OWLS-TC

OWL-S Service Retrieval Test Collection

Version 3.0 revision 1

User Manual

November 10, 2009

Matthias Klusch, Mahboob Alam Khalid, Patrick Kapahnke,
Benedikt Fries, Martin Vasileski

Saarbrücken, Germany

Table of contents

Introduction

This is the third version of the OWL-S service retrieval test collection named OWLS-TC3. The collection is intended to support the evaluation of the performance of OWL-S service matchmaking algorithms. It provides 1007 semantic Web services written in OWL-S 1.1 (and for backward compatibility OWLS 1.0) from seven different domains (education, medical care, food, travel, communication, economy, weapons). It provides a set of 29 test queries which are associated with relevance sets to conduct performance evaluation experiments.

Major improvements of OWLS-TC3 to former versions are the following:

- Binary relevance sets have been improved based on collaborative relevance assessments.
- Graded relevance sets as for example used in NTCIR[1] are now available.
- Groundings in WSDL for all OWL-S 1.1 files.
- All relevance sets are provided in XML to allow easy computational access. The provided XML file can for example be used in the Semantic Web Service Matchmaker Evaluation Environment (SME2) available at http://semwebcentral.org/projects/sme2.
- Many bugs like syntax errors have been fixed.

Part of the services contained in OWLS-TC3 were retrieved from public IBM UDDI registries, and semi-automatically transformed from WSDL to OWL-S. Up to this version, around 30 people (DFKI, University of Jena, University of Thessaloniki and others) have worked on improvements and extensions. Relevance sets have been defined collaboratively.

Please note, that no *standard* test collection for OWL-S service retrieval does exist yet. As a consequence, OWLS-TC can only be considered as one possible starting point for any activity towards achieving such a standard collection by the community as a whole.

OWLS-TC is available at semwebcentral.org: http://projects.semwebcentral.org/projects/owls-tc/

[1] NII Test Collection for IR Systems, *http://research.nii.ac.jp/ntcir/*

Installation

1. Install a local web server like the “Apache HTTP Webserver”: We recommend to install **XAMPP** (see following subsection “Install XAMPP”).[2]

2. Copy the folders “services”, “ontology”, “queries” and “wsdl” of the OWLS-TC3 release to the public http root folder (e.g. apache/htdocs)

The usage of a local web server is necessary, since all ontologies used by a matchmaking tool, e.g. OWLS-MX, are expected to be available locally, and problems with incorrect paths to access ontologies used to describe service I/O concepts can thus be avoided.

Install XAMPP

Homepage: http://www.apachefriends.org/en/xampp.html

Download:
http://prdownloads.sourceforge.net/xampp/xampp-win32-1.5.0-pl1-installer.exe?download

XAMPP is an easy to install Apache HTTP web server distribution also containing MySQL, PHP and Perl. Despite the fact that it contains more applications than needed for OWLS-TC it allows a first time user to avoid the difficult configuration of the Apache HTTP web server.

The following manual should guide you through the process of installing XAMPP for the use with OWLS-TC, while using the above linked Windows version.

1. The installer starts with the language selection. Once you selected your desired language it will prompt you for an installation directory. You can select anything here, but we will consider the default “c:\apachefriends\xampp\” in this walkthrough.

2. The installation and configuration will take some minutes and temporarily open a command prompt. Once it is finished the installer will ask you if you want to install the web server as a windows service.
   Unless you want to start the local web server at windows startup you should deny this question. For the sole purpose of OWLS-TC the usual manual startup suffices.

3. Now you can start the Apache control panel, which will allow you to start the server by answering yes to the next question.

4. Using the control panel you can start the server by simply pressing the according Start” button next for the Apache.
   If you use Windows XP SP2 (or have any other firewall installed) your firewall will ask you if you if the program “Apache HTTP Server” should be blocked from the internet. It is necessary that you allow access to the program.

[2] The installer available from the Apache homepage in the Web, unfortunately, caused some problems with OWLS-MX on a WinXP Professional/Java1.5 equipped laptop.

5. The installation of the web server is finished at this point. To install the test collection you can simply copy the directory “ontology”, “services”, “queries” and “wsdl” to the apache web root directory (unless you specify something else it is “htdocs” in the XAMPP directory, so in our case “c:\apachefriends\xampp\htdocs”).

