Aligning Ontologies using Multi Agent Technology

Gihan M. Rajakaruna
Asoka S. Karunananda
Shantha Jayalal

ICTer - 2012
14th December 2012
Overview

- Introduction
- Major Problem
- Aim & Objectives
- Existing Solutions for Ontology Alignment
- Technologies Adapted
- Multi-Agent Based Approach
- Multi-Agent System Design
- OntoMAS System Development
- Evaluation Results
- Conclusion
Introduction

- Modern information systems extensively use ontologies to model domain knowledge.
- Ontologies represent the meaning of the data and could be processed by both humans and machines.
- High demand for sharing and reusing the knowledge in existing ontologies.
Introduction...

- Ontologies are complex structures and thus, sharing of knowledge coming from various ontologies has become a tedious task.
- This resulted in a research area called ontology alignment.
Introduction...

- Definition of Ontology Alignment
  - Exploring the semantic relationships between two existing ontologies
  - Used to create a common vocabulary that could facilitate the interoperability across multiple heterogeneous information systems
Introduction…

- Definition of Ontology Alignment…
Major Problem

- Ontology Alignment process takes considerable amount of man hours, and it is being a tedious task for domain experts.
- Alignment decisions involves several factors such as structure of the ontologies, linguistic meaning of the terms etc.
- Any ontology alignment solution should also consider these different factors when making accurate alignment decisions.
Aim and Objectives

The aim of the project is to develop a system for ontology alignment with the use of multi-agent system technology

Main Objectives:

- Study of existing solutions for ontology alignment
- Study and selection of ontology aligning techniques and tools
- Develop an approach to ontology alignment by using agent communication, coordination, and negotiation
- Develop a prototype for emulating the approach
- Evaluate the approach using the prototype
Existing Solutions for Alignment

- Major research area in Artificial Intelligence
  - Numerous techniques have been developed
  - Three main categories:
    - Human Driven Techniques
    - Domain Dependent Techniques
    - Generic Techniques
## Existing Solutions…

<table>
<thead>
<tr>
<th>Name of the system</th>
<th>Category</th>
<th>Technology/Theory used</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROMPT</td>
<td>Human Driven</td>
<td>Class name similarities</td>
</tr>
<tr>
<td>PROMPTDIFF</td>
<td>Human Driven</td>
<td>Heuristic matchers and a fixed-point algorithm</td>
</tr>
<tr>
<td>SAMOA</td>
<td>Human Driven</td>
<td>Integrate the algorithms used in PROMPT, FORM etc…</td>
</tr>
<tr>
<td>Ontology Alignment Using Rough Sets</td>
<td>Domain Dependent</td>
<td>String based, linguistic based, and structural based matchers with rough sets</td>
</tr>
<tr>
<td>IF-Map Algorithm</td>
<td>Domain Dependent</td>
<td>Barwise-Seligman theory of Information Flow and Kent’s virtual ontology of community connections</td>
</tr>
<tr>
<td>Memetic Algorithm</td>
<td>Domain Dependent</td>
<td>Generic Algorithms and Hill Climbing search</td>
</tr>
<tr>
<td>PSO Algorithm</td>
<td>Domain Dependent</td>
<td>Particle Swarm Optimization</td>
</tr>
<tr>
<td>Ontology Alignment Using Reference Ontology</td>
<td>Domain Dependent</td>
<td>probability distribution measures and matrices</td>
</tr>
<tr>
<td>DSSim</td>
<td>Generic</td>
<td>Multi-Agent, Fuzzy Logic</td>
</tr>
<tr>
<td>RiMOM</td>
<td>Generic</td>
<td>Adaptive strategy using Similarities of labels and structure</td>
</tr>
<tr>
<td>Ontology Alignment Using Upper Ontologies</td>
<td>Generic</td>
<td>Use an global ontology as a semantic bridge</td>
</tr>
</tbody>
</table>
Existing Solutions...

