Semantic Annotations for WSDL

Amit P. Sheth
Wright State University - Main Campus, amit.sheth@wright.edu

Jacek Kopecky

Follow this and additional works at: http://corescholar.libraries.wright.edu/knoesis

Part of the Bioinformatics Commons, Communication Technology and New Media Commons, Databases and Information Systems Commons, OS and Networks Commons, and the Science and Technology Studies Commons

Repository Citation
SAWSDL: Tools and Applications

Amit P. Sheth
Kno.e.sis Center
Wright State University, Dayton, OH
Knoesis.wright.edu

Key contributors: Karthik Gomadam, Kunal Verma, Ajith Ranabahu, Meena Nagarajan
What does Semantics bring to the table?

- **Better Reuse**
  - Semantic descriptions of services to help find relevant services

- **Better Interoperability**
  - Beyond syntax to semantics, mapping of data exchanged between the services (very time consuming without semantics, just as XML in WSDL gives syntactic interoperability, SAWSDL gives semantic interoperability)

- **Configuration/Composition**
  - Enable dynamic binding of partners

- **Some degree of automation across process lifecycle**
  - Process Configuration (Discovery and Constraint analysis)
  - Process Execution (Addressing run time heterogeneities and exceptions)
What can we support or demonstrate today

- API for handling SAWSDL documents: **SAWSDL4J**
- Tool for annotating WSDL services to produce SAWSDL: **Radiant** and for discovery: **Lumina**
- Using SAWSDL with UDDI for Discovery: **SemBowser**
- Using SAWSDL with Apache Axis for Data Mediation
- Using SAWSDL with WS-BPEL for run-time binding
- Early Examples of SAWSDL annotated services: biomedical research

Also:
- **Semantic Tools for Web Services** by IBM alphaWorks
- **WSMO Studio**, more mentioned by Jacek
Semantic Web Services Discovery: Illumia
Syntactic and Semantic Match do not suffice

DATA MEDIATION REQUIRED

Web service 1
Address Lookup

OUTPUT FROM WEB SERVICE 1
Address line 1
Address line 2
City_State_zip
Listing Name
First Name
Last Name
Address
City
State
Postal Code
Phone Number
Published

INPUT TO WEB SERVICE 2

Web service 2
Geocode Enhancer

OUTPUT FROM WEB SERVICE 2
Census Track
State Number
County Number
Block Number
Block Group

INPUT TO WEB SERVICE 1
Telephone Number
Mediation approach

- User specified mappings from Web service message element to semantic model concept (say OWL Ontology)
  - upcast : from WS message element to OWL concept
  - Downcast : from OWL concept to WS message element

```xml
<complexType name="Address">
  <sequence>
    <element name="StreetAd1" type="xsd:string"/>
    <element name="StreetAd2" type="xsd:string"/>
    ...........
  </sequence>
</complexType>
```

```xml
<POOntology:has_StreetAddress rdf:datatype="xs:string">
{ fn:concat($a/streetAddr1 , " ", $a/streetAddr2 ) }
</POOntology:has_StreetAddress>
```
<table>
<thead>
<tr>
<th>Heterogeneities / Conflicts</th>
<th>Examples - conflicted elements shown in color</th>
<th>Suggestions / Issues in Resolving Heterogeneities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain Incompatibilities – attribute level differences that arise because of using different descriptions for semantically similar attributes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naming conflicts</td>
<td>Web service 1: Student(Id#, Name)</td>
<td>Web service 2: Student(SSN, Name)</td>
</tr>
<tr>
<td></td>
<td>Web service 1: Student(Id#, Name)</td>
<td>Web service 2: Book (Id#, Name)</td>
</tr>
<tr>
<td>Data representation conflicts</td>
<td>Web service 1: Student(Id#, Name) Id# defined as a 4 digit number</td>
<td>Web service 2: Student(Id#, Name) Id# defined as a 9 digit number</td>
</tr>
<tr>
<td>Data scaling conflicts</td>
<td>Web service 1: Marks 1-100</td>
<td>Web service 2: Grades A-F</td>
</tr>
<tr>
<td><strong>Entity Definition – entity level differences that arise because of using different descriptions for semantically similar entities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naming conflicts</td>
<td>Web service 1: EMPLOYEE (Id#, Name)</td>
<td>Web service 2: WORKER (Id#, Name)</td>
</tr>
<tr>
<td></td>
<td>Web service 1: TICKET (TicketNo, MovieName)</td>
<td>Web service 2: TICKET (FlightNo, Arr. Airport, Dep. Airport)</td>
</tr>
<tr>
<td>Schema Isomorphism conflicts</td>
<td>Web service 1: PERSON (Name, Address, HomePhone, WorkPhone)</td>
<td>Web service 2: PERSON (Name, Address, Phone)</td>
</tr>
<tr>
<td><strong>Abstraction Level Incompatibility – Entity and attribute level differences that arise because two semantically similar entities or attributes are represented at different levels of abstraction</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalization conflicts</td>
<td>Web service 1: GRAD-STUDENT (ID, Name, Major)</td>
<td>Web service 2: STUDENT (ID, Name, Major, Type)</td>
</tr>
<tr>
<td>Aggregation conflicts</td>
<td>Web service 1: PROFESSOR (ID, Name, Dept)</td>
<td>Web service 2: FACULTY (ID, ProfID, Dept)</td>
</tr>
<tr>
<td>Attribute Entity conflicts</td>
<td>Web service 1: COURSE (ID, Name, Semester)</td>
<td>Web service 2: DEPT (Course, Sem, ...)</td>
</tr>
</tbody>
</table>