6. Now you can test if you can open the directory http://127.0.0.1/services/ in your web browser.

The **archive of the test collection** contains the following **subdirectories**,

- ***services***
  All services of the test collection (subdirectories for the different supported OWLS-Versions). *Supposed to be copied to the root directory of the local web server.*

- ***queries***
  All services requests of the test collection (subdirectories for the different supported OWLS-Versions). *Supposed to be copied to the root directory of the local web server*

- ***ontology***
  Ontologies used by the services and service requests. *Supposed to be copied to the root directory of the local web server*

- ***domains***
  Contains all services sorted according to their domains

- ***wsdl***
  Contains the WSDL groundings for all of the services. *Supposed to be copied to the root directory of the local web server*

Structure

Domains

OWL-S Service Retrieval Test Collection version 3.0 consists of 1007[3] indexed OWL-S services from the following 7 domains:

1- education
2- medical care
3- food
4- travel
5- communication
6- economy
7- weapon

For information on the OWL-S web service description language standard of the W3C, please check, for example, http://www.daml.org/services/.

Services

The current status of the number of services and queries related to each of these domains are as follows.

DOMAIN	#services	#queries
education	286	6
medical care	73	1
food	34	1
travel	197	6
communication	59	2
economy	395	12
weapon	40	1

The restrictions of OWL-S 1.1 service descriptions of this version of OWLS-TC are as follows:

(a) The services are not *describedby* Process Model, but instead are directly *describedby* Atomic process.

(b) Profile:
Set of profile parameters is restricted to minimal service I/O information (ServiceName, textDescription, hasInput, hasOutput);
No precondition and effect parameter values.

(c) Grounding in WSDL is provided for all of the OWL-S 1.1 services. These WSDL groundings are created from the existing OWL-S services using the tool OWLS2WSDL[4]. For each service, by using the option *owls2wsdl* from this tool, the WSDL file was created

[3] Some services appears in more than one category. Therefore the number of services is 1007 if just the first occurrence of each service is considered and 1080 if we consider repetitions accross different categories.
[4] *http://semwebcentral.org/projects/owls2wsdl/*

automatically (although some manual changes had to be done, i.e. fixing the namespaces) and by using the option *owl-s* the groundings for the OWL-S services were created.

The concepts used in the service input/output parts refer to ontologies (.owl files) that are stored in the /ontology directory.

Requests and Relevance sets

Overview

All relevance sets created for OWLS-TC3 are available in XML format in **owls-tc3.xml** provided in the main directory of the distribution. These were created using the SWSRAT (**S**emantic **W**eb **S**ervice **R**elevance **A**ssessment **T**ool) tool which was developed by Steffen Metzger in collaboration with Patrick Kapahnke and Matthias Klusch at the R&D division I2S of the German Research Center for Artificial Intelligence (DFKI Saarbruecken, Germany). It contains binary as well as graded relevance sets for all 29 queries. Different users of the SWSRAT tool subjectively assessed the service request/offer pairs. Not all possible combinations have been rated, thus the relevance sets can not be considered as *complete*. A pooling strategy as used in TREC[5] based on the top-100 results of participants of the S3 contest[6] in 2008 has been conducted.

Graded Relevance

The graded relevance sets have been created using the following 4-graded scale:

- **highly relevant** (value: 3) - Any service offer that is exactly what the user asked for (or even better for him, e.g. by giving additional information)
- **relevant** (value: 2) - Any service offer that *might* answer the request completely or does the requested job at least partially
- **potentially relevant** (value: 1) - Any service offer that *may* be helpful.
- **nonrelevant** (value: 0) - Anything totally irrelevant to the service request

Binary Relevance

The collaboratively created graded relevance assessments were projected onto a binary relevance scale using SWSRAT. The approach of *relaxed* binary relevance is choosen for this, which means that a service offer is considered as binary relevant to a query, if it is at least potentially relevant according to the graded scale given above.

