Faculty of Computer Science

University of New Brunswick

Thesis Proposal for MCS Degree Program

Design and Implementation of a Distributed Query System
Supporting Conference Organization Committees

By

Chaudhry Usman Ali

Supervisors: Dr. Harold Boley, Dr. Weichang Du

September 30, 2011


1 Introduction

When studying conference organization systems online, it is obvious that conference organization involves a multitude of procedures, consuming much time and effort of the organizing body. Individual tasks which might not seem laborious at first, bundled together, form a formidable challenge for any Organizing Committee (OC) to tackle. Conference organization typically involves organization partner coordination, sponsoring correspondence, panel participants management, etc. The importance of conference organization systems cannot be overemphasized as they attempt to alleviate the burden of repetitive tasks through the automation of rule-based organization processes. A careful review of the conference management systems shows various systems developed for different purposes.

- Non-Query-Based Conference Organization Systems

  - Manuscript Central
    Manuscript Central \(^1\), developed by ScholarOne, Inc., is the online submission and peer review system used to handle manuscript submissions to journals. This system is currently used by most IEEE and Association of Computing Machinery (ACM) journals. The system is used by authors to upload papers and have them reviewed.

  - Microsoft Research Conference Management Tool (MSRCMT)
    Firstly developed for ACM SIGKDD 1999, the MSRCMT \(^2\) is an academic conference management service sponsored by Microsoft Research. Surajit Chaudhury, a Research Area Manager at Microsoft Research is the architect of MSRCMT. Similar to Manuscript Central, the MSRCMT is also a fully-developed system. It is free and hosted by Microsoft Research, but with

---

\(^1\)http://scholarone.com/products/manuscript/
\(^2\)http://cmt.research.microsoft.com/cmt/
limited support, since it is developed and managed by a small team.

- EasyChair
  EasyChair ³ is capable of supporting two models: (1) the standard model for conferences having one program committee and (2) the multi-track version for conferences having multiple tracks that have their own program committee.

- WitanWeb
  WitanWeb ⁴ is a system designed to allow for the online submission, referring and decision making of proposals. The proposals could be papers submitted to a referred publication, grant proposals, award nominations, etc.

- Query-Based Conference Organization Systems
  - SymposiumPlanner-2010
    After three earlier SymposiumPlanner systems, SymposiumPlanner-2010 has been implemented using a combination of EMERALD ⁵ and Rule Responder ⁶. EMERALD is a multi-agent knowledge-based framework, which offers flexibility, reusability and interoperability of behavior between agents, based on Semantic Web and FIPA language standards. The main advantage of this approach is that it provides a safe, generic, and reusable framework for modeling and monitoring agent communication and agreements. Rule Responder is a tool for creating virtual organizations as multi-agent systems that support collaborative teams on the Semantic Web.

  - SymposiumPlanner-2011
    After SymposiumPlanner-2010, the 2011 instantiation is im-

³http://www.easychair.org/
⁴http://witanweb.ca/cascon2010/WitanWebFAQ.jsp
⁵http://lpis.csd.auth.gr/systems/EMERALDRR/
⁶http://ruleml.org/RuleResponder/
implemented using Rule Responder. The SymposiumPlanner-2011 has been implemented for organizing two conferences i.e. RuleML-2011@BRF  and RuleML-2011@IJCAI. Distributed users (humans or agents) can interact with Rule Responder by query answering conversations [CB08] based on an Enterprise Service Bus (ESB). Rule Responder agents will process events, queries and requests according to their rule-based decision and behavioural logic. It can also delegate subtasks to other agents, collect answers, and send the validated answer(s) back to the requester. Since the Rule Responder framework has been conceived, many instantiations of it have been developed in areas as diverse as Service Level Management, Business Process Management, Symposium Planning, and Health Care systems [BP11]. The main benefit from such use is the alleviation of the burden of repetitive tasks through the automation of rule-based organization processes.

Conference organization involves various procedures which make the task difficult both for the users and the developers of such systems. Tasks like responsibility assignment, contact information retrieval, and partner management are complex, and usually involve redundant transactions. One of the foremost challenges of developing systems to perform these tasks is the necessary trade-off between load balancing and speed. A request overload can burden the system and the speed of response of a system to a query can make the difference in attracting more organizations to use the system. For this purpose, it is necessary to ensure that the system does not spend too much time on one query, and a powerful enterprise service bus capable of supporting different protocols can help transport messages among disparate knowledge bases without delay and error. Another technical challenge

is to provide a user friendly user client, where users can query the system easily without getting confused, and retrieve relevant information.

Query based systems are receiving more attention lately because they can simulate a real world organization by responding to queries from users. A system capable of interacting with Enquiry Users as well as participants can provide an intelligent solution to perform the redundant tasks of an organization. EMERALD depends on external rule base and knowledge base. SymposiumPlanner uses local rule base and external semantic knowledge repositories on the Internet such as DBpedia, Freebase, YAGO, Semantic Web Dog Food to avoid redundancy in knowledge bases. However, SymposiumPlanner -2011 has added complexity to cater to requirements of the multiple events related to RuleML Symposia. This effects system performance as well as making it harder to maintain. The system also suffers from some faults in design which can be fixed to help it perform to its full potential.

Earlier conference support systems include SymposiumPlanner-2007 through -2011 series of Rule Responder instantiations\(^9\), which focus on question answering as part of the official websites of the RuleML Symposia. The SymposiumPlanner instantiations span various implementations from initial versions in 2007, 2008, and 2009, to the SymposiumPlanner-2010 (in EMERALD and Rule Responder) instantiation based on EMERALD (hence Jade) [KOBB11], to the 2011 double-instantiation using the latest versions of Mule and Prova, as well as a more user friendly interface involving Suborganizational Agents for covering both sites.

This thesis will design and implement, in Rule Responder, a distributed query system supporting symposium Organization Committees (OCs). Our improved SymposiumPlanner design and implementation will be explored in a use case, the RuleML-2012 Symposium. Its support will build on the earlier SymposiumPlanners and an extensive

\(^9\)http://ruleml.org/SymposiumPlanner
literature survey, distilling design principles and implementation techniques for future SymposiumPlanners from the lessons learned so far, as started in [ZPAB11]. In order to evaluate the SymposiumPlanner-2012 system, we will solicit feedback from the emerging RuleML-2012 OC, which will lead to continuous revisions of the knowledge bases, and perhaps the SymposiumPlanner-2012 system architecture as an instantiation of Rule Responder. This will then allow further improved designs and implementations for the SymposiumPlanner systems of 2013 and beyond.

2 Background & Related Work

Rule Responder is a service-oriented middleware tool that can be used by virtual organizations for automated rule-based collaboration
Distributed users (humans or agents) can interact with Rule Responder by query answering conversations [CB08] based on an Enterprise Service Bus (ESB). Rule Responder’s architecture realizes a system of personal agents (PAs) and organizational agents (OAs), accessed by external agents (EAs), on top of an Enterprise Service Bus (ESB) communication middleware [CB08]. The rule-based PAs represent, as their ‘dynamic profiles’, all of the participating human members of the virtual organization modelled by the Rule Responder. An OA constitutes an intelligent query dispatching system, using a rule engine execution environment for selectively delegating queries to one of the PAs.

2.1 Approaches to Knowledge Representation

In Rule Responder we make use of two paradigms of knowledge representation: Ontologies and rules. The OA uses a Responsibility Assignment Matrix to delegate incoming queries to particular PAs based on their responsibilities. This is achieved through the use of an ontology (OWL Lite). The Organizational Agent is implemented using the Prova Semantic Web rule engine. The Organizational Agent uses reaction rules that describe its policies, regulations, opportunities, etc. Each Personal Agent uses derivation rules to answer queries relevant to the supported chair’s role. Hence, our framework uses a combination of rules and ontologies.

2.2 Hierarchical Distributed Architecture

A star topology (hierarchy) constitutes simple but effective distributed topology. In our SymposiumPlanner, the star topology connects the specialized PAs ‘spokes’ to the centralized OA ‘hub’. A star topology is partially fault tolerant: If a spoke is broken then it will not
affect the rest of the spokes. However, if the hub is broken then the system ceases to function.

2.3 Organizational Agent

An Organizational Agent is used to describe the goals of the symposium as a whole and contains a knowledge base that describes the symposium’s policies, regulations, and opportunities. This knowledge base contains condition/action/event rules as well as derivation rules. An OA manages its local PAs, providing control of their life cycles and ensuring overall goals and policies of the organization and its semiotic structures [BP11].

2.4 Personal Agents

In the SymposiumPlanner system, each organization committee chair is designed as a Personal Agent, which contains a knowledge base that represents its chair’s responsibilities to answer corresponding queries. Personal agents are chair’s roles in the symposium organization. Personal Agents contain FOAF-extending profiles for each person of the organizational team.

2.5 Communication Among Agents

A lot of the early ESB products had a history in the enterprise application integration (EAI) market. We can identify two major differences between enterprise service bus (ESB) and EAI products. The first is the change from the hub-and-spoke model in EAI products to a bus-based model in ESB products. The hub-and-spoke model is a centralized architecture, where all data exchange is processed by a hub. The bus model uses a distributed architecture, in which the ESB functionality can be implemented by several physically separated functions.

A second difference between EAI and ESB products is the use of open standards.
Mule open-source ESB allows the deployment of rule-based agents on the Mule object broker and supports the communication in this rule-based agent processing network via a multitude of transport protocols. Several agent services which at their core run a rule engine are installed as Mule components which listen at configured endpoints, e.g., JMS message endpoints, HTTP ports, SOAP server/client addresses or JDBC database interface. The large variety of transport protocols provided by Mule can be used to transport the messages to the registered endpoints or external applications/tools. Usually, JMS is used for the internal communication between distributed agent instances, while HTTP and SOAP is used to access external Web services. The use of Mule allows architectural flexibility by decoupling the functional components of Rule Responder from the communication components [Cha04].
2.6 Interchange Languages and Rule Engines

For SymposiumPlanner System, we use Reaction RuleML as our interchange language between agents. Reaction RuleML\textsuperscript{11} is a general, practical, compact and user-friendly XML-serialized sublanguage of RuleML. The SymposiumPlanner System uses two representative rule engines, namely Prova\textsuperscript{12} and OO jDREW\textsuperscript{13}. Prova is mostly used to process reaction rules. OO jDREW is used to process derivation rules. There exists a bi-directional translator between RuleML/XML and POSL\textsuperscript{14}(human-readable, Prolog like syntax). The POSL language is faster to write and easier to read than any XML syntax.

3 Thesis Objectives

Symposium organization typically involves submission reviewing, organization partner coordination, sponsoring correspondence, panel participants management, etc. Our overarching objective is to design, implement and evaluate a distributed query system supporting symposium Organization Committees (OCs). This is achieved by the following specific objectives.

1. Our system will distill design principles and implement techniques for SymposiumPlanner-2012 from the lessons learnt so far.

2. Our system will provide Personal Agents which can assist the users and participants of a virtual organization. The Personal Agents are semi-autonomous in nature; they can provide help in dealing with routine tasks. The Personal Agents use knowledge bases that encode knowledge of each human chair.

\textsuperscript{11}http://reaction.ruleml.org
\textsuperscript{12}http://prova.ws
\textsuperscript{13}http://ruleml.org/oojdrew/
\textsuperscript{14}http://ruleml.org/submission/ruleml-shortation.html
3. Our system will allow different rule engines to execute global and local rulebases.

4. Our system will allow rules as well as queries and their answers to be transmitted over an Enterprise Service Bus (ESB).

5. As a demonstration of our methodology, our deployed system will support the Question&Answer (Q&A) parts of the official websites for the RuleML Symposia.

6. In order to evaluate our system, as we will solicit feedback from the real-world OC. This practice can result in continuous revisions of the knowledge bases of the Personal Agents, the Organizational Agent and perhaps even the system architecture.

4 Proposed Work

The main approach of this thesis is as follows: We will build a new Rule Responder instantiation, SymposiumPlanner-2012 for conference planning.

- **Architecture Design**
  - In our system the agent communication will be achieved in a star-like (hierarchical) manner (see Figure 3). In particular, when an Enquiry User asks a question to the organization via the External Agent, it does not know any Personal Agent such as the one that ultimately might answer the question. The query will be forwarded to the Organizational Agent, which will then delegate it to an appropriate Personal Agent.

- **Ontology Design and Rules**
  - Our system will employ both ontologies and rulebases. The use of an ontology for the Responsibility Assignment Matrix (OWL Lite) and rulebases in the Organizational Agent and Personal Agents
gives the framework a hybrid nature with both forms of knowledge representation.

- Implementation of the System:

  1. We will use the Mule open-source ESB to handle message-based interactions between the rule engines PROVA and OO jDREW. JMS (Java Message Service) will be used for internal communication between distributed agent instances while the HTTP and SOAP will be used to access external Web services.

  2. Our system will include the Prova rule engine to execute the Knowledge Bases (KBs) rules for the Organizational Agent. Reaction RuleML will be used by PROVA and act as an interchange language between distributed rule agents in the SymposiumPlanner.

  3. A single Organizational Agent will be implemented in PROVA to deliver and filter queries to the Personal Agents. The se-
lection logic for the dissemination of queries to PAs will be described by Responsibility Assignment Matrix implemented using an OWL Lite Ontology.

4. PAs represent a committee member including the following ones:
   - General Chair
   - Program Chair
   - Publicity Chair
   - Liaison Chair
   - Panel Chair

The Positional-Slotted Language (POSL), will be used to build various profiles of Personal Agents. Personal Agents for each OC member will be implemented in OO jDREW having their own rule-based decision logic on top of their personal information sources. The OO jDREW rule engine will be used by each of the Personal Agents.

5. An External Agent will be implemented to provide a user the Web interface to interact with the system and use queries to find information regarding the Symposium.

6. Our system will integrate existing factual information from the hosting conference on the Internet. This will help us avoid redundancy in the knowledge bases of the SymposiumPlanner agents.

- SymposiumPlanner-2012 Use Case

In the development of the SymposiumPlanner-2012 use case we will be taking a proactive approach. We are designing, implementing, and deploying the system well in advance to the actual August 2012 symposium collocated with ECAI-2012, so that OC roles will be well-prepared, based on the 2007-2011 experience. ‘Corporate’
know-how of the RuleML directors will be available to the thesis developer. The human RuleML-2012 chairs can then fully benefit from the prepared PA knowledge bases, building on and modifying/extending the encoded knowledge. This will also help the OC chairs to select and possibly change roles if they fit into another role in a more efficient manner. This gradual adoption will serve to evaluate and tune the system before it is utilized it for participant query answering.

Our Web interface should allow participants to issue queries via web forms, which will generate both RuleML/XML and controlled natural language. By reusing and integrating existing factual information about hosting ECAI-2012 conference on the Internet, we envision avoiding redundancy in the knowledge bases of the SymposiumPlanners agents\textsuperscript{15}. Since there will be only one symposium site in 2012, we will need to reunify the 2-SubOrganizational Agents’ business logic of SymposiumPlanner 2011 into one Organizational Agent.

In order to evaluate the SymposiumPlanner-2012 system, we will solicit feedback from the emerging RuleML-2012 OC, which will lead to continuous revisions of the knowledge bases, and perhaps the SymposiumPlanner-2012 system architecture as an instantiation of Rule Responder. This will then allow further improved designs and implementations for the SymposiumPlanner systems of 2013 and beyond.

One of the main advantages of SymposiumPlanner is that it answers user’s queries and reduces user’s burden of finding the relevant information by themselves. The queries include the information about the Symposia. For its usability, the SymposiumPlanner

\textsuperscript{15}There are large semantic data repositories on the Internet, such as: DBpedia (http://de.dbpedia.org), Freebase (http://www.freebase.com), YAGO (http://www.mpi-inf.mpg.de/yago-naga/yago), and Semantic Web Dog Food (http://data.semanticweb.org)
Figure 4: SymposiumPlanner Instantiations

user client provides an interface to distributed Personal Agents, allowing users to query the available interfaces, describe and submit the queries, and retrieve the answers from a standard Web browser for the External Agent.

SymposiumPlanner user client allows users to query the SymposiumPlanner agents via the SymposiumPlanner interface [FKS06].

● Evaluation of the System

In order to evaluate our system, implementation will be subjected to testing in the following scheme,

– Mule ESB Testing

* Level One: Unit Testing

JUnit is a well-known test framework and is the industry standard to implement unit tests for Java applications. For message flows developed with Mule, we can unit test the system with various inputs. A sample output from a JUnit test is as follows,

```java
import org.mule.umo.*;
import org.mule.tck.FunctionalTestCase;
import org.mule.extras.client.MuleClient;
import org.mule.providers.NullPayload;

public class MuleServerTest extends FunctionalTestCase {
```
public MuleServerTest() {
    super();
}

protected String getConfigResources() {
    return "mule-config.xml";
}

public String getName() {
    return "Mule Server Test";
}

public static void testTestCase() throws Exception {
    MuleClient client = new MuleClient();
    UMOMessage message = client.send("vm://test","hello", null);
    assertNotNull(message);
    assertFalse(message.getPayload() instanceof NullPayload);
    System.out.println("Got Response: " + message.getPayloadAsString());
}

Testing: Mule Server Test(MuleServerTest)=

log4j:WARN No appenders could be found for logger (org.mule.config.MuleDtdResolver).
log4j:WARN Please initialize the log4j system properly.

Got Response: OK - Completed Event: hello

* Level Two: Functional Testing
This testing involves testing the transformation scripts (XSL) to check if the scripts used in the system produce expected results for known inputs. This can be easily achieved by defining a set of typical acceptable messages, applying the desired transformation and comparing the outcome with what was expected.

* Level Three: Integration Testing
This testing involves deploying the whole configuration in a testing environment that is as close to the production version as possible and exercising the ESB with different happy and failing messaging scenarios. Since Mule supports heterogenous inbound endpoints, for example,

  <http:inbound-endpoint address=http://&$(http.host):&$(BBF2011_PORT)/&>
  <http:inbound-endpoint address=http://&$(http.host):&$(IJCAI2011_PORT)/&>

It can be tested through it's responses for the correctness of the integration of endpoints, to see whether they match the input or not.

  – Correctness of Responses: The system can be tested with vari-
ous inputs to check whether the query is delegated to the right Personal Agent and the response(s) are as expected.

- The Rule Responder Benchmark and Testing tool can be used to test the response time of the various use cases of Rule Responder, and provide the user with the ability to gather execution time data for specific instantiation queries of their choice\(^\text{16}\).
- While developing the infrastructure for SymposiumPlanner-2012, we will solicit feedback from the OC for evaluating the system including aspects like knowledge bases and architecture. The evaluation process will also include testing from end user perspective to gauge the completeness, correctness and consistency of the system responses.

## 5 Schedule

<table>
<thead>
<tr>
<th>Time Schedule</th>
<th>2011, April - August</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature Study and Background Research</td>
<td>2011, August</td>
</tr>
<tr>
<td>Proposal Submission</td>
<td>2011, September - Oct</td>
</tr>
<tr>
<td>SymposiumPlanner Design</td>
<td>2011, September</td>
</tr>
<tr>
<td>- Ontology Design</td>
<td>2011, September - Oct</td>
</tr>
<tr>
<td>- System Design</td>
<td>2011, September - Oct</td>
</tr>
<tr>
<td>SymposiumPlanner Implementation</td>
<td>2011, Oct - Jan</td>
</tr>
<tr>
<td>- PAs Implementation</td>
<td>2011, Oct - Nov</td>
</tr>
<tr>
<td>- OA Implementation</td>
<td>2011, Nov - Dec</td>
</tr>
<tr>
<td>- EA and interface Implementation</td>
<td>2011, Dec - Jan</td>
</tr>
<tr>
<td>Testing and Evaluation</td>
<td>2012, Feb - March</td>
</tr>
<tr>
<td>Thesis Defense</td>
<td>2012, June</td>
</tr>
</tbody>
</table>

\(^\text{16}\)https://mandarax.svn.sourceforge.net/svnroot/mandarax/RRBenchmarkingAndTesting/
References


[FKS06] Norbert E. Fuchs, Kaarel Kaljurand, and Gerold Schneider. Attempto controlled english meets the challenges of knowledge representation, reasoning, interoperability and


[PBKCb] Adrian Paschke, Harold Boley, Alexander Kozlenkov, and Benjamin Craig. Rule Responder: A RuleML-Based Prag-


