Assignment 2 (Partial Solutions)

Camilo Thorne

Room 00.012
Institut für Maschinelle Sprachverarbeitung
Universität Stuttgart
+49 (0) 711 685-81369
camilo.thorne@ims.uni-stuttgart.de

Semantic Web, SS 2017

Institut für Maschinelle Sprachverarbeitung

Outline

Institut für Maschinelle Sprachverarbeitung

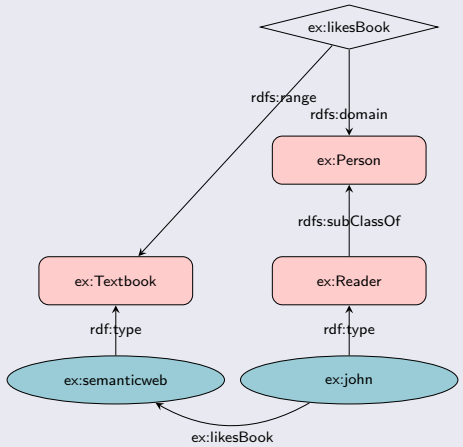
Outline

Institut für Maschinelle Sprachverarbeitung

RDFS Models

Models and Entailment

Show that `ex:machinelearning ex:likes ex:semanticweb.` is not entailed by RDFS ontology $\mathcal{O}_b$



by constructing **two** set-theoretical models $\mathcal{I}_1$ and $\mathcal{I}_2$ of $\mathcal{O}_b$, such that

- `ex:machinelearning ex:likes ex:semanticweb.` is true in $\mathcal{I}_1$
- `ex:machinelearning ex:likes ex:semanticweb.` is false in $\mathcal{I}_2$

Institut für Maschinelle Sprachverarbeitung

RDFS Models (Answer) I

- Two steps: **(1)** define $^{\mathcal{I}}$: define D and $\cdot^{\mathcal{I}}$ over the nodes of $\mathcal{O}_b$ and check if $^{\mathcal{I}}$ verifies the edges of $\mathcal{O}_b$; **(2)** Check if it is a model of the additional triple `ex:machinelearning ex:likes ex:semanticweb.`

  1. $\mathcal{I}_1$:

$$
\begin{aligned}
&D_1 = \{john, semantic\_web, machine\_learning\},\\
&\texttt{ex:john}^{\mathcal{I}_1} = john,\\
&\texttt{ex:semanticweb}^{\mathcal{I}_1} = semantic\_web,\\
&\texttt{ex:machinelearning}^{\mathcal{I}_1} = machine\_learning,\\
&\texttt{ex:Reader}^{\mathcal{I}_1} = \{john\},\\
&\texttt{ex:Person}^{\mathcal{I}_1} = \{john\},\\
&\texttt{es:Textbook}^{\mathcal{I}_1} = \{semantic\_web\},\\
&\texttt{ex:likesBook}^{\mathcal{I}_1} = \{(john, semantic\_web)\},\\
&\texttt{ex:likes}^{\mathcal{I}_1} = \{(machine\_learning, semantic\_web)\}
\end{aligned}
$$

$\Rightarrow I_1$ is a model of $\mathcal{O}_b$ and a model of the triple.

Institut für Maschinelle Sprachverarbeitung

RDFS Models (Answer) II

2. $\mathcal{I}_2$:

$$D_1 = \{john, semantic\_web, machine\_learning\},$$
$$\texttt{ex:john}^{\mathcal{I}_2} = john,$$
$$\texttt{ex:semanticweb}^{\mathcal{I}_2} = semantic\_web,$$
$$\texttt{ex:machinelearning}^{\mathcal{I}_2} = machine\_learning,$$
$$\texttt{ex:Reader}^{\mathcal{I}_2} = \{john\},$$
$$\texttt{ex:Person}^{\mathcal{I}_2} = \{john\},$$
$$\texttt{es:Textbook}^{\mathcal{I}_2} = \{semantic\_web\},$$
$$\texttt{ex:likesBook}^{\mathcal{I}_2} = \{(john, semantic\_web)\},$$
$$\texttt{ex:likes}^{\mathcal{I}_2} = \{\}$$

$\Rightarrow I_2$ is a model of $\mathcal{O}_b$ but is not a model of the triple.

- For $\mathcal{O}_b$ to entail `ex:machinelearning ex:likes ex:semanticweb.`, all the models of $\mathcal{O}_b$ must be models of the triple as well, but $\mathcal{I}_2$ doesn't satisfy this property; we therefore conclude that $\mathcal{O}_b$ does not entail the triple.

Institut für Maschinelle Sprachverarbeitung

RDFS Rules

RDFS Rule Soundness

Show that the following RDFS rules seen in course are **sound**:

1. RDFS range rule
2. RDFS domain rule
3. RDFS subclass transitivity rule
4. RDFS subproperty rule

Institut für Maschinelle Sprachverarbeitung

RDFS Rules (Answer)

I'll solve only case 3., all the others being analogous.

The subclass rule transitivity is sound if for all ontologies $\mathcal{O}$ containing triples of the form

(1) `A rdfs:subclassOf B.` and (2) `B rdfs:subclassOf C.`

$\mathcal{O}$ entails

(3) `B rdfs:subclassOf C.`

This will hold when **every** model $\mathcal{I}$ of $\mathcal{O}$ and (1) and (2) is a model of (3) as well. Let $\mathcal{I}$ be an arbitrary model of (1) and (2). If such is the case, by definition of $\mathcal{I}$

(1.1) $\texttt{A}^{\mathcal{I}} \subseteq \texttt{B}^{\mathcal{I}}$ and (2.1) $\texttt{B}^{\mathcal{I}} \subseteq \texttt{C}^{\mathcal{I}}$.

As the $\subseteq$ relation among sets is transitive, (1.1) and (2.1) imply that $\texttt{A}^{\mathcal{I}} \subseteq \texttt{C}^{\mathcal{I}}$, i.e., that $\mathcal{I}$ is a model of (3). Since $\mathcal{I}$ was arbitrary, this property holds for **every** model of $\mathcal{O}$. Whence, by definition of entailment, it follows that (1) and (2) entail (3), and that the rule is sound.

Institut für Maschinelle Sprachverarbeitung

Outline

Institut für Maschinelle Sprachverarbeitung

SPARQL (Main Features)

- SPARQL is the formal query language used to query RDF knowledge bases/triplestores
- It syntax and semantics mirrors that of SQL3
- Queries consist of:
  1. a head (`SELECT`) specifying projected attributes (answer relation/table)
  2. a body (`WHERE`) specifying relation/table joins
  3. selections (`FILTER`), that specify Boolean conditions or regular expressions used to filter answers
  4. grouping constructs (e.g., `GROUP BY`, `ORDER BY`) to further structure the answer relation/table
  5. extra constructs specific to semantic web standards

Institut für Maschinelle Sprachverarbeitung

SPARQL - TV Shows (1)

DBpedia

Test this query:

```
SELECT ?episodeName, ?season, ?chalkboardGag
WHERE {
 ?episode <http://dbpedia.org/property/blackboard> ?chalkboardGag.
 ?episode <http://dbpedia.org/property/season> ?season.
 ?episode <http://dbpedia.org/property/episodeName> ?episodeName.
}
ORDER BY ?season
```

What can you say about it?

Institut für Maschinelle Sprachverarbeitung

SPARQL - TV Shows (1)

This query requests to list TVs shows with their episodes (with number and season) and backboard gag if available.

However, it returns empty answers due to a change in the DBpedia schema (DBpedia namespaces).

By looking at the vocabulary of the current version, we build the non-empty query:

```
SELECT ?episodeName ?gag ?season
WHERE{
 ?episode <http://dbpedia.org/property/episode> ?episodeName.
 ?episode <http://dbpedia.org/ontology/seasonNumber> ?season.
 OPTIONAL{?episodeName
       <http://dbpedia.org/property/blackboard> ?gag.}
}
ORDER BY ?season
```

Institut für Maschinelle Sprachverarbeitung

SPARQL - TV Shows (2)

List TV shows with their episodes (with number and season) and backboard gag if available; use prefixes instead of full URIs:

```
SELECT ?episodeName ?gag ?season
WHERE{
 ?episode <http://dbpedia.org/property/episode> ?episodeName.
 ?episode <http://dbpedia.org/ontology/seasonNumber> ?season.
 OPTIONAL{?episodeName
       <http://dbpedia.org/property/blackboard> ?gag.}
}
ORDER BY ?season
```

Institut für Maschinelle Sprachverarbeitung

SPARQL - Simpsons Query (1)

The Simpsons

In order to work with RDF-based data, you need to learn about the things available for querying. DBpedia offers a search page

http://dbpedia.org/fct/

where you can find out more about given entries. From the results of the previous query, look up at least one Simpson episode and find the information that enables you to do the following:

- Extend the query to list stars that appeared in that episode.

Institut für Maschinelle Sprachverarbeitung

SPARQL - Simpsons Query (2)

We run again into the problem that the DBpedia schema doesn't allow to extend the previous query to ask about the Simpson's (and many other TV shows).

We observe that in animated shows, by "actors", **voice actors** are meant. The current version of DBpedia (2017) connects one the hand the episodes of The Simpson's to their seasons, and on the other hand the voice actors to the general entry for the show. Both pieces of information can be connected via the **shared resource**

`<http://dbpedia.org/page/Category:The_Simpsons>`

(i.e., their shared category "The Simpson's"), using the **join variable** `?catShow` in the SPARQL query below.

We can also use filters to remove unwanted information (e.g., episode names in French and so forth) and `distinct` to remove redundancies in answers

Institut für Maschinelle Sprachverarbeitung

SPARQL - Simpsons Query (3)

```
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX dbr: <http://dbpedia.org/resource/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT distinct ?show ?episode ?season ?actor
WHERE{
  ?episodes    dct:subject    ?seasons .
  ?episodes    dct:subject    ?seasons .
  ?seasons     skos:broader   ?category .
  ?category    skos:broader   ?catShow .
  ?shows       dct:subject    ?catShow .
  ?shows       dbo:voice      ?actors .
  ?shows       dbp:showName   ?show .
  ?actors      dbo:birthName   ?actor .
  ?episodes    rdfs:label     ?episode .
  ?seasons     rdfs:label     ?season .
  FILTER(?category    = <http://dbpedia.org/resource/Category:The_Simpsons_episodes>)
  FILTER(?shows       = <http://dbpedia.org/resource/The_Simpsons>)
  FILTER(lang(?episode)      = "en")
  FILTER(lang(?actor)        = "en")
  FILTER(lang(?show)         = "en")
  FILTER(lang(?season)       = "en")
}
```

Institut für Maschinelle Sprachverarbeitung

SPARQL - Companies in Stuttgart

List companies in Stuttgart:

```
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?company where
WHERE{
  ?c rdf:type dbo:Company .
  ?c foaf:name ?company .
  ?c dbo:location dbr:Stuttgart .
}
```

Institut für Maschinelle Sprachverarbeitung

SPARQL - Companies in Stuttgart

List companies in Stuttgart with their homepages:

```
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?company ?homepage
WHERE{
  ?c rdf:type dbo:Company .
  ?c foaf:name ?company .
  ?c dbo:location dbr:Stuttgart .
  ?c foaf:homepage ?homepage .
}
```

Institut für Maschinelle Sprachverarbeitung

SPARQL - Companies in Stuttgart

List companies in Stuttgart with their homepage and founding year:

```
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?company ?homepage ?founded
WHERE{
  ?c rdf:type          dbo:Company .
  ?c foaf:name         ?company .
  ?c dbo:location      dbr:Stuttgart .
  ?c foaf:homepage     ?homepage .
  ?c dbo:foundingYear   ?founded .
}
```

Institut für Maschinelle Sprachverarbeitung

SPARQL - Movies

List movies, directors and their birthdates, order by director name and limit to 10,000 answers:

```
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dbp: <http://dbpedia.org/property/>
SELECT ?movie ?director ?birthdate
WHERE{
 ?movie rdf:type    dbo:Film .
 ?movie dbo:director ?dir .
 ?dir   dbo:birthDate ?birthdate .
 ?dir   dbp:birthName ?director .
}
ORDER BY ?director
LIMIT 10000
```

Institut für Maschinelle Sprachverarbeitung

SPARQL - Movies

List movies starring John Wayne, with their runtimes, lasting more than 4,000, sort (desc.) by running time:

```
PREFIX dbo: <http://dbpedia.org/ontology/>
SELECT ?movie ?runtime
WHERE{
  ?movie dbo:starring <http://dbpedia.org/resource/John_Wayne> .
  ?movie dbo:runtime ?runtime .
}
ORDER BY DESC(?runtime)
```

Institut für Maschinelle Sprachverarbeitung

SPARQL - Movies

List movies starring John Wayne, with their runtimes, lasting at least 4,000 seconds, sort (desc.) by running time:

```
PREFIX dbo: <http://dbpedia.org/ontology/>
SELECT ?movie ?runtime
WHERE{
  ?movie dbo:starring  <http://dbpedia.org/resource/John_Wayne> .
  ?movie dbo:runtime  ?runtime .
  FILTER(?runtime > 4000)
}
ORDER BY DESC(?runtime)
```

Institut für Maschinelle Sprachverarbeitung