

# Domain Knowledge Model for Embodied Conversation Agent

<sup>1</sup>Goh Ong Sing, <sup>2</sup>Chun Che Fung, <sup>3</sup>Arnold Depickere, <sup>4</sup>Kok Wai Wong, <sup>5</sup>Wilson Wong

School of Information Technology, Murdoch University, Perth, Western Australia

{<sup>1</sup>os.goh, <sup>2</sup>l.fung, <sup>3</sup>a.depickere, <sup>4</sup>k.wong} @murdoch.edu.au

Faculty of Information and Communication Technology, Kolej Universiti Teknikal Kebangsaan

Malaysia, Melaka, Malaysia

<sup>5</sup>wilson@kutkm.edu.my

## Abstract

*This paper presents and discusses several technical issues in implementing the conversation intelligent agent. In this work, we focus on domain knowledge models that enable structured queries on the field of epidemic crisis. We will simulate this system with sample conversation taken from our target Domain-Specific knowledge on the recent world epidemic crisis, Severe Acute Respiratory Syndrome (SARS). A solution would be to incorporate the conversation system embodied intelligent agent which will allow a more user-friendly interaction. The system does not only match a query against a database of keywords from our domain knowledge model but also help the user navigate through the document space until the appropriate information are found using the “URLs push” technique.*

## 1 Introduction

The turn of the millennium has brought with it the wind of change to the community of conversation system based on World Wide Web online information. According to Google Inc [1], the size of the static web has reached 8.058 billion pages in the year 2005 thus, making the World Wide Web one of the most populous sources of information ever encountered. Researchers in the field are slowly seeing a shift in approach, a shift towards the adoption of knowledge-base, higher level of natural language processing and advanced reasoning for development of the conversation system. Several related research on embodied conversation agent are investigating the construction of domain knowledge models. There has been some pioneering work on conversational interfaces for simple tasks in limited domains such as [2] [3] [4]. The REA [5] uses an approach on discourse modeling and the conversations are story-related. In the DIVA II project [6], the conversation is modeled implicitly within the video and audio data annotations. P. Tarau and E. Figa [7] used the

Prolog-based conversational agent which integrates more than a GigaByte of knowledge base data from Open-Domain knowledge base such as WordNet, FrameNet and Open Mind. InCA [8] is a personal coach agent equipped with pedagogical models as well as specific domain knowledge.

## 2 Conversation Engine

This research project involves the establishment of a domain knowledge model with an architecture aiming for the possibility of practical applications in nearer future conversation robot or chatterbot, called Artificial Intelligent Neural-network Identity (AINI) [9]. Our real-time prototype relies on distributed agent architecture specifically at the Web. A software agent, such as the conversation engine, multi-domain knowledge model, multi-modal human-computer communication interface, etc, communicates with one another via TCP/IP. AINI is a conversation agent that is capable of involvement in a fairly meaningful conversation with users who interact with her. AINI is a software conversation robot, which uses human-computer communication system, a combination of natural language processing and multi-modal communication. A human user can communicate with the developed system using typed natural language conversation. The system embodied conversation agent system will reply text-prompts or Text-to-Speech Synthesis and with appropriate facial-expressions.

For the purposes of this research, the application area of designing the conversation agent is primarily grounded in an ability to communicate based upon scripting and/or artificial intelligence programming on the field of epidemic crisis. A sample of the communication interface between a user and AINI in the CCNet portal [10] is depicted in Figure 1.



Figure 1: Natural language interface in CCNet Portal

AINI employs an Internet three-tier, thin-client architecture (Figure 2) that may be configured to work with any web application. Composed of a data server, application and client layers, this Internet specific architecture offers a flexible solution to the unