A multimodal restaurant finder for semantic web

Conference or Workshop Item

How to cite:


For guidance on citations see FAQs.

© 2007 The Authors

Version: Not Set

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data policy on reuse of materials please consult the policies page.
A Multimodal Restaurant Finder for Semantic Web

Yulan He, Thanh Tho Quan and Siu Cheung Hui
School of Computer Engineering
Nanyang Technological University
Singapore
Email:{asylhe,pa0218164b,asschui}@ntu.edu.sg

Abstract—Multimodal dialogue systems provide multiple modalities in the form of speech, mouse clicking, drawing or touch that can enhance human-computer interaction. However, one of the drawbacks of the existing multimodal systems is that they are highly domain-specific and they do not allow information to be shared across different providers. In this paper, we propose a semantic multimodal system, called Semantic Restaurant Finder, for the Semantic Web in which the restaurant information in different city/country/language are constructed as ontologies to allow the information to be sharable. From the Semantic Restaurant Finder, users can make use of the semantic restaurant knowledge distributed from different locations on the Internet to find the desired restaurants.

Index Terms—Multimodal dialogue systems, Semantic Web, Ontology

I. INTRODUCTION

Conversation or dialogue is the most natural way for people to exchange ideas, convey information, and express feelings, etc. Therefore, speech interfaces offer great potential for enhancing human-computer interaction (HCI). In recent years, lots of research work have been done to enhance the speech communication by using complimentary information from gestures. The first known multimodal interface was built in 1980 by Bolt [1] in which shapes could be created, moved, copied, removed, and named using the combination of speech and pointing to resolve deictic references such as that and there. Cheyer [2] describes a multimodal map-based application for a travel planning domain by combining handwriting, gesture and speech. More recently, Johnston et al. [3]–[5] developed a multimodal system called MATCH (Multimodal Access to City Help) which provides a mobile multimodal speech-pen interface to restaurant and subway information for New York City. Pelé et al. [6] developed a virtual human called Nesto which is an interactive 3D embodied conversational agent.

One of the drawbacks of the existing multimodal systems is that they are highly domain-specific and they do not allow information to be shared across different providers. In this paper, we propose a semantic multimodal dialogue system, called Semantic Restaurant Finder (SRF), using Semantic Web services in which the restaurant information in different city/country/language are constructed as ontologies to allow the information to be sharable.

The rest of the paper is organized as follows. Section II describes the current development of Semantic Web and Semantic Web services. Section III discusses the system support for the multimodal Semantic Restaurant Finder. Section IV gives the Semantic Web architecture for SRF. The approach for automatic generation of restaurant ontology is then discussed in section V followed by the performance analysis presented in section VI. Finally, section VII concludes the paper.

II. SEMANTIC WEB AND SERVICES

Vast information are available on the Web. The Semantic Web extends the ability of the World Wide Web by developing standards and tools that allow meaning to be added to the content of web pages. Ontology is a formal, explicit specification of a shared conceptualization. It
is the key enabling technology for the Semantic Web which provides a framework for the sharing and reuse of data on the Web.

Web Services [7] have recently become an important feature in the Web community since they provide a new level of interoperability between applications. Web Services are based on agent technology [8]. A Web Service is identified by a URI (Universal Resource Identifier) [9], whose public interfaces and bindings are defined and described using XML-based service description languages. Typical Web Service techniques such as Web Services Description Services (WSDS) [10] or Universal Description, Discovery and Integration (UDDI) [11] provide technical information on service accessibility (including the service address and the service port), but it lacks the description on service capability. In addition, functions offered by typical Web Services are accessible, but not fully understandable by other programs.

This problem can be tackled by the Semantic Web. The Semantic Web uses ontology as a standard for knowledge presentation. Therefore, knowledge can be shared, reused and exchanged on the Semantic Web. Based on OWL (Web Ontology Language) [12], an ontology language description language, OWL-S (former DAML-S) [13] was introduced to describe Web Services in a semantic manner.

III. MULTIMODAL DIALOGUE SUPPORT

Figure 1 depicts the multimodal dialogue support for Semantic Restaurant Finder. Users can interact with the SFR system in natural ways such as speech, typing, or mouse clicking and drawing to enquire restaurant information. For example, users can request a list of restaurants by saying "show me seafood or general restaurants near Orchard Road". An example of restaurant finding result is displayed in Figure 2.

The multimodal Semantic Restaurant Finding system consists of eight main components. The Open Agent Architecture (OAA2) [14], [15] is used to facilitate multiple asynchronous communications among various components. The components are briefly discussed as follows.

- **Speech Recognizer.** Implemented using the ATK toolkit [16], which is the real-time API for HTK [17].
- **Speech Semantic Parser.** Implemented using the Hidden Vector State (HVS) model [18], [19]. This is illustrated in Figure 3 which shows the sequence of HVS stack states corresponding to the given parse tree. The details of how this is done are given in [19].
- **Gesture Recognizer.** The gesture recognizer used is a sample C++ Win32 application called GestureApp [20].
- **Gesture Semantic Parser.** The gesture semantic parser can be rather treated as an interpreter to map the recognized raw gesture or symbol into a semantic partial frame.
- **Multimodal Input Fuser.** Multimodal input fusing is done by using the time correlates between speech events (words) and the pointing, drawing or mouse clicking events.
- **Dialogue Manager.** The dialogue manager controls the interaction with the user.
- **Response Generator.** The response generator produces an appropriate response according the instructions given by the dialogue manager.
- **Speech Synthesizer.** The Festival 2.0 speech synthesizer [21] is used to transform text into speech.

IV. SEMANTIC WEB ARCHITECTURE FOR SEMANTIC RESTAURANT FINDER

Figure 4 shows the system architecture of SRF using Semantic Web services. There might be several restaurant ontologies available on the Semantic Web. We will discuss the ontology generation process in section V. The Semantic Web architecture consists of three major components: Web Service Provider, Matchmaking Agent and Web Service Requester.

There might be many Web Service Providers on the Web. Each Web Service Provider accesses the restaurant ontology to retrieve restaurant-related information for a given user request. These Web Service Providers must register its services with the Matchmaking Agent which serves as a registry and look-up service repository by
Fig. 1. Multimodal dialogue support for SRF.

<table>
<thead>
<tr>
<th>No.</th>
<th>Description</th>
<th>Distance (km)</th>
<th>Note/Thumbnail</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Mass Restaurants</td>
<td>0.05</td>
<td><img src="image1.png" alt="Mass Restaurants" /></td>
</tr>
<tr>
<td>2.</td>
<td>Fish &amp; Co</td>
<td>0.12</td>
<td><img src="image2.png" alt="Fish &amp; Co" /></td>
</tr>
</tbody>
</table>

Fig. 2. An example of restaurant finding result.

Fig. 3. Example of a parse tree and its vector state equivalent.

Fig. 4. Semantic Web architecture for Semantic Restaurant Finder.
sending a profile file. The information in
the profile is described using the ontology-
based service description language OWL-
S. The Web Service Requester connects
the client Web Browser of the Semantic
Restaurant Finder system with the restaur-
ant ontology. It creates a profile upon re-
ceipt of the user’s request. Such a profile is
then sent as a request to the Matchmaking
Agent to locate the appropriate Web Service
Provider to answer the request.

V. RESTAURANT ONTOLOGY
GENERATION

Restaurant information can be extracted
from the Web and stored in a database,
known as the Restaurant Database. Then,
the information on the Restaurant Database
will be converted into ontology formalism.
This is shown in Figure 5.

A. Knowledge Discovery

The Knowledge Discovery process dis-
covers hidden knowledge from the Restau-ant Database using a clustering technique.
In this research, we have applied one of the
most effective clustering techniques known
as Kohonen Self-Organizing Map (KSOM)
Neural Network [22] for knowledge dis-
covery. Figure 6 shows the Knowledge
Discovery process.

- **Feature Selection** - Information on
restaurants’ street addresses, regional
locations and cuisine types are ex-
tracted as feature factors.
- **Preprocessing** - Information on the
extracted features (i.e. the restaurant
addresses, locations and cuisine types)
is converted into representative binary
strings.
- **Transformation** - The restaurant
records are converted into restaurant
vectors.
- **KSOM Clustering** - This step clus-
ters the restaurant vectors using the
KSOM neural network. The training
algorithm for the KSOM network is
given in [22]. After training, a set of
clusters for restaurants (or restaur-
ant vectors) are generated, and they
are referred to as restaurant clusters.
Information on restaurant clusters is
then stored in the Restaurant Cluster
Database.

B. Ontology Generation

The Ontology Generation process cre-
ates the Restaurant Ontology to represent
the knowledge on restaurants, which can be
obtained from the Restaurant Database and
Restaurant Cluster Database. The Ontology
Generation process is shown in Figure 7.
Figure 8 shows the generated Restaurant
Ontology.

- **Class Mapping** maps the schematic
and discovered knowledge into ontology’s
schematic and discovered classes.
- **Schematic Relation Generation** gener-
ates relations between schematic classes.
- **Composed Relation Generation** gener-
ates the relations between the schematic
classes and discovered classes.
- **Instances Generation** generates instances
for the schematic and discovered classes
based on the data available on the Restau-
rant Cluster Database and Restaurant
Database.

VI. PERFORMANCE ANALYSIS

To evaluate the performance of the pro-
posed approach, an experiment has been
conducted. In the experiment, a Restaurant
Database is constructed. The database con-
tains 2100 restaurants in Singapore, which
are located over 66 locations based on
MRT (Mass Transit Railway) stations and
provides a variety of 76 cuisine types.

As discussed earlier, KSOM neural net-
work is adopted for the proposed sys-
tem. For comparison purpose, a supervised
neural network, Learning Vector Quantiza-
tion Version 3 (or LVQ3) and the k-nearest
neighbor (kNN) technique, are also used
in the experiment for evaluation. There are
two popular variations of the kNN tech-
nique. The first variation, denoted as kNN1,
is based on vector’s normalized Euclidean
distance to find the matches closest to the
input string, while the second variation,
denoted as kNN2, uses the fuzzy-trigram
technique.

Table I gives the retrieval accuracy of
different techniques. The accuracy of the
two neural networks KSOM and LVQ3 are
better than those of the kNN tech-
niques. And LVQ3 has achieved better per-
formance than KSOM. This is expected as
supervised learning techniques often per-
form better than an equivalent unsupervised

Fig. 5. Restaurant ontology generation.

Fig. 6. Knowledge Discovery process.

Fig. 7. Ontology Generation process.

Fig. 8. Restaurant Ontology.

one when applied to the same domain. However, as LVQ3 always require human interpretation in training, it will encounter difficulties when training a large dataset.

Based on the clustering results from KSOM, the Restaurant Ontology is then generated. The generated ontology can be represented using an ontology description language that enables it to be sharable over the Semantic Web by Web-based applications.

VII. CONCLUSION

In this paper, we have proposed the Semantic Restaurant Finder system, in which restaurant information in different city/country/language are constructed as ontologies for knowledge sharing. As such, users can make use of the semantic restaurant knowledge distributed from differ-
ent locations on the Internet to find the desired restaurants. To achieve this, this paper has presented the multimodal dialogue support and the semantic web architecture for Semantic Restaurant Finder. In the proposed architecture, Web Service Requester, Matchmaking Agent and Web Service Provider are constructed to allow users to access the multiple Restaurant Ontologies available on the Semantic Web environment. A framework for ontology generation using the KSOM neural network has also been discussed. Finally, an experiment has been conducted to evaluate the performance of the proposed system.

REFERENCES