

Description Logics as Ontology Languages for the Semantic Web

Franz Baader, Ian Horrocks, Ulrike Sattler

Beschreibungslogik
Martin Hofmann, Hans Leiß
Hauptseminar, Sommersemester 2008
CIS + Institut für Informatik, LMU München

The semantic Web and Ontologies

Motivation:

Automatisch erzeugte Suchergebnisse aus dem Web sind unbefriedigend.

Grund:

Informationen im Web sind prinzipiell auf menschliche Verarbeitung ausgelegt. Maschinell zu verarbeitende Information ist nur auf ein Set von Schlüsselwörtern begrenzt.

Ziel:

Ressourcen im Web sollen von menschlichen und maschinellen Agenten gleichermaßen gefunden und verstanden werden können.

Methode:

Verwendung eines Semantic Mark-Up bei den Web-Ressourcen und Ontologien, mittels denen eine einheitliche Terminologie für die Agenten festgelegt wird.

The semantic Web and Ontologies

Ontology Requirements:

Intuitive Syntax

Menschen- und maschinenverständlich

Kompatibel mit existierenden Web-Standards (XML, RDF, RDFS)

Interoperabilität: Verwendung einer allgemeinen Ontologie bzw.
Übersetzbarkeit

Formale Definition der Semantik

Ausreichende Ausdrucksstärke

Entscheidbarkeit

Reasoning für die Entwicklungsphase (Konsistenzprüfung), die Integration
(Ableitung und Test von Beziehungen über versch. Ontologien) und nach
Auslieferung (Ableitung von geschlossenen Beziehungen zwischen Instanzen)

Description Logics

Ontology Requirements:

DLs sind eine Familie von Wissensrepräsentationssprachen

Sie repräsentieren Wissen begrenzter Anwendungsbereiche in einer strukturierten, formal definierten Weise

Sie unterscheiden sich von ihren Vorgängern (Semantische Netze und Frames) durch eine logikbasierte, formale Semantik.

Sie bilden Beschreibungskonzepte aus unter der Verwendung von atomaren Konzepten und atomaren Rollen.

Konzept „Glücklicher Mann“: Ein Mann, der mit einer Doktorin verheiratet ist und mindestens 5 Kinder hat, die alle Professoren sind.

HappyMan $\equiv \neg\text{Female} \sqcap \exists\text{married.Doctor} \sqcap (\geq 5\text{hasChild}) \sqcap \forall\text{hasChild.Professor}$

Description Logics

Atomare Konzepte:

Female

Atomare Rollen:

hasWife.Female

Komplexe Konzepte:

HappyMan $\equiv \neg$ Female \sqcap \exists married.Doctor \sqcap (≥ 5 hasChild) \sqcap \forall hasChild.Professor

Constraints:

\exists hasChild.Human \sqsubseteq Human

Assertions:

HappyMan(BOB), hasChild(BOB,MARY)

The expressive DL *SHIQ* - Definition 1

Syntax von *SHIQ* Rollen

Primitive Rollen

Transitive Rollen

Deren inverse Rollen

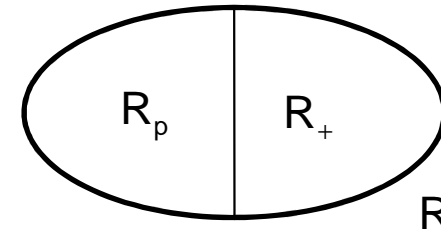
Somit ist die Gesamtheit aller *SHIQ* Rollen:

$$R \cup \{r^- \mid r \in R\}$$

Ein Rolleninklusionsaxiom ist von der Form:

$r \sqsubseteq s$, wobei r und s *SHIQ* Rollen sind.

Eine Rollenhierarchie \mathcal{R} ist ein endliches Set von Rolleninklusionsaxiomen.



