7.

Knowledge Representation in Social Context

CS227
Spring 2011
Outline

- Vision for Social Machines
- From Web to Semantic Web
- Two Use Cases
- Summary
The Beginning of Social Machines

Image credit: http://www.lifehack.org
Social Machines

Real life is and must be full of all kinds of social constraint – the very processes from which society arises. Computers can help if we use them to create abstract *social machines on the Web: processes in which the people do the creative* work and the machine does the administration. . . The stage is set for an evolutionary growth of new social engines. The ability to create new forms of social process would be given to the world at large, and development would be rapid.

--- T. Berners-Lee, M Fischetti

*Weaving the Web: The Original and Ultimate Destiny of the World Wide Web*

1999
Limitations of Current Social Machines

• The current social machines function isolated from one another
  – Just as human communities are linked to each other online communities must be linked on the web
  – There is no single set of policies or mechanisms that would work across all domains
What’s Needed?

• Technology must be developed to allow communities to construct, share and adapt social machines
  – Communities of scientists, educators, school children, citizen activists
• Problems that our society faces today are such that only the concerted effort of groups of people operating with a joint power much greater than that of the individual can hope to provide solutions
• A new generation of web technologies is needed
Key Questions

- What is the function of a social machine and what KR&R it needs?
- What KR&R formalism encourages multiple independent authors to contribute content?
- How can you use social processes during the content authorship process?
- What incentive mechanisms encourage useful content?
- What is the quality of resulting content? What is the minimum threshold for the content to be useful?
Example Social Machines

- Current World Wide Web
- Facebook
- Wikipedia
Current World Wide Web

- What is the function of a social machine and what KR&R it needs?
  - Information sharing of documents
- What KR&R formalism encourages multiple independent authors to contribute content?
  - Hyperlinked Text
- How can you use social processes during the content authorship process?
  - Very little social process
- What incentive mechanisms encourage useful content?
  - If many other people link to you, you will show up at the top of searches
- What is the quality of resulting content? What is the minimum threshold for the content to be useful?
  - Spotty, but still useful for finding things
Facebook

- What is the function of a social machine and what KR&R it needs?
  - Social networking, friends, likes, …
- What KR&R formalism encourages multiple independent authors to contribute content?
  - Links between friends, events, pictures, etc.
- How can you use social processes during the content authorship process?
  - Commenting on content of others
- What incentive mechanisms encourage useful content?
  - Opportunity of staying in touch
- What is the quality of resulting content? What is the minimum threshold for the content to be useful?
  - Certainly depends on the person
• What is the function of a social machine and what KR&R it needs?
  – To serve as an encyclopedia
• What KR&R formalism encourages multiple independent authors to contribute content?
  – Text + Infoboxes
• How can you use social processes during the content authorship process?
  – Extensive editorial process
• What incentive mechanisms encourage useful content?
  – Opportunity to lead an area and getting your work read
• What is the quality of resulting content? What is the minimum threshold for the content to be useful?
  – Excellent content
Two New Social Machines

- Web of Data as opposed to web of Documents
  - Linked Data
- Answering questions over Wikipedia
  - Semantic Wikipedia

- Let us consider these two social machines in the context of key questions for social KR&R
Function of Linked Data

Function of Linked Data
``Connect the Walled Gardens``

KR&R for Linked Data

- Everything has a URI
  - Don't say "color" say http://example.com/2002/std6#col
- Connect URIs using labeled edges:

Can be encoded in XML

Simplicity and mathematical consistency

This is called Resource Description Framework (RDF)

Slide adapted from Sir Tim Berners-Lee
Linked Data Principles

Set of best practices for publishing structured data on the Web in accordance with the general architecture of the Web.

1. Use **URIs** as names for things.
2. Use **HTTP URIs** so that people can look up those names.
3. When someone looks up a URI, provide useful **RDF** information.
4. Include RDF statements that **link** to other URIs so that they can discover related things.

The Basis: RDF Data Model

Flexible graph-based data model.
Data items are identified with HTTP URIs

HTTP URIs take the role of global primary keys.

