PI: Dr. Paul Kogut

Yui Leung, Ted Mielczarek, Kathleen Ryan, Linda Gohari, Roger Lee

Key Researchers:

Dr. Jeff Heflin – Lehigh University

Dr. Mitch Kokar, Dr. Chris Matheus, Dr. Ken Baclawski - VIS/Northeastern University

Dr. Richard Waldinger - Kestrel/SRI

2004 Research Thrusts

May 26, 2004
Lockheed Martin contributions to 2004 DAML Program thrusts:

- Mature OWL tools
  - AeroSWARM - OWL markup generation service
  - ConsVISor and BugVISor – OWL consistency checking/debugging
  - DLDB – hybrid semantic web/relational database reasoning infrastructure

- Semantic Web Services
  - C4ISR service discovery and composition experiments

- OWL standardization support
  - Formally verified OWL axiomatic semantics
Mature OWL Tools

- AeroSWARM
  - Use cases – markup pages for posting or ingestion into KB
  - 44 common properties (vs. 6 in 2003)
  - Web service on load-balanced servers for integration with other tools

- ConsVISor
  - Easy to use web-based tool for checking ontologies and markup
  - Full XSD support
  - Support for debugging ontologies via OWL symptom ontology

- DLDB
  - Scaleable open source infrastructure
    - Quantitative evaluation of DLDB, Sesame, OWLJessKB
  - Tools/techniques for benchmarking OWL applications
Integrated Tool Architecture

- Web Page URL
- Web Service Invocation
- OWL Ontologies
- Markup Creator
- OWL Markup
- AeroText™
- Website Crawler
- OWL-S Markup
- Ontology Mapping Tool
- Predefined Ontology Mappings
- Markup Converter
- AeroText™
- Consistency Reasoning Service
- BugVISor
- OWL Ontologies
- Ontology Engineer
- Consistency checking results

http://ubot.lockheedmartin.com
Student --> Person who takes courses
Graduate Student --> person who takes graduate courses
Graduate course ∈ Course

CREATE VIEW Student_1_view AS
SELECT * FROM Student_1 UNION
SELECT * FROM UndergraduateStudent_1_view UNION SELECT * FROM GraduateStudent_1_view;$
Lehigh University Benchmark

**Synthetic Data Generation**
- data generated in unit of university
- random
  - class instances and property instances are randomly decided
- “Realistic”
  - some constraints are applied
- customizable
  - can select # of univ, start index, and seed
- arbitrary size
- repeatable
Mature Tools - Experiments

How do we demonstrate robustness of tools?

by applying them to hard knowledge management problems in a realistic web context:

- cross document co-reference
  - Is it plausible that Al Smith in document 1 is the same as Al Smith in document 2?
- GOWLgle – Google results filtering
Problem: When collecting information about a person or an organization need to check if assertions are referring to same entity

- linguistic clues do not work – need reasoning

Approach:

- assert that entity X sameAs entity Y
- apply logical reasoning to check if all assertions about these entities in set of documents and constraints in the ontology are consistent
  
  ◆ If consistent then co-reference is plausible
  ◆ If not then co-reference is not plausible
Problem: Google does not use semantics and there is not enough OWL content yet – need hybrid information retrieval techniques

Approach:
- user chooses semantic keywords and relations
- markup web pages in list of N best Google results
- filter out pages that do not match semantic keywords
C4ISR OWL-S Experiments

- OWL-S for Net-centric warfare
  - Develop OWL-S descriptions of current and future Air Force, Navy, Army and Intel systems/services – identify KR issues
  - Experiment with OWL-S discovery and composition approaches – identify OWL-S tool/architecture issues
## Deliverables

<table>
<thead>
<tr>
<th>Deliverable</th>
<th>IP</th>
<th>SemWebCentral</th>
</tr>
</thead>
<tbody>
<tr>
<td>AeroSWARM</td>
<td>Open service</td>
<td>Currently registered</td>
</tr>
<tr>
<td>ConsVISor/BugVISor</td>
<td>Open service</td>
<td>Currently registered</td>
</tr>
<tr>
<td>DLDB</td>
<td>Open source</td>
<td>June 2004</td>
</tr>
<tr>
<td>OWL axiomatic semantics</td>
<td>Open source</td>
<td>June 2004</td>
</tr>
</tbody>
</table>

- **Open services:**
  - **Advantages:**
    - No need to download and install
    - Use expensive software/hardware infrastructure for free
  - **Disadvantages:**
    - Limited customization options


**Plans**

- **In 2004**
  - Wrap-up current tool development efforts
    - Refine AeroSWARM based on SemWebCentral feedback
    - DLDB inference and query interface enhancements
  - Continue OWL-S experiments in Net-centric warfare

- **In 2005 we would like to:**
  - Help transition OWL-S to DoD for Net-centric warfare
  - Refine GOWLgle into deliverable open service to show value added of Semantic Web
  - Refine co-reference into deliverable service integrated with AeroSWARM
  - Semi-automatic generation of customized benchmarks
  - Develop community symptom ontology and bug ontology