

OWL (odvozování)



ISKM89 Organizace dat - sémantický web | podzim 2023
Zuzana Nevěřilová | Centrum zpracování přirozeného jazyka

RDF(S) a úrovně OWL

RDF/RDFS

OWL je vždy OWL 2 (i když OWL 1 existovalo)

OWL Lite - minimální užitečná expresivita a “síla” inference

OWL DL - deskripční logika

OWL Full - bez jakýchkoliv omezení (obsahuje i RDFS)

OWL Lite, OWL EL, OWL RL, OWL QL \subseteq OWL2 DL \subseteq OWL2 Full

https://en.wikipedia.org/wiki/Description_logic
<https://www.w3.org/TR/owl-ref/#Sublanguage-def>

RDF(S) a úrovně OWL

OWL Full - V OWL Full jsou třídy (také) individua, tudíž na ně lze aplikovat rovnost tříd (`owl:sameAs`).

Protože datové hodnoty jsou (také) individua, jsou datatype properties podtřídami object properties.

`owl:ObjectProperty = rdf:Property`

RDF je OWL Full (bez jakýchkoliv omezení).

RDF(S) a úrovně OWL

OWL DL - musí být odděleny třídy, instance, datové typy, datatype properties, object properties, annotation properties, ontology properties, datové hodnoty (data values).

Vlastnosti

- inverse of,
- inverse functional,
- symmetric, a
- transitive

nikdy nejsou datatype properties.

Kardinalita nesmí omezovat tranzitivní vlastnosti.

Všechny třídy a vlastnosti musí být explicitně deklarovány (buď jako třídy, nebo jako vlastnosti).

RDF(S) a úrovně OWL

OWL Lite -

navíc k omezením OWL DL přidává následující:

Neobsahuje vlastnosti:

- [owl:oneOf](#)
- [owl:unionOf](#)
- [owl:complementOf](#)
- [owl:hasValue](#)
- [owl:disjointWith](#)

Objektem trojice s `rdfs:domain` a `rdfs:range` musí být třídy a třídy nebo datové typy

Objektem trojice `rdf:type` musí být třída (i anonymní).

Objektem trojice `owl:allValuesFrom` a `owl:someValuesFrom` musí být třída nebo datový typ.

Úložiště trojic a inferenční stroje

Úložiště trojic často obsahuje inferenční stroj. Jsou to ale rozdílné věci.

Cíle úložiště:

- ukládat trojice (v některé ze serializací RDF(S)/OWL)
- poskytovat trojice
 - jako RDF (v některé ze serializací)
 - SPARQL - projekce (tabulka) nebo podgraf
- + inferenční stroj: poskytovat odvozené trojice
 - vygenerovat je při dotazu (on the fly)
 - dopředu (materializované trojice)

Úložiště trojic a inferenční stroje

- AllegroGraph
- AnzoGraph
- Apache Jena + Fuseki + TDB
- Apache Marmotta + Postgres
- ArangoDb
- Blazegraph
- Cosmos
- GraphDB
- MarkLogic
- Neo4j
- Neptune
- Stardog
- Tiger
- Virtuoso

Rozdíly:

- Škálovatelnost (distributed system, multiple pods)
- Způsob uložení dat (native triple store, relační databáze, grafová databáze)
- Inference (RDF, OWL DL 2, vlastní inferenční pravidla)
- Konektory k jiným nástrojům (ElasticSearch, Lucene, SOLR)
- Spuštění v cloudu nebo on premises (AWS, Azure, GCP, other)
- Podpora (logování, vizualizace, import, export)
- Cena

Úložiště trojic Inferenční stroje

RDF Store	Features											
	SPARQLv.	Inference	3/4-uple	Spatial search	Full-text search	Storage	Temporal search	Size (million/billion triples)	License	Dev. Language	Cluster support	Active project
Virtuoso 7.2.4 OS	1.1	RDFS+	4	Adv	Y	RDBMS	N	50BT	OS	C	N	Y
Virtuoso 7.2.4 Comm	1.1	RDFS+	4	Adv	Y	RDBMS	N	50BT	Cm	C	H	Y
Graph DB SE 7.0.1	1.1	OWL2RL	4	Bas	Y	custom	N	10BT	Cm	Java	H	Y
Stardog 4	1.1	OWL2	4	Adv	Y	custom	N	10BT	Cm	Java	H	Y
Oracle 12c	1.1	RDFS, OWL2	4	Adv	Y	custom	N	1TT	Cm	C/Java	H	Y
Apache Jena-Fuseki	1.1	RDFS OWL-Lite	3	Bas	Y	custom (TDB)	N	1.7BT	OS	Java	N	Y
Apache Jena-Fuseki	1.1	No	4	Bas	Y	custom (TDB)	N	1.7BT	OS	Java	N	Y
Blazegraph 2.1.2	1.1	RDFS+	3	Bas	Y	custom	N	50BT	OS	Java	V&H	Y
Blazegraph 2.1.2	1.1	No	4	Bas	Y	custom	N	50BT	OS	Java	V&H	Y
CumulusRDF	1.1	No	3	No	N	Cassandra 1.2	N	120MT	OS	Java	V	(Y)
Strabon	1.1	No	3	Adv	N	RDBMS	Y	500MT	OS	Java	N	(Y)
4store	1.1	No	4	No	N	custom	N	15BT	OS	C	V	N
h2rdf+	1.0	No	3	No	N	HBase	N	2.7BT	OS	Java	H&V	N

Dedukce a RDF - kde najdeme sylogismus o Sókratovi?

Všichni lidé jsou smrtelní.

Sókratés je člověk.

Sókratés je smrtelný.

backward-chaining

odvození na dotaz Je Sókratés smrtelný?

forward-chaining

všechna možná odvození se předpočítají

(total materialization)