**pd:cygri** = http://richard.cyganiak.de/foaf.rdf#cygri

**dbpedia:Berlin** = http://dbpedia.org/resource/Berlin
Resolving URIs over the Web

The HTTP protocol brings together identification and retrieval again.
Following Links deeper into the Web
From XML to OWL

XML

Issue addressed: how to express data in text?
XML Solution: "wrap" data within start tag/end tags, and empower users to create their own tags
Example:

```
<altitude>1500 feet</altitude>
```

RDF and RDF Schema

Issue addressed: how can data support statements?
RDF Solution: use a subject, property, object pattern
Example:

```
<Fighter rdf:ID="F16">
  <altitude>50,000 feet</altitude>
  <builder>Lockheed</builder>
</Fighter>
```

XML Schema (XMSL)

Issue addressed: how should the type structure of the data be expressed?
XML Schema Solution: XML templates
Example:

```
<element name="altitude" type="integer"/>
```

OWL

Issue addressed: how to express data semantics?
OWL Solution: use inheritance and a description logic to express restrictions and describe entailment
Example:

```
<owl:Class rdf:ID="Fighter">
  <rdfs:subClassOf rdf:resource="#Aircraft"/>
</owl:Class>
```

HTML → XML & XMSL → RDF → OWL

Slide Credit: Mark Greaves
Semantic Web Language Stack

Slide Credit: Sir Tim Berners-Lee
LOD Datasets on the Web: May 2007

- Over 500 million RDF triples
- Around 120,000 RDF links between data sources

Slide Credit: Christian Bizer
LOD Datasets on the Web: September 2010

- Over 24.7 billion RDF triples
- Over 436 million RDF links between data sources

Slide Credit: Christian Bizer
Reuse Terms from Common Vocabularies

• Common Vocabularies
  – Friend-of-a-Friend for describing people and their social network
  – SIOC for describing forums and blogs
    • Semantically linked online communities http://sioc-project.org/
  – SKOS for representing topic taxonomies
    • Simple knowledge organization system http://www.w3.org/2004/02/skos/
  – Organization Ontology for describing the structure of organizations
  – GoodRelations provides terms for describing products and business entities
  – Music Ontology for describing artists, albums, and performances
  – Review Vocabulary provides terms for representing reviews

Slide Credit: Christian Bizer
Multiple Ontologies

- It is not one person's ontology
- It is not several people common ontology
- It is many people's many ontologies
- So it is a mess, but a meaningful mess.

Slide credit: Jerome Euzenat
Ontology Heterogeneity

Slide credit: Jerome Euzenat
Addressing Ontology Heterogeneity

Slide credit: Jerome Euzenat
Ontology Alignment

Slide credit: Jerome Euzenat
Ontology Alignment Evaluation Initiative (OAEI)

- Formal comparative evaluation of different ontology-matching tools;
- Run every year since 2004;
- Variety of test cases (in size, in formalism, in content);
- Results consistent across test cases;
- Results very dependent on the tasks and the data (from under 50% of precision and recall to well over 80% if ontologies are relatively similar);
- Progress every year!

http://oaei.ontologymatching.org
**Publish Schema Mappings on the Web**

- **Simple Mappings:** OWL, RDFS, SKOS
  - `owl:equivalentClass`, `owl:equivalentProperty`, `rdfs:subClassOf`, `rdfs:subPropertyOf`
  - SKOS Mapping Vocabulary

- **Complex Mappings:** R2R
  - provides value transformation functions
  - structural transformations

```xml
<http://xmlns.com/foaf/0.1/Person>
owl:equivalentClass
<http://dbpedia.org/ontology/Person>.
```
Task involved in Linked Data Consumption

Application Layer
- Application Code

Data Access, Integration and Storage Layer
- Web Data Access Module
- Vocabulary Mapping Module
- Identity Resolution Module
- Quality Evaluation Module
- Integrated Web Data

Web of Linked Data

Publication Layer
- LD Wrapper
- Database A
- LD Wrapper
- Database B
- RDFa
- Legacy App C
- RDF/XML

Slide Credit: Christian Bizer
Linked Data Summary

- What is the function of a social machine and what KR&R it needs?
  - To provide data interoperability
    • For example, between LinkedIn, Facebook and Wikipedia

- What KR&R formalism encourages multiple independent authors to contribute content?
  - Simpler the better
  - RDF is the simplest layer in the KR&R

- How can you use social processes during the content authorship process?
  - Still evolving and unclear

- What incentive mechanisms encourage useful content?
  - Still evolving and unclear

- What is the quality of resulting content? What is the minimum threshold for the content to be useful?
  - Still evolving and unclear
Semantic Wikipedia

• What is the function of a social machine and what KR&R it needs?
  – Answering questions over the Wikipedia content
What Wikipedia knows

- Wikipedia has articles about
  - all cities
  - their populations
  - their mayors

But, can I ask for a list of the world’s largest city with a female mayor?
Search results
From Wikipedia, the free encyclopedia

Chile (section Largest cities)
Chileans elected their first female president, Michelle Bachelet Jeria, of the Socialist Party, defeating Sebastián ... Largest cities: Culture ...
124 KB (16,703 words) - 17:20, 7 April 2011

Chicago (redirect from The Windy City)
date April 2011) is the largest city in the U.S. state of Illinois ... The successes of the Bike Program are due in large part to Mayor ...
131 KB (17,317 words) - 22:51, 7 April 2011

Rockville, Maryland (section Mayor)
the city's population is 50,734, making it the second-largest ... whose members, along with the Mayor, serve as the legislative body of the city ...
25 KB (3,405 words) - 14:51, 7 April 2011

Toronto (redirect from Accordion City)
most populous city in North ... is the seventh largest urban region ... with Mel Lastman as its first mayor (after being mayor of North York) ...
115 KB (14,990 words) - 00:29, 9 April 2011

Baltimore (redirect from Baltimore City)
The "Leland Stanford Junior University", commonly referred to as "Stanford University" or "Stanford", is a [Private university|private research university] on an
{{convert|8180|acre|ha|sing=on}} campus located near [[Palo Alto, California]].
Semantic Media = Media Wiki + KR&R

- Extend wiki with **typed links**
- So the computer “understands” it

Karlsruhe is a city in
[[located in::Germany]]
with an area of
[[area::173 km²]].
# Mapping of SMW Data Model to OWL

<table>
<thead>
<tr>
<th>Term</th>
<th>OWL Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Article</td>
<td>owl:Individual</td>
</tr>
<tr>
<td>Category</td>
<td>owl:Class</td>
</tr>
<tr>
<td>Relation (page in range)</td>
<td>owl:ObjectProperty</td>
</tr>
<tr>
<td>Relation (datatype in range)</td>
<td>owl:DatatypeProperty</td>
</tr>
<tr>
<td>Typed link [[[relation::object]]]</td>
<td>Object property instance</td>
</tr>
<tr>
<td>[[[relation::value]]]</td>
<td>Datatype property instance</td>
</tr>
<tr>
<td>[[[Category::class]]] (on article page)</td>
<td>rdf:type class</td>
</tr>
<tr>
<td></td>
<td>Class instantiation</td>
</tr>
<tr>
<td>[[[Category::class]]] (on category page)</td>
<td>rdfs:subClassOf class</td>
</tr>
<tr>
<td></td>
<td>Subsumption</td>
</tr>
</tbody>
</table>
Karlsruhe

Karlsruhe is a city with an area of 173.46 km² on the Rhine river in Germany. Further background information can be found in the Wikipedia article about Karlsruhe.

Karlsruhe on semanticweb.org

The following information is automatically extracted from the contents of semanticweb.org. To add or remove entries, create or edit the respective pages.

Events


Organizations and Institutions

AIFB, FZI, and Universität Karlsruhe (TH)

Facts about Karlsruhe

- Adjacent to: Rhina
- Area: 173,460,000 m² (173.46 km², 17,346 ha, 66.973 miles²)
- Located in: Germany

Category: City
Located in Germany

A list of all pages that have property "Located in" with value "Germany"

- Celle
- Elbe
- Hannover
- Worms
- Frankfurt am Main
- Baden-Württemberg
- Hamburg
- Halle (Saale)
- Saxony-Anhalt
- Freiburg
- Marburg
- Saarland
- Bremen
- Hesse
- Rhineland-Palatinate
- Wiesbaden
- Munich
- Isar
- Dresden
- Karlsruhe

Property: Located in
Value: Germany
Making it Easier to Enter KR&R

- Templates
- Semantic Forms
- Side Bar Editor
_templates

{{City
| Country=Germany
| Population=285812
| State=Baden-Württemberg
| Founded=1715
| Mayor=Heinz Fenrich
| Area=173 km²
}}
## Semantic Forms

### Input forms

**Edit City: Karlsruhe**

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Germany</td>
</tr>
<tr>
<td>State</td>
<td>Baden-Württemberg</td>
</tr>
<tr>
<td>Population</td>
<td>285,812</td>
</tr>
<tr>
<td>Area</td>
<td>173 km²</td>
</tr>
<tr>
<td>Mayor</td>
<td>Heinz Fennich</td>
</tr>
</tbody>
</table>

**Free text:**

"Karlsruhe" is the seat of the federal highest court of Germany.
Halo Side Bar

Halo sidebar

Editing Karlsruhe

Karlsruhe (population: 285,812) in 2006 is a city in the south west of Germany, in the States of Germany (Bavaria). It is located near the German border.