* Interoperation between services needs transformation rules (mapping) in addition to annotation of the entities and/or attributes indicating their semantic similarities (matching).
• Web services interoperate by re-using these mappings.
  – Ontologies now a vehicle for Web services to resolve message level heterogeneities
DM Architecture components

- **METEOR-S Middleware**
  - EPR handler – End Point Resolution handler
    - For clients to use the middleware
    - Reroute SOAP messages to middleware
  - DM handler – Data Mediation handler
    - Main component for facilitating data mediation
    - Works with the EPR handler + a mapping processing engine (SAXON for XQuery / XSLT)

- **Uses extensibility support offered by Axis 2 (handlers)**
Semantic Templates

- SAWSDL + Enhanced policy descriptions to model the data, functional and non-functional semantics at the various tiers
  - Business Process Tier: Capture process level requirements
  - Implementation Tier: Capture partner level requirements
- Non-functional semantics captured at template and operation levels.
- XML representation for interoperability.
Semantic Templates

- SAWSDL for data and functional semantics
- Semantic Policy Descriptions for non-functional semantics
Example of a semantic template in the supply chain domain

Semantic Template

- **ServiceLevelMetaData (SLM)**
  - Category: NAICS:Electronics
  - ProductCategory: DUNS:RAM
  - Location: Atlanta, GA

- **SemanticOperation Template (SOPT1)**
  - Action: Rosetta:RequestPurchaseOrder
  - Input: Rosetta:PurchaseOrder:Input
  - Output: Rosetta:PurchaseOrder:Output
  - OLP: (Encryption = RSA, ResponseTime< 5 Sec)

- **SemanticOperation Template (SOPT2)**
  - Action: Rosetta:CancelOrder
  - Input: Rosetta:CancelOrder:Input
  - Output: Rosetta:CancelOrder:Output
  - OLP: (Encryption = RSA, ResponseTime< 5 Sec)
Semantic Discovery

- Finds actual services matching semantic templates
- Implemented as a layer over UDDI
- Current implementation based on ontological representation of operations, inputs and outputs.
- Returns ranked of services for each semantic template
USING SAWSDL WITH WS-BPEL FOR RUN-TIME BINDING
Dynamic configuration Problem

Find optimal partners for the process based on process constraints – cost, supply time, etc.

