Testing OWL Axioms Against RDF Facts
A possibilistic approach

Andrea G. B. Tettamanzi, Catherine Faron-Zucker, and Fabien Gandon

Univ. Nice Sophia Antipolis, CNRS, INRIA, I3S, UMR 7271, France

EKAW 2014, Linköping
Introduction: Ontology Learning

Top-down construction of ontologies has limitations
  - aprioristic and dogmatic
  - does not scale well
  - does not lend itself to a collaborative effort
Bottom-up, *grass-roots* approach to ontology and KB creation
  - start from RDF facts and learn OWL 2 axioms
Recent contributions towards OWL 2 ontology learning
  - FOIL-like algorithms for learning concept definitions
  - statistical schema induction via association rule mining
  - light-weight schema enrichment (DL-Learner framework)
All these methods apply and extend ILP techniques.
Introduction: Ontology validation, Axiom Scoring

Need for evaluating and validating ontologies
- General methodological investigations, surveys
- Tools like OOPS! for detecting pitfalls
- Integrity constraint validation

Ontology learning and validation rely on axiom scoring
- We tackle the problem of testing a single axiom
- Most popular scoring heuristics based on statistical inference

Research Questions:
1. Can we apply a possibilistic approach to axiom testing for ontology learning?
2. Could this be beneficial to ontology and KB validation?
Count-Based Axiom Scoring

Given axiom $\phi$, let us define

- $u_\phi$ the support or sample size for $\phi$
- $u_\phi^+$ the number of confirmations of $\phi$
- $u_\phi^-$ the number of counterexamples (falsifiers) of $\phi$

Score from statistical inference: $\Pr(\phi \text{ is true} \mid \text{evidence})$

- Simple statistics: $\hat{p}_\phi = \frac{u_\phi^+}{u_\phi}$
- Refinements are possible, e.g., confidence intervals

Implicit assumption that we know how to estimate $\Pr(e \mid \phi)$ in

$$\Pr(\phi \mid e) = \frac{\Pr(e \mid \phi) \Pr(\phi)}{\Pr(e \mid \phi) \Pr(\phi) + \Pr(e \mid \neg\phi) \Pr(\neg\phi)}$$

$\Rightarrow$ Alternative scoring heuristics based on possibility theory, weaker than probability theory
Possibility Theory

Definition (Possibility Distribution)

\[ \pi : \Omega \rightarrow [0, 1] \]

Definition (Possibility and Necessity Measures)

\[ \Pi(A) = \max_{\omega \in A} \pi(\omega); \]
\[ N(A) = 1 - \Pi(\overline{A}) = \min_{\omega \in \overline{A}} \{1 - \pi(\omega)\}. \]

For all subsets \( A \subseteq \Omega \),

1. \( \Pi(\emptyset) = N(\emptyset) = 0, \quad \Pi(\Omega) = N(\Omega) = 1; \)
2. \( \Pi(A) = 1 - N(\overline{A}) \) (duality);
3. \( N(A) > 0 \) implies \( \Pi(A) = 1, \quad \Pi(A) < 1 \) implies \( N(A) = 0. \)

In case of complete ignorance on \( A \), \( \Pi(A) = \Pi(\overline{A}) = 1. \)
Support of an Axiom

**BS:** finite set of *basic statements*, i.e., assertions that may be tested by a SPARQL ASK query.

**Definition (Content of Axiom \( \phi \))**

\[
content(\phi) = \{ \psi : \phi \models \psi \} \cap BS.
\]

- \( content(\phi) \) is finite \( \iff \) BS is finite
- every \( \psi \in content(\phi) \) may be tested by a SPARQL ASK query

**Definition (Support of Axiom \( \phi \))**

\[
u_\phi = |content(\phi)|.
\]
Possibility and Necessity of an Axiom

\[
\Pi(\phi) = 1 - \sqrt{1 - \left(\frac{u_\phi - u_-}{u_\phi}\right)^2}
\]

\[
N(\phi) = \sqrt{1 - \left(\frac{u_\phi - u_+}{u_\phi}\right)^2} \quad \text{if } \Pi(\phi) = 1, \ 0 \ \text{otherwise.}
\]
Combination of possibility and necessity of an axiom:

**Definition**

$$\text{ARI}(\phi) = N(\phi) - N(\neg\phi) = N(\phi) + \Pi(\phi) - 1$$

- $-1 \leq \text{ARI}(\phi) \leq 1$ for all axiom $\phi$
- $\text{ARI}(\phi) < 0$ suggests rejection of $\phi$ ($\Pi(\phi) < 1$)
- $\text{ARI}(\phi) > 0$ suggests acceptance of $\phi$ ($N(\phi) > 0$)
- $\text{ARI}(\phi) \approx 0$ reflects ignorance about the status of $\phi$
To test axioms, we define a mapping $Q(E, x)$ from OWL 2 expressions to SPARQL graph patterns such that

$$\text{SELECT DISTINCT } ?x \text{ WHERE } \{ Q(E, ?x) \}$$

returns $[Q(E, x)]$, all known instances of class expression $E$ and

$$\text{ASK } \{ Q(E, a) \}$$

checks whether $E(a)$ is in the RDF base.

For an atomic concept $A$ (a valid IRI),

$$Q(A, ?x) = ?x \text{ a } A.$$
Concept Negation: $Q(\neg C, ?x)$

$$Q(\neg C, ?x) = \{ ?x \ ?p \ ?o . \quad \text{FILTER NOT EXISTS } Q(C, ?x) \}$$  \hspace{1cm} (1)

has the problem of treating negation as failure.

$$Q(\neg C, ?x) = \{ ?x \ a \ ?dc . \quad \text{FILTER NOT EXISTS } \{ ?z \ a \ ?dc . \ Q(C, ?z) \} \}$$  \hspace{1cm} (2)

Better than (1), but just pushes the problem one step further

$$Q(\neg C, ?x) = \{ ?x \ a \ ?dc . \ ?dc \ \text{owl:disjointWith } C \}$$  \hspace{1cm} (3)

OK, but very few DisjointClasses axioms currently found in ontologies!
Concept Negation: Discussion

\[ [Q(C, x)] \]

\[ [Q(D, x)] \]

\[ [Q(D'', x)] \]

\[ [Q(D', x)] \]
SubClassOf($C \sqsubseteq D$) Axioms

To test SubClassOf axioms, we must define their logical content based on their OWL 2 semantics:

$$(C \sqsubseteq D)\mathcal{I} = C\mathcal{I} \subseteq D\mathcal{I}$$

$$\equiv \forall x \ x \in C\mathcal{I} \Rightarrow x \in D\mathcal{I}$$

Therefore,

\[
\text{content}(C \sqsubseteq D) = \{ D(a) : C(a) \text{ is in the RDF store} \} \\
\]

because, if $C(a)$ holds,

$$C(a) \Rightarrow D(a) \equiv \neg C(a) \lor D(a) \equiv \top \lor D(a) \equiv D(a)$$
Support, Confirmations and Counterexamples of $C \sqsubseteq D$

$u_{C \sqsubseteq D}$ can be computed by

\[
\text{SELECT (count(DISTINCT ?x) AS ?u) WHERE \{ Q(C, ?x) \}.}
\]

As for the computational definition of $u_{C \sqsubseteq D}^+$ and $u_{C \sqsubseteq D}^{-}$:

- confirmations: $a$ s.t. $a \in [Q(C, x)]$ and $a \in [Q(D, x)]$;
- counterexamples: $a$ s.t. $a \in [Q(C, x)]$ and $a \in [Q(\neg D, x)]$.

Therefore,

- $u_{C \sqsubseteq D}^+$ can be computed by

  \[
  \text{SELECT (count(DISTINCT ?x) AS ?numConfirmations) WHERE \{ Q(C, ?x) Q(D, ?x) \}}
  \]

- $u_{C \sqsubseteq D}^-$ can be computed by

  \[
  \text{SELECT (count(DISTINCT ?x) AS ?numCounterexamples) WHERE \{ Q(C, ?x) Q(\neg D, ?x) \}}
  \]
Experiments

Experimental Setup:

- DBpedia 3.9 in English as RDF fact repository
- Local dump (812,546,748 RDF triples) loaded into Jena TDB
- Method coded in Java, using Jena ARQ and TDB
- 12 6-core Intel Xeon CPUs @2.60GHz (15,360 KB cache), 128 GB RAM, 4 TB HD (128 GB SSD cache), Ubuntu 64-bit OS.

Two experiments:

1. Explorative test of systematically generated subsumption axioms
2. Exhaustive test of all subsumption axioms in the DBpedia ontology.

Results at [http://www.i3s.unice.fr/~tettaman/RDFMining/](http://www.i3s.unice.fr/~tettaman/RDFMining/).
Experiments & Results

Explorative Experiment

Systematically generate and test SubClassOf axioms involving atomic classes only

- For each of the 442 classes $C$ referred to in the RDF store
- Construct all $C \sqsubseteq D : C$ and $D$ share at least one instance
- Classes $D$ are obtained with query

    SELECT DISTINCT ?D WHERE \{ Q(C, ?x). ?x a ?D \}

Due to the sheer number of axioms thus generated and to the long time it takes to test them, this experiment is still underway.
Explorative Experiment: Test Time

Time to test an axiom inversely proportional to its ARI
Good news: time-out on test $\Rightarrow$ ARI($\phi$) < 0 likely!
Assessment:

1. sort the 380 tested axioms by their ARI
2. manually tag each of them as either true or false

Findings:

- ARI(ϕ) > 1/3 as the optimal acceptance criterion for ϕ
- This would yield 4 FP and 6 FN (97.37% accuracy)
- Misclassifications to blame on mistakes in DBpedia
- Pr score w/ 0.7 threshold yields 13 FN (+7) and 4 FP (=)
Explorative Experiment: Comparison w/ Probabilistic Score

The graphs show the distribution of Acceptance/Rejection Index and Bühmann and Lehmann’s Score frequencies. The bar chart on the left illustrates the frequency distribution, while the scatter plot on the right compares the two scores across different index values. The plots help in understanding the comparative performance of the probabilistic score against the possibilistic one.
Exhaustive Experiment

Test all SubClassOf axioms in DBpedia ontology

- Functional syntax, with query

  ```sql
  SELECT DISTINCT
  concat("SubClassOf(<" , str(?x),
    ">
  ")")
  WHERE { ?x a owl:Class . ?x rdfs:subClassOf ?y }
  ```

- 541 axioms

- Testing “only” took 1 h 23 min 31 s
Exhaustive Experiment: Results

- For 143 axioms, $u_{\phi} = 0$ (empty content!): $\text{ARI}(\phi) = 0$
- For 28 axioms, $\text{ARI}(\phi) < 0 \Rightarrow \exists$ erroneous facts

Examples of axioms $C \sqsubseteq D$ with their counterexamples:

<table>
<thead>
<tr>
<th>Axiom</th>
<th>Counterexamples</th>
</tr>
</thead>
<tbody>
<tr>
<td>dbo:LaunchPad $\sqsubseteq$ dbo:Infrastructure</td>
<td>:USA</td>
</tr>
<tr>
<td>dbo:Brain $\sqsubseteq$ dbo:AnatomicalStructure</td>
<td>:Brain [sic]</td>
</tr>
<tr>
<td>dbo:Train $\sqsubseteq$ dbo:MeanOfTransportation</td>
<td>:New_Jersey_Transit_rail_operations, :ALWEG</td>
</tr>
<tr>
<td>dbo:ProgrammingLanguage $\sqsubseteq$ dbo:Software</td>
<td>:Ajax</td>
</tr>
<tr>
<td>dbo:PoliticalParty $\sqsubseteq$ dbo:Organisation</td>
<td>:Guelphs_and_Ghibellines, :-, :New_People’s_Army, :Syrian</td>
</tr>
</tbody>
</table>

N.B.: counterexamples are instances $a$ such that $C(a)$ and $E(a)$ with $E^I \cap D^I = \emptyset$: in this case, either $C(a)$ is wrong or $E(a)$ is.
Conclusions & Future Work

Contributions

- Axiom scoring heuristics based on possibility theory
- Solid basis for axiom induction, ontology learning
- A framework based on the proposed heuristics

The proposed heuristics

- is suitable for axiom induction
- may be useful for ontology and KB validation
- no less rigorous or objective than probability-based scoring

Future work

- extend experimental evaluation to more general axioms
- enlarge test base w/ additional RDF datasets from the LOD
- improve framework by setting a time-out on query evaluation
Thank you for your attention!