Founded in 1715 as Karlsruhe Palace, the surrounding town became the seat of two of the highest courts in Germany. The Federal Constitutional Court of Germany, whose decisions have the force of a law, and the Federal Court of Justice of Germany.

[[Category:City]]
Implementation in Semantic MediaWiki

Wiki users

Questions

Feedback

Annotations

Semantic search & browsing

Data sources

Wiki content

Other sources

Initial seed

Extraction tools

Data sources for Information Extraction

QuestionAPI

Semantic MedialWiki
• Available as RDF
• Via SPARQL endpoint
• RDFa export in development
Pay Offs for Adding the Markup

- Automatic Tables and Summaries
- Automatic Visualization
- Most importantly: queries
Automatic Tables and Lists

- Many pages answer questions
  - list of female tennis players
  - asteroids named after people
  - countries sorted by area, population, …
- They can be generated automatically
- Less maintenance tasks
- Higher consistency
Timeline of upcoming events

ISWC2008
- The 7th International Semantic Web Conference
- City: Karlsruhe
- Dates: 26 October 2008 to 30 October 2008
- Time: Sun, 26 Oct 2008 00:00:00 GMT to Thu, 30 Oct 2008 00:00:00 GMT

SDoW2008
- OWLED 2008
- OM-2008
- ISWC2008

Events with articles in this wiki.

RR2008
Inline Queries

{{#ask:
[[Category:City]]
[[located in::Germany]]
|?state
|?population
|?area#km²
}}
Inline Query Results

|| State       | Population | Area     |
|-------------|------------|----------|
| Berlin      | Berlin (state) | 3,301,407 | 891.60 km² |
| Celle       | Lower Saxony | 72,000   | 176.01 km²  |
| Dresden     |             | 508,351   | 328.8 km²   |
| Freiburg    | Baden-Württemberg | 217,547 | 153.07 km²  |
| Halle (Saale)| Saxony-Anhalt  | 240,000  | 135 km²      |
| Hannover    | Lower Saxony | 522,944   | 204.01 km²   |
| Karlsruhe   | Baden-Württemberg | 285,812 | 173.46 km²   |
| Koblenz     |             | 105,668   | 105.02 km²   |
| Leipzig     | Saxony      | 510,274   | 297.5 km²    |
| Marburg     | Hessen      | 79,139    | 124.5 km²    |
| Munich      | Bavaria     | 1,300,000 | 310.799 km²  |
| Wiesbaden   | Hessen      | 300,427   | 204.1 km²    |
| Worms       | Rhineland-Palatinate | 85,829  | 108.73 km²   |
Homework 2 / Problem 2

- Create a Wiki about cars
  - To be authored by the whole class
  - Two pre-defined properties
  - Rest will emerge as part of the social authorship process

- Social queries
  - Who owns the same car as me?
  - Which is the most popular car?
  - etc.
Semantic Wikipedia Summary

• What is the function of a social machine and what KR&R it needs?
  – To answer questions over the content of Wikipedia
• What KR&R formalism encourages multiple independent authors to contribute content?
  – Text + Infoboxes + RDF + OWL
• How can you use social processes during the content authorship process?
  – The existing Wikipedia process needs to be extended for ontologies
• What incentive mechanisms encourage useful content?
  – Still to be designed
• What is the quality of resulting content? What is the minimum threshold for the content to be useful?
  – Yet to be discovered
Real life is and must be full of all kinds of social constraint – the very processes from which society arises. Computers can help if we use them to create abstract social machines on the Web: processes in which the people do the creative work and the machine does the administration. . . The stage is set for an evolutionary growth of new social engines. The ability to create new forms of social process would be given to the world at large, and development would be rapid.

--- T. Berners-Lee, M Fischetti

Weaving the Web: The Original and Ultimate Destiny of the World Wide Web
1999
Related Conferences

The 10th International Semantic Web Conference
October 23-27, 2011
Bonn, Germany

2011 Semantic Technology Conference
JUNE 5 - 9 SAN FRANCISCO, CA
Presented by semanticweb.com & Mediabistro

20th international world wide web conference
28th March - 1st April 2011
hicc, hyderabad, india
### Reading

**Required**
- From the Semantic Web to social machines: A research challenge for AI on the World Wide Web by Jim Hendler and Tim Berners-Lee *Artificial Intelligence, Volume 174, Issue 2, February 2010, Pages 156-161*

**Optional**