Conceptual Approach

1. Create framework to capture represent domain knowledge
2. Represent constraints on the domain knowledge
3. Ability to reason on the constraints and configure the process
Dynamic Binding: Guiding principles

- Semantic templates to capture the requirements for each partner.
- Partners are selected during the run time of the process and the process is configured
  - Semantically Enhanced UDDI Registries for discovery of partners.
  - Approaches to match enhanced policies (Sem-Pol) and agreements (SWAPS)
- Execution environment supporting discovery, configuration and invocation.
Example of a process with semantic templates

Semantic Template 1 (ST1)
- ServiceLevelMetaData (SLM)
  - Category: NAICS: Electronics
  - ProductCategory: DUNS: MB
  - Location: Athens, GA
  - Action: getQuote
  - OLP: {Encryption = RSA, ResponseTime<5 Sec}
  - Action: Order
  - OLP: {Encryption = SHA1, supplyTime<5 days}
  - Action: Cancel
  - OLP: {Encryption = RSA, Penalty<25%}

Semantic Template 2 (ST2)
- ServiceLevelMetaData (SLM)
  - Category: NAICS: Electronics
  - ProductCategory: DUNS: RAM
  - Location: Athens, GA
  - Action: getQuote
  - OLP: {Encryption = RSA, ResponseTime<5 Sec}
  - Action: Order
  - OLP: {Encryption = SHA1, supplyTime<4 days}
  - Action: Cancel
  - OLP: {Encryption = RSA, Penalty<20%}

Semantic Template 3 (ST3)
- ServiceLevelMetaData (SLM)
  - Category: NAICS: Electronics
  - ProductCategory: DUNS: Processor
  - Location: Athens, GA
  - Action: getQuote
  - OLP: {Encryption = RSA, ResponseTime<5 Sec}
  - Action: Order
  - OLP: {Encryption = SHA1, supplyTime<4 days}
  - Action: Cancel
  - OLP: {Encryption = RSA, Penalty<15%}
Semantic Biological Web Services Registry

STARGATE Web services discovery using task name

This allows the user to search for Web services with given 'Task Name'. This requires an exact match between the user defined term and task concept of the service.

Task Name: Raw to mzXML

Submit
Reset

STARGATE

Welcome to SemBROWSER, Semantic Biological Web Services Registry

Name: SysJavaRawmzXMLService
WSDL Location: http://192.168.168.100:8080/axis2/PWF/SysCommv0.21/SysJavaRawmzXML.jws
Business Entity: glycomics
Semantic Web Services

- Formalize description and classification of Web Services using ProPreO concepts

```xml
<?xml version="1.0" encoding="UTF-8"?>
<wsdl:definitions targetNamespace="urn:ngp"
    xmlns:wssem="http://www.ibm.com/xmlns/WebServices/WSSemantics"
    xmlns:ProPreO="http://lsdis.cs.uga.edu/ontologies/ProPreO.owl">
    <wsdl:types>
        <schema targetNamespace="urn:ngp" xmlns="http://www.w3.org/2001/XMLSchema">
            ...
        </schema>
    </wsdl:types>
    <wsdl:message name="replaceCharacterRequest">
        <wsdl:part name="in0" type="soapenc:string"/>
        <wsdl:part name="in1" type="soapenc:string"/>
        <wsdl:part name="in2" type="soapenc:string"/>
    </wsdl:message>
    <wsdl:message name="replaceCharacterResponse">
        <wsdl:part name="replaceCharacterReturn" type="soapenc:string"/>
    </wsdl:message>
</wsdl:definitions>
```
ISiS – Integrated Semantic Information and Knowledge System

Semantic Web Process to incorporate provenance

- Biological Sample Analysis by MS/MS
- Raw Data to Standard Format
- Data Pre-process
- DB Search (Mascot/Sequest)
- Results Post-process (ProValt)

Raw Data → Standard Format Data → Filtered Data → Search Results → Final Output

Semantic Annotation Applications

Storage

Biological Information

Knowledge Enabled Information and Services Science
• Evaluate the specific effects of changing a biological parameter: Retrieve abundance data for a given protein expressed by three different cell types of a specific organism.

• Retrieve raw data supporting a structural assignment: Find all the raw ms data files that contain the spectrum of a given peptide sequence having a specific modification and charge state.

• Detect errors: Find and compare all peptide lists identified in Mascot output files obtained using a similar organism, cell-type, sample preparation protocol, and mass spectrometry conditions.

A Web Service Must Be Invoked

ProPreO concepts highlighted in red
Some Relevant Papers


Stargate Portal: SemBowser and example SAWSDL service:
http://glycomics.ccrc.uga.edu/stargate/index.jsp