Web Reasoning Using Fact Tagging
Mehdi Terdjimi, Lionel Médini and Michael Mrissa
Introduction
Context

- Reasoning on the Web / WoT
  - Resource-limited devices
  - Complex models
  - Dynamic Web applications
  - Scenario: Smart Home temperature regulation
**Context**

- **OWL 2 RL reasoning**
  - **Facts** (triples)
    - Explicit / Implicit facts
  - Conjunctive rules
    - \( E_1 \land E_2 \land E_3 \rightarrow I_1 \)
  - Loop
    - Until no more facts are produced

- **Complexity**
  - Depends on expressivity + « intricacy level »
  - Transitive closure can be EXPTIME

- **Dynamic KB Maintenance**
  - Insertions / deletions / re-insertions

---

```
window.isSecured \land (t_{out} < t_{in}) \land window.isOpen
↓
cooling.isActivated
```
Related Work
Related work

**Reasoning on the Web**
- EYE [Verborgh et al., 2015]
- CHR.js [Nogatz, 2015]
- Javascript Semantic Web Toolkit [Stepanov, 2011]
- HyLAR [Terdjimi et al. 2015], [Terdjimi et al., 2016]

**Reasoning optimizations**
- Limiting expressivity [Grimm et al., 2012]
- Axioms rewriting [Kollia and Glimm, 2014]
- Triple Pattern Fragments [Verborgh et al., 2014]

**Maintenance**
- Fact counting [Gupta et al., 1993]
- Fact dependency [Goasdoué et al., 2013]
- Delete-Rederive (DRed) [Gupta et al., 1993]
- Incremental Reasoning [Motik et al., 2012]
DRed and Incremental Reasoning

- Used in HyLAR [Terdjimi et al. 2016]
- Re-inferring overhead
  - Common in smart-* applications
  - On cyclic (re-occurring) data
    - Ex: temperature, time, location, etc.
- Costly deletions

→ Can we Improve incremental maintenance?
Contribution
Proposition

- Improve incremental maintenance
  - For reoccurring situations
  - Approach: « keep track » of previous inferences
    - Store previously encountered facts
    - Avoid recalculating previous inferences
    - Filter actually valid facts at selection
Tag-based reasoning

- **Explicit facts**
  - `valid` tag (insertion)
  - `invalid` (deletion)
  - $f_e \cdot \text{valid} \in \{\text{true, false}\}$

- **Implicit facts**
  - Tagged using their explicit antecedents
  - $f_i \cdot \text{derivedFrom} = \{(f_{e1}, f_{e2}), \ldots, (f_{eN})\}$

- **Selection (filtering)**
  - Explicit facts being valid
  - Disjunction of antecedents validity for implicit facts
  - $f_i \cdot \text{isValid}() = (f_{e1} \cdot \text{valid} \land f_{e2} \cdot \text{valid}) \lor \ldots \lor (f_{en} \cdot \text{valid})$
Tag-based reasoning: illustration

**Rules**
- $r_1 : E_1 \rightarrow I_1$
- $r_2 : E_2 \rightarrow I_2$
- $r_3 : I_2 \rightarrow I_1$

**E1 deletion / re-insertion**
- Incremental Reasoning

**Tag-based Reasoning**

Diagram showing the relationships between $E_1$, $I_1$, $E_2$, and $I_2$ with the rules $r_1$, $r_2$, and $r_3$.
Tag-based reasoning: illustration

**Rules**
- r1 : E1 $\rightarrow$ I1
- r2 : E2 $\rightarrow$ I2
- r3 : I2 $\rightarrow$ I1

**I1 selection**
- **Incremental Reasoning**
  - If I1 $\in$ KB $\rightarrow$ I1
  - If I1 $\notin$ KB $\rightarrow$ $\emptyset$

- **Tag-based Reasoning**
  - If I1 $\in$ KB
    - If $E1.valid \lor E2.valid$ $\rightarrow$ I1
    - Otherwise $\rightarrow$ $\emptyset$
  - If I1 $\notin$ KB $\rightarrow$ $\emptyset$
Tag-based reasoning: complexity

- **Time complexity**
  - *Poly()* at first insertion (single iteration) wrt.
    - Number of rules
    - Number of facts
    - Max. number of causes
  - *O(n)* at deletion and re-insertion
  - *O(n^3)* at selection

- **Space complexity**
  - Storing causes: \( C \)
  \[
  |C| \leq \left( \frac{|Fe|}{|Fe|_2} \right) = \frac{|Fe|!}{\left( \frac{|Fe|!}{2} \right)^2}
  \]
  - \( Fe : KB \) explicit facts → limit KB density
  - Limited in the case of cyclic data
Implementation

**HyLAR**
- Parsing interface
  - *Standard Turtle/N3/JSON-LD parsers*
- Storage manager
  - *Includes rdfstore.js triplestore [Hernandez & Garcia 2012]*
- Reasoner
  - *Tag-based and incremental reasoning algorithms*
- Dictionary & Logics
  - *Storage and processing of logic facts*
Evaluation

Comparison with the Incremental Reasoning (Motik et al.)

Experimental conditions

- Schema: Lehigh University Benchmark Ontology [Guo et al., 2005]
- Datasets: O1, O2 et O3 (resp. 5759, 7394 et 8824 triples)
- Rules: subsumption, transitivity, inverse, equivalence, equality
- 10 cycles = 1 classification and 1 insertion, followed by 10 x (deletion, re-insertion and selection)

<table>
<thead>
<tr>
<th>TB</th>
<th>CLASSIF+INIT</th>
<th>INSERT</th>
<th>DELETE</th>
<th>RE-INSERT</th>
<th>SELECT</th>
<th>10 CYCLES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>O1</td>
<td>O2</td>
<td>O3</td>
<td>O1</td>
<td>O2</td>
<td>O3</td>
</tr>
<tr>
<td>Rsu</td>
<td>5121</td>
<td>4650</td>
<td>3617</td>
<td>1313</td>
<td>1333</td>
<td>1179</td>
</tr>
<tr>
<td>Perf. (%)</td>
<td>-97</td>
<td>-4</td>
<td>-1</td>
<td>-34</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Rs-trans-inv</td>
<td>4302</td>
<td>4222</td>
<td>3305</td>
<td>1156</td>
<td>1028</td>
<td>802</td>
</tr>
<tr>
<td>Perf. (%)</td>
<td>-93</td>
<td>-2</td>
<td>-3</td>
<td>-1</td>
<td>-6</td>
<td>-5</td>
</tr>
<tr>
<td>Requiv</td>
<td>1782</td>
<td>1597</td>
<td>1048</td>
<td>375</td>
<td>344</td>
<td>360</td>
</tr>
<tr>
<td>Perf. (%)</td>
<td>-1</td>
<td>-5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-25</td>
</tr>
<tr>
<td>Requal</td>
<td>4050</td>
<td>3812</td>
<td>2877</td>
<td>408</td>
<td>480</td>
<td>454</td>
</tr>
<tr>
<td>Perf. (%)</td>
<td>-83</td>
<td>-7</td>
<td>-5</td>
<td>-21</td>
<td>-3</td>
<td>-17</td>
</tr>
<tr>
<td>Rall</td>
<td>7510</td>
<td>6118</td>
<td>4863</td>
<td>3519</td>
<td>2529</td>
<td>2001</td>
</tr>
<tr>
<td>Perf. (%)</td>
<td>-96</td>
<td>-7</td>
<td>-8</td>
<td>-19</td>
<td>-11</td>
<td>-3</td>
</tr>
</tbody>
</table>

LIRIS
Discussion
**Goal fulfilled**

**Advantage**
- Performs well for reoccurring incoming facts

**Overheads**
- At first insertion (to store causes)
- At selection
  - May take time in highly intricated graphs

→ **Use the right level of abstraction**
Conclusion
Conclusion

Contribution

- Tag-based reasoning
  - Implemented in the Web reasoner HyLAR
  - Improved KB maintenance
    - For re-occurring data scenarios
    - At re-insertion and deletion times

Perspectives

- "Fact forgetting"
- Discretizing fact sets
Any questions