On the Empirical Evaluation of SWS Approaches

Towards Common SWS Test Collections

Ulrich Küster, Birgitta König-Ries

(Friedrich Schiller University Jena, Germany)
The Grand Vision of SWS

Web Activities

- Discovery
- Binding
  
Automated Activities

- Selection
- Composition
- Orchestration
- Mediation
The Problem with SWS

SWS Cycle

- Mature Technology
  - Improve
- Reliable Evaluations
  - Used in
- Real Data
  - Create
- Real Applications
  - Empowers
The Problem with SWS

Mature Technology requires Reliable Evaluations

SWS Deadlock requires Realistic Test Data

Real Applications require Realistic Test Data
Evaluation of SWS

- Different tasks
  - Discovery, Composition, Mediation, ...
- Different criteria
  - Usability, Efficiency, Automation, ...
- Different context assumptions
  - Service complexity, involved actors, involved business value, involved transaction type, ...

→ Different evaluations
→ Different needs for data
The „ideal“ SWS Test Collection

1. Is not specific to a single formalism
2. Contains unambiguous natural language descriptions
3. Contains different semantic descriptions for the same services
4. Covers a variety of domains, use cases, and services
5. Is large enough to allow statistically significant tests
6. Is built by different people to avoid biases
7. ...

⇒ Tough requirements!
How are we doing so far?

- **OWLS-TC 2.2:**
  - ~1000 advertisements, 29 requests, only IO, mostly artificial (but not generated!), relevance judgments
- **SWS-TC 1.1:**
  - 241 advertisements, fairly realistic, only IO, little natural language documentation, no relevance judgments
- **WSML/WSMO? SAWSDL?**
  - None or next to none available
- **Natural language / neutral formalism?**
  - None or next to none available
- **WSDL?**
  - Lots of test data, no test collection though
The „ideal“ Test Collection revisited

1. Is not specific to a single formalism
2. Contains unambiguous natural language descriptions
3. Contains different semantic descriptions for the same services
4. Covers a variety of domains, use cases, and services
5. Is large enough to allow statistically significant tests
6. Is built by different people to avoid biases

7. ...

→ Tough requirements!
→ Community involvement is crucial!
→ Did not happen sufficiently so far!
1. Promote exchange and reuse of existing data
   - Wiki-like collaboration model
   - GPL-licensed data

2. Improve structure, documentation, and usability
   - Extensible, open source web portal
   - Relational database instead of flat files

3. Support reuse and comparison across formalisms
   - Service as primary concept, descriptions secondary
   - WordNet as neutral formalism
OPOSSum's Relational Data Model

Diagram showing relationships between Resource, Service Description, Formalism, Collection, Service Implementation, Category, Tag, Parameter, Request, Offer, and Relevance Judgement.
Integration of OWLS-TC and SWS-TC

1. Extracted
   - Service domain
   - Imported ontologies
   - Service name from profile
   - Textual description from profile
   - IO parameters
     - Name
     - Type (ontological reference)
     - Parameter description

2. Manually mapped parameter types to WordNet

3. Added service entries to OPOSSum

4. Attached original description to service entry
Status of OPOSSum / Collection

- 1426 service entries
  - 1032 from OWLS-TC, 241 from SWS-TC, 113 real geocoding services, some others
- 1477 service descriptions
  - mostly OWL-S, some WSDL, no WSML/SAWSDL
- Full binary relevance judgments for OWLS-TC services
- Graded relevance judgments for OWLS-TC Education subset

- Online portal to search, access, edit, and upload data
- Programmatic API through MySQL access

➤ Demoed yesterday
➤ Available at http://fusion.cs.uni-jena.de/OPOSSum
Conclusions and Future Work

- No standard test collections so far
- Standard test collection must be built collaboratively
- OPOSSum supports this 😊

Please:
- Upload and share your data
- Edit and improve existing data
- Provide WSML/SAWSDL (and other) descriptions
- Extend the portal
- Share your tools
  - WordNet based search engines
  - Semantic search engines
  - SWS formalism translators
  - Other tools and services
  - ...

ICSC 2008, Santa Clara, CA, USA - August 4-7 2008
Thank You!

Questions?

Ulrich Küster (Ulrich.Kuester@uni-jena.de)
http://fusion.cs.uni-jena.de/ukuester
http://fusion.cs.uni-jena.de/OPOSSum