The expressive DL *SHIQ* - Definition 1

Semantik von *SHIQ* Rollen: Interpretationen

Eine Interpretation $I = (\Delta^I, \bullet^I)$ besteht aus:

einem Individuenbereich Δ^I und
einer Funktion \bullet^I .

Diese Funktion bildet jede Rolle auf eine Untermenge von $\Delta^I \times \Delta^I$ ab,
so dass für jedes $p \in R$ und $r \in R_+$ gilt:

$$\langle x, y \rangle \in p^I \text{ iff } \langle y, x \rangle \in (p^-)^I$$
$$\text{if } \langle x, y \rangle \in r^I \text{ and } \langle x, z \rangle \in r^I \text{ then } \langle x, z \rangle \in r^I.$$

Eine Interpretation I erfüllt eine Rollenhierarchie R , wenn gilt:

$$r^I \subseteq s^I \text{ für jedes } r \sqsubseteq s \in R.$$

So eine Interpretation nennen wir ein Modell von R .

The expressive DL *SHIQ* - Definition 1

Semantik von *SHIQ* Rollen: Weitere Definitionen

1. Rückgabe der ursprünglichen Relation bei inversen Rollen:

$$\text{Inv}(r) := \begin{cases} r^- & \text{wenn } r \text{ ein Rollenname ist} \\ s & \text{wenn } r = s^- \text{ für einen Rollennamen } s. \end{cases}$$

2. Da eine Rolleninklusion zu ihrer Inversion führt, impliziert eine Rollenhierarchie R stets zusätzliche Inklusionsbeziehungen. Da Transitivität gilt, definieren wir \equiv^* als reflexiv transitive Hülle aus:

$$\equiv_{\mathcal{R}} := R \cup \{ \text{Inv}(r) \equiv \text{Inv}(s) \mid r \equiv s \in R \}$$

Es gilt: $r \equiv_{\mathcal{R}} s$ als Abkürzung für $r \equiv^*_{\mathcal{R}} s$ und $s \equiv^*_{\mathcal{R}} r$

3. Eine binäre Relation ist transitiv, wenn ihre Inversion transitiv ist.

Wenn also $r \equiv_{\mathcal{R}} s$ und r oder $\text{Inv}(r)$ transitiv ist, interpretiert jedes Modell \mathcal{S} als transitive Relation. Somit definieren wir:

$$\text{Trans}(\mathcal{S}, \mathcal{R}) := \begin{cases} \text{true} & \text{wenn } r \in R_+ \text{ oder } \text{Inv}(r) \in R_+ \text{ für ein } r \text{ mit } r \equiv_{\mathcal{R}} s \\ \text{false} & \text{ansonsten.} \end{cases}$$

4. Eine Rolle ist primitiv in \mathcal{R} , wenn $\text{Trans}(\mathcal{S}, \mathcal{R}) = \text{false}$ für alle $s \equiv^*_{\mathcal{R}} r$

The expressive DL *SHIQ* - Definition 2

Syntax von *SHIQ* Konzepten

Sei N_C eine Menge von Konzeptnamen. Die Menge der *SHIQ* Konzepte ist die kleinste Menge, für die gilt:

- (1) Jeder Konzeptname $A \in N_C$ ist ein *SHIQ* Konzept.
- (2) Sind C und D *SHIQ* Konzepte und r eine *SHIQ* Rolle, so sind auch $C \sqcap D$, $C \sqcup D$, $\neg C$, $\forall r.C$, $\exists r.C$ *SHIQ* Konzepte.
- (3) Ist C ein *SHIQ* Konzept, r eine primitive *SHIQ* Rolle und $n \in \mathbb{N}$, so sind auch $(\leq n r.C)$ und $(\geq n r.C)$ *SHIQ* Konzepte.