Summary

- Most of the existing techniques are domain dependent and require considerable amount of human interaction

- Very few generic solutions exist but, those algorithmic approaches couldn’t handle the complexity with diversify ontologies
Technologies Adapted

- Protégé Ontology Modeling Environment
  - Capability to create, edit, visualize, and reasoning on ontologies
  - Has a flexible modular design

- Multi-Agent System Technology
  - JADE - Java Agent Development Environment
  - Assures effective communication between the agents
Technologies Adapted…

- **WordNet**
  - Lexical database for English language

- **Upper Ontologies**
  - Cover large amount of concepts in numerous domains
    - Ex: Cyc, DOLCE, SUMO
Multi-Agent based Approach

Hypothesis

- Ontology alignment problem can be solved using Multi-Agent system technology, in the manner of communication, coordination, and negotiation
- Inspired by Marvin Minsky’s idea of the Society of Mind
Multi-Agent based Approach…

- Users
  - Semantic web application developers
  - Ontology modeling engineers

- Inputs
  - Ontologies in RDF, RDFS, or OWL formats

- Output
  - Generates a common ontology in OWL format
Multi-Agent based Approach…

Process

- User choose a base ontology
- Assigns an agent to represent each and every concept in input ontologies
- Initiate the agents to perform the matching
- Let the agents communicate, coordinate, and negotiate to explore the relationships
- Generate the output ontology by considering the explored results
Multi-Agent based Approach...

Features

- Non-algorithmic approach
  - No predefined algorithms
- More error tolerable
  - Most of the incorrect decisions were ignored in the final output
- Efficient and scalable
- Runs on protégé ontology modeling environment
Multi-Agent System Design

- Two main modules to convert the inputs to output
  - Agent Module
  - Knowledge Module

- System is Designed based on the Request-Resource-Message Space-Ontology architecture
Multi-Agent System Design...

- System architecture
Multi-Agent System Design...

- Agent Module
  - Consists of six types of agents
    - Alignment Request Agent
    - Concept Resource Agent
    - String Matching Resource Agent
    - Structure Matching Resource Agent
    - Linguistic Matching Resource Agent
    - Upper Ontology Matching Resource Agent
Multi-Agent System Design...

- **Alignment Request Agent**
  - Represents a concept of the base ontology
  - Initiates the process by broadcasting a request message to all concept resource agents
  - Knows only about the own concept details
  - When received the responses updates the Protégé interface
Multi-Agent System Design…

- Concept Resource Agent
  - Represents a concept of the non-base ontology
  - Listens to the responses from Alignment Request Agents and response
  - Coordinates with matching resource agents
Multi-Agent System Design…

- **String Matching Resource Agent**
  - Accepts pairs of concepts as inputs
  - Explores the shared substring, common prefixes, common suffixes etc.
  - Negotiates with Linguistic Matching Resource Agent
  - Sends the final result to the Concept Resource Agent
Multi-Agent System Design…

- Linguistic Matching Resource Agent
  - Accepts pairs of concepts as inputs
  - Uses WordNet database to explore the synonym and antonym information
  - Negotiates with String Matching Resource Agent
  - Sends the final result to the Concept Resource Agent
Multi-Agents System Design...

- Structure Matching Resource Agent
  - Accepts pairs of concepts as inputs
  - Explores the sub class, super class relationships
  - Negotiates with Upper Ontology Matching Resource Agent
  - Sends the final result to the Concept Resource Agent
Multi-Agent System Design…

- Upper Ontology Matching Resource Agent
  - Accepts pairs of concepts as inputs
  - Explores the upper ontologies to find the concept similarity
  - Negotiates with Structure Matching Resource Agent
  - Sends the final result to the Concept Resource Agent
Multi-Agent System Design…

- Knowledge Module
  - Consists of following knowledge sources
    - Rules for Agents to Operate
    - Input Ontologies
    - Linguistic Information
    - Upper Ontologies
    - Generated Results
OntoMAS System Development

- Agent Module Development with JADE
  - Agent Development
    - Agents are developed as Java Classes
  - Agent Initialization
    - Similar to creating an object from any Java class by using the `new` keyword
  - Agent Business Logic
    - Doesn’t contain any complex business logic
OntoMAS System Development...