[5] Text REtrieval Conference, *http://trec.nist.gov/*
[6] Semantic Service Selection, *http://www-ags.dfki.uni-sb.de/~klusch/s3/*

Structure of the XML Relevance File

```
<binaryrelevanceset>     //for the binary relevance sets
    <request ...>
        <name ... />
        <uri .../>
        <ratings>
            <offer ...>     //each request contains multiple offers
                <name ... />
                <uri .../>
                <relevant>value</relevant>     //here the value can be 0 or 1
            </offer>
            ...
         </ratings>
    </request>
    ...
</binaryrelevanceset>


<relevancegrades>     //for the graded relevance assessments
    <services>
        <request ...>
            <name ... />
            <uri .../>
            <ratings>
                <offer ...>
                    <name ... />
                    <uri .../>
                    <grade ...>
                        <name>NameOfGrade</name>     //highly relevant (value 3); relevant (value 2);
                        <value>ValueOfGrade</value>     //potentially relevant (value 1); nonrelevant (value 0)
                    </grade>
                </offer>
                ...
             </ratings>
        </request>
        ...
    </services>
</relevancegrades>
```

Used Ontologies

To describe the semantics of I/O concepts of services in this test collection, 23 different ontologies have been used. They were retrieved from various public sources in the Web for non-commercial use, and are included in the collection for your convenience. Some of them are listed in the following.

Most ontologies were taken from Teknowledge, including

1. the "**Suggested Upper Merged Ontology (SUMO)**"
   upon which several other ontologies rely.
   Project: http://reliant.teknowledge.com/
   Ontology: http://reliant.teknowledge.com/DAML/

2. **Larflast** ontology (University of Sofia)
   Project: http://www-it.fmi.uni-sofia.bg/larflast/
   Ontology: http://www.larflast.bas.bg/

3. **Wine and food ontology** (W3C)
   Project: http://www.w3.org/
   Ontology: http://www.w3.org/TR/2003/PR-owl-guide-20031215/

4. The **AKT Reference Ontology** v2 (AKT Technologies)
   Project: http://www.aktors.org/
   Ontology: http://d3e.open.ac.uk/akt/2002/ref-onto.html
   http://www.aktors.org/ontology/

5. **University ontology** by Zhengxiang Pan
   Dept. of Computer Science and Engineering, Leigh University
   Project: http://www.lehigh.edu/~zhp2/
   Ontology: http://www.lehigh.edu/~zhp2/2004/0401/

6. Taken from the Protege demo owl ontology site: **travel.owl** A tutorial ontology for a Semantic Web of tourism. Contributed by Holger Knoblauch.
   Project: http://protege.stanford.edu
   Ontology: http://protege.stanford.edu/plugins/owl/owl-library/
   http://www.owl-ontologies.com/

7. Taken from German Research Center for Artificial Intelligence website
   Project: http://www.dfki.de/scallops/
   Ontology: http://www.dfki.de/scallops/health-scallops/

8. Simple **Book ontology** (DFKI): book.owl

9. **Simplified SUMO ontology** (DFKI): simplified_sumo.owl
   Extract from ontology 1 to allow faster parsing

10. My Ontology (DFKI): anonymous ontology

PLEASE NOTE: If you are using the ontologies 8, 9, or 10, please make sure that you change their local import URLs to your needs.

Disclaimer

Copyright (C) 2009, DFKI

OWLS-TC is free for non-commercial use only. You can redistribute it and/or modify it under the terms of the GNU General Public License (GPL) as published by the Free Software Foundation; version 2 of the License.

DISCLAIMER OF WARRANTY

This test collection is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of fitness for a particular purpose. See the GNU General Public License for more details. For a copy of the GNU General Public License, please write to the Free Software Foundation, Inc., 59 Temple Place, Suite 330, Boston, MA 02111-1307 USA

Support and contact

Please report any bugs and errors you encounter to us; we are continuously working to improve OWLS-TC! Errors can be reported by emailing the developers or preferably by using the online tracker system available at the project homepage at semwebcentral.org: http://projects.semwebcentral.org/tracker/?group_id=89

For OWLS-TC support and help, please contact

- Patrick Kapahnke at patrick.kapahnke@dfki.de
- Martin Vasileski at martin.vasileski@dfki.de

For general and scientific questions, please contact:

- Dr. Matthias Klusch at klusch@dfki.de, http://www.dfki.de/~klusch

Release history

- 10.11.2009 OWLS-TC version 3, revision 1
- 24.06.2009 Third release, OWLS-TC version 3
- 23.05.2008 OWLS-TC version 2.2, revision 2
- 17.12.2007 OWLS-TC version 2.2, revision 1
- 9.11.2007 OWLS-TC version 2.2
- 6.10.2006 OWLS-TC version 2.1
- 15.11.2005 Second release, OWLS-TC version 2
- 6.04.2005 Initial release, OWLS-TC version 1