The expressive DL *SHIQ* - Definition 2

Semantik von *SHIQ* Konzepten: Interpretationen

Die Interpretationsfunktion \bullet^I einer Interpretation $I = (\Delta^I, \bullet^I)$ bildet zusätzlich jedes Konzept zu einer Untermenge von Δ^I ab, so das gilt:

$$(C \sqcap D)^I = (C^I \cap D^I), \quad (C \sqcup D)^I = (C^I \cup D^I), \quad \neg C^I = \Delta^I \setminus C^I,$$

$$(\exists r.C)^I = \{x \in \Delta^I \mid \text{es gibt mindestens ein } y \in \Delta^I \text{ mit } \langle x, y \rangle \in r^I \text{ und } y \in C^I\},$$

$$(\forall r.C)^I = \{x \in \Delta^I \mid \text{für alle } y \in \Delta^I, \text{ wenn } \langle x, y \rangle \in r^I, \text{ dann } y \in C^I\},$$

$$(\leq n r.C)^I = \{x \in \Delta^I \mid \# r^I(x, C) \leq n\},$$

$$(\geq n r.C)^I = \{x \in \Delta^I \mid \# r^I(x, C) \geq n\},$$

wobei $\#M$ die Kardinalität der Menge M bezeichnet,

und $r^I(x, C) := \{y \mid \langle x, y \rangle \in r^I \text{ und } y \in C^I\}$.

Gilt $x \in C^I$, sagen wir, das x eine Instanz von C in I ist,

und wenn $\langle x, y \rangle \in r^I$, dann nennen wir y einen r -successor von x in I .

The expressive DL *SHIQ* - Definition 3 and 4

Definition 3:

Eine „general Concept Inclusion“ (GCI) ist von der Form $C \sqsubseteq D$, wobei es sich bei C, D um *SHIQ* Konzepte handelt. Eine endliche Menge GCIs nennen wir eine TBox. Eine Interpretation I ist ein Modell einer TBox T wenn es alle GCIs in T erfüllt.

Definition 4:

Das Konzept C ist erfüllbar in bezug auf die Rollenhierarchie \mathcal{R} und einer TBox wenn es ein Modell I von \mathcal{R} und T gibt, mit $C^I \neq \emptyset$.

Eine solche Interpretation nennen wir Modell von C bezüglich \mathcal{R} und T .

Das Konzept D subsumiert das Konzept C mit Bezug auf $\langle \mathcal{R}, T \rangle$, wenn für alle Modelle I von \mathcal{R} und T gilt: $C^I \subseteq D^I$

Zwei Konzepte C, D sind äquivalent bezüglich \mathcal{R} , wenn sie sich gegenseitig subsumieren.

DAML+OIL - Constructors

Constructor	DL Syntax	Beispiele
intersectionOf	$C_1 \sqcap \dots \sqcap C_n$	Human \sqcap Male
unionOf	$C_1 \sqcup \dots \sqcup C_n$	Doctor \sqcup Lawyer
complementOf	$\neg C$	\neg Male
oneOf	$\{x_1 \dots x_n\}$	{john,mary}
toClass	$\forall P.C$	\forall hasChild.Doctor
hasClass	$\exists r.C$	\exists hasChild.Lawyer
hasValue	$\exists r.\{x\}$	\exists citizenOf.{USA}
minCardinalityQ	$(\geq n r.C)$	$(\geq 2$ hasChild.Lawyer)
maxCardinalityQ	$(\leq n r.C)$	$(\leq 1$ hasChild.Male)
inverseOf	r^-	hasChild $^-$

