),\quad \mathrm{Poor} \doteq \mathrm{tri}(1,2.5,4),\quad \mathrm{Fair} \doteq \mathrm{tri}(3,5,7),\quad \mathrm{Good} \doteq \mathrm{tri}(6,7.5,9),\quad \mathrm{VeryGood} \doteq \mathrm{rs}(8,10)$$

$$\mathrm{a}_1 : \mathrm{Alternative} \sqcap \exists \mathrm{hasCost}.\mathrm{VeryPoor} \sqcap \exists \mathrm{hasDeliveryTime}.\mathrm{Fair} \sqcap \exists \mathrm{hasQuality}.\mathrm{Good}$$
$$\mathrm{a}_2 : \mathrm{Alternative} \sqcap \exists \mathrm{hasCost}.\mathrm{Good} \sqcap \exists \mathrm{hasDeliveryTime}.\mathrm{VeryGood} \sqcap \exists \mathrm{hasQuality}.\mathrm{Poor}$$
$$\mathrm{a}_3 : \mathrm{Alternative} \sqcap \exists \mathrm{hasCost}.\mathrm{Fair} \sqcap \exists \mathrm{hasDeliveryTime}.\mathrm{Fair} \sqcap \exists \mathrm{hasQuality}.\mathrm{Poor}$$

$$\mathrm{Alternative} \doteq (= \mathrm{hasRankValue}\ 0.258 \cdot \mathrm{hasCost} + 0.105 \cdot \mathrm{hasDeliveryTime} + 0.637 \cdot \mathrm{hasQuality})$$

- Final Rank Value: $\mathrm{rank}(\mathcal{K}, \mathrm{a_i}) = \mathit{mom}(\mathcal{K}, \mathrm{Alternative}, \mathrm{a_i}, \mathrm{hasRankValue})$

$$\mathrm{rank}(\mathcal{K}, \mathrm{a}_1) = 5.301 \qquad \mathrm{rank}(\mathcal{K}, \mathrm{a}_2) = 4.577 \qquad \mathrm{rank}(\mathcal{K}, \mathrm{a}_3) = 3.408$$

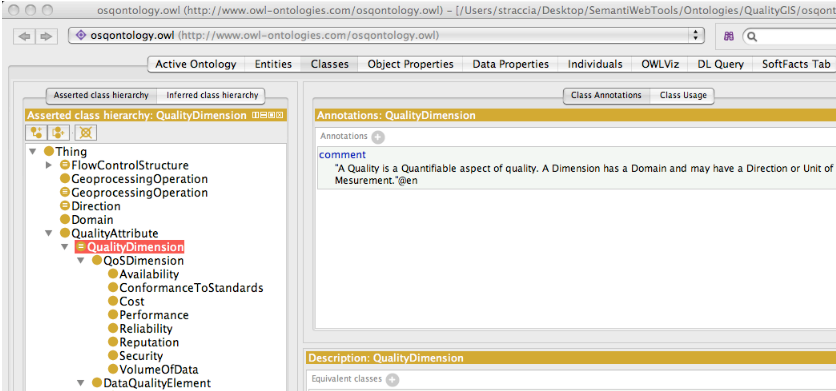
$$\mathrm{a}^* = \arg\max_{\mathrm{a_i}} \mathrm{rank}(\mathcal{K}, \mathrm{a_i}) = \mathrm{a}_1$$

- Encoding nicely extends if background knowledge is involved such as, *e.g.*,
  - Criteria taxonomy

$$\text{Consistency} \sqsubseteq \text{DataQualityElement}$$

  - Properties of alternatives, *e.g.*,

$$a_1 : \text{Alternative} \sqcap \exists \text{hasSecurity.VeryPoor}$$



Contribution & Future Work

- We have added extended data type restrictions to fuzzy DLs
- FUZZYDL reasoner supports the language proposed here
- A major task is:
  - How to encode fuzzy DLs into OWL 2 ?
- Integrate other features from the fuzzy literature

For Further Reading I

Fernando Bobillo, Miguel Delgado, and Juan Gómez-Romero.
Delorean: A reasoner for fuzzy OWL 1.1.
In *Proceedings of the 4th International Workshop on Uncertainty Reasoning for the Semantic Web (URSW 2008)*, volume 423. CEUR Workshop Proceedings, 10 2008.

Fernando Bobillo and Umberto Straccia.
fuzzyDL: An expressive fuzzy description logic reasoner.
In *2008 International Conference on Fuzzy Systems (FUZZ-08)*, pages 923–930. IEEE Computer Society, 2008.

Petr Hájek.
Metamathematics of Fuzzy Logic.
Kluwer, 1998.

Ian Horrocks and Peter Patel-Schneider.
Reducing OWL entailment to description logic satisfiability.
Journal of Web Semantics, 2004.

Carsten Lutz and Maja Milicic.
A tableau algorithm for description logics with concrete domains and general tboxes.
J. Autom. Reasoning, 38(1-3):227–259, 2007.

Thomas Lukasiewicz and Umberto Straccia.
Managing uncertainty and vagueness in description logics for the semantic web.
Journal of Web Semantics, 6:291–308, 2008.

For Further Reading II

Richard Onchaga, Ing Widya, Javier Morales, and J. M. Lambert.
An ontology framework for quality of geographical information services.
In *GIS '08: Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*, pages 1–4, New York, NY, USA, 2008. ACM.

Giorgos Stoilos, Nikolaos Simou, Giorgos Stamou, and Stefanos Kollias.
Uncertainty and the semantic web.
IEEE Intelligent Systems, 21(5):84–87, 2006.

Umberto Straccia.
Description logics with fuzzy concrete domains.
In Fahiem Bachus and Tommi Jaakkola, editors, *21st Conference on Uncertainty in Artificial Intelligence (UAI-05)*, pages 559–567, Edinburgh, Scotland, 2005. AUAI Press.