- Agent Module Development with JADE...
  - Agent Communication
OntoMAS System Development

- Plugin Development for Protégé Environment
  - Protégé plugins are implemented as Java classes
  - Uses Protégé-OWL API to access the core functionality
  - Introduced Menu Item, Wizard, View, and Tab
  - Can be distributed as a single Jar file
Evaluation

Evaluation Strategy

- Compare the generated results against the actual relationships between the two input ontologies
- Four parameters observed
  - **TP** (True Positive) - True relationships correctly identified by OntoMAS
  - **FP** (False Positive) – Incorrect relationships marked by OntoMAS
  - **FN** (False Negative) – Correct relationships missed by OntoMAS
  - **TN** (True Negative) – False relationships correctly identified by OntoMAS
Evaluation...

- Relationships between the Parameters

```
<table>
<thead>
<tr>
<th>Actual Relationships</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
</tr>
<tr>
<td>FP</td>
</tr>
<tr>
<td>TN</td>
</tr>
<tr>
<td>FN</td>
</tr>
</tbody>
</table>

- TP: Correctly Identified Matching’s
- FP: Incorrectly Identified Matching’s
- TN: Correctly Identified none Matching’s
- FN: Missing Matching’s

Generated Relationships
```
Evaluation…

- Calculate the precision, recall, and accuracy

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

\[
\text{Recall} = \frac{TP}{TP + FN}
\]

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]
Evaluation...

- OntoMAS on Conference Organizing Domain

<table>
<thead>
<tr>
<th>Ontology</th>
<th>Classes</th>
<th>Data Type Properties</th>
<th>Object Type Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDAS</td>
<td>104</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>ACM-SIGKDD</td>
<td>49</td>
<td>11</td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Actual Generated</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>16</td>
<td>70</td>
</tr>
<tr>
<td>False</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>
Evaluation...

- OntoMAS on Conference Organizing Domain...

\[
\text{Precision} = \frac{16}{16 + 8} = \frac{16}{24} = 0.667
\]

Precision = 66.7 \%

\[
\text{Recall} = \frac{16}{16 + 10} = \frac{16}{26} = 0.615
\]

Recall = 61.5 \%

\[
\text{Accuracy} = \frac{16 + 70}{16 + 8 + 70 + 10} = \frac{86}{104} = 0.827
\]

Accuracy = 82.7 \%
Evaluation...

- OntoMAS on Agricultural Domain
  - WikiGoviya ontology has 54 concepts, while HARTI ontology has 38 concepts

<table>
<thead>
<tr>
<th>Actual</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>True</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>False</td>
<td>9</td>
<td>13</td>
</tr>
</tbody>
</table>

Precision = \( \frac{11}{11 + 9} = \frac{11}{20} = 0.55 \)

Precision = 55.0%

Recall = \( \frac{11}{11 + 13} = \frac{11}{24} = 0.458 \)

Recall = 45.8%

Accuracy = \( \frac{11 + 21}{11 + 9 + 21 + 13} = \frac{33}{54} = 0.611 \)

Accuracy = 61.1%
Conclusion

- Generally, OntoMAS achieved considerable amount of precision, recall, and accuracy on both ontological domains.

- But, OntoMAS has shown better results with Conference Organizing domain.

- All stated objectives were achieved.
Conclusion…

- **Limitations**
  - Use of WordNet and Upper Ontologies are effective only on more structured domains

- **Future Directions**
  - Consider other OWL properties during the alignment
  - Use the proposed solution in practical applications
Thank You!