DAML+OIL - Axioms

Axiom	DL Syntax	Beispiele
subClassOf	$C_1 \sqsubseteq C_2$	Human \sqsubseteq Animal \sqcap Biped
sameClassAs	$C_1 \equiv C_2$	Man \equiv Human \sqcap Male
subPropertyOf	$P_1 \sqsubseteq P_2$	hasDaughter \sqsubseteq hasChild
samePropertyAs	$P_1 \equiv P_2$	cost \equiv price
disjointWith	$C_1 \sqsubseteq \neg C_2$	Male $\sqsubseteq \neg$ Female
sameIndividualAs	$\{x_1\} \equiv \{x_2\}$	{President_Bush} \equiv {G_W_Bush}
differentIndividualFrom	$\{x_1\} \sqsubseteq \neg \{x_2\}$	{john} $\sqsubseteq \neg$ {peter}
transitiveProperty	$P \in R_+$	hasAncestor ⁺ $\in R_+$
uniqueProperty	$T \sqsubseteq (\leq 1 P.T)$	T $\sqsubseteq (\leq 1$ hasMother.T)
unambiguousProperty	$T \sqsubseteq (\leq 1 P^- .T)$	T $\sqsubseteq (\geq 1$ isMotherOf ⁻ .T)

Reasoning in *SHIQ*

Schwierigkeit 1: Qualified Number Restriction

Durch die Einführung der Konzepte $(\leq nr.C)$ und $(\geq nr.C)$ erhöht sich die Komplexität des Entscheidungsproblems. War EXPTIME zunächst ein Worst-Case-Fall, wird es bei *SHIQ* nun im Gegensatz zu weniger ausdrucksstarken DLs zum Standard.

Schwierigkeit 2: Termination

Durch die Existenz reflexiver, transitiver Rollen wird die Terminierung des Algorithmus zu einer nicht trivialen Angelegenheit. Dies wird durch einen Mechanismus namens Blocking gelöst, durch den Äste, die zu infiniten Bäumen führen, erkannt und geblockt werden.

Beispiele: Rollen (zu Folie 9,10,11)

Rollen, Semantik:

r: „ist ein Kind von“ $\in R \langle \text{Kind}, \text{Elternteil} \rangle$
s: „ist verwandt mit“ $\in R^+ \langle \text{Verwandter}, \text{Verwandter} \rangle$
t: „ist verheiratet mit“ $\in R \langle \text{Ehepartner}, \text{Ehepartner} \rangle$
u: „ist Sohn von“ $\in R \langle \text{Sohn}, \text{Elternteil} \rangle$
v: „ist Vater von“ $\in R \langle \text{Vater}, \text{Kind} \rangle$
w: „ist Geschwister von“ $\in R^+ \langle \text{Kind}, \text{Kind} \rangle$

Rolleninklusion:

$r \sqsubseteq s$
 $t \sqsubseteq s$
 $u \sqsubseteq r$
 $v \sqsubseteq s$
 $w \sqsubseteq s$

Beispiele: Konzepte (zu Folie 6,7,8)

Konzepte:

$N_C = \{\text{Kind, Verwandter, Elternteil, Sohn, Vater}\}$

GCI: General Concept Inclusions:

Kind \sqsubseteq Verwandter

Ehepartner \sqsubseteq Verwandter

Sohn \sqsubseteq Kind

Vater \sqsubseteq Verwandter

Interpretation und Modell: Die Familie Simpson:

Kind := {Bart, Lisa, Maggie}

Verwandter := {Homer, March, Bart, Lisa, Maggie}

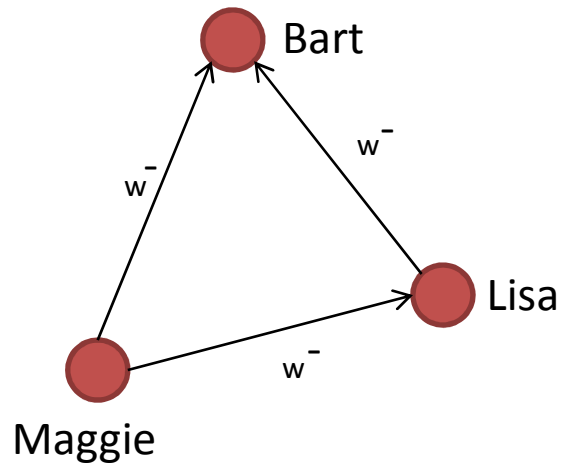
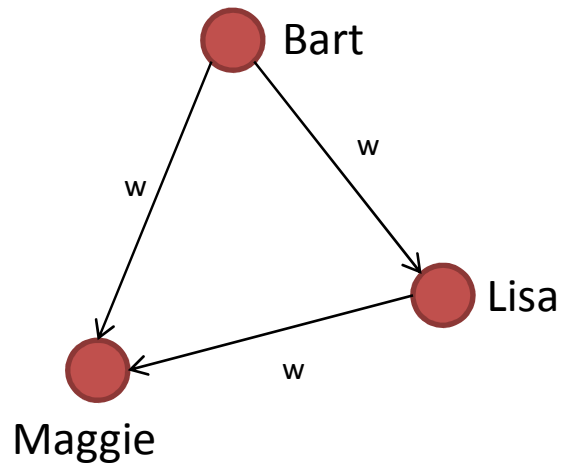
Ehepartner := {Homer, March}

Sohn := {Bart}

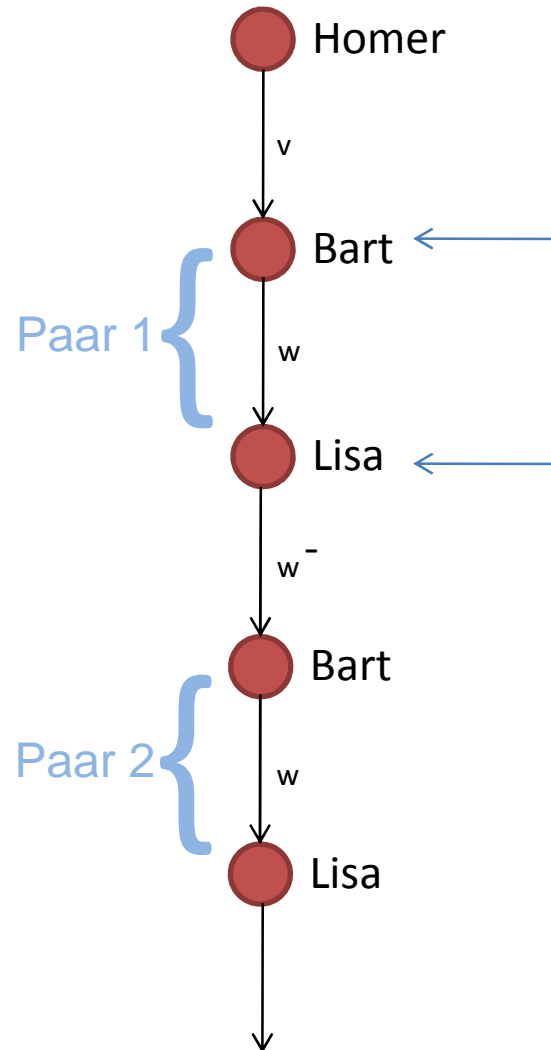
Vater := {Homer}

Die Simpsons als infiniter Baum

Transitiv reflexive Rollen



Infiniter Pfad



Abbruchkriterium

Zwei Paare Knoten
 mir paarweise
 identischen Labels
 verbunden durch
 Kanten mit ebenfalls
 den gleichen Labels
 führen zu einem
 Block des Pfades.

Conclusion

Die mit Nachdruck betriebenen Forschungen bezüglich Logik-basierenden Semantiken in Zusammenarbeit mit hochoptimierten Systemen für ausdrucksstarke DLs sowie die Verfügbarkeit von Reasoning-Algorithmen machen diese Familie von Wissensrepräsentationssystem zu einem idealen Ausgangspunkt für das semantische Web.

Ein wichtiger Punkt wird der Entwurf, die Entwicklung, die Integration und die Wartung von Ontologien. Hierfür gibt es Bedarf an Werkzeugunterstützung um auch Nicht-DL-Experten die Entwicklung von Ontologien zu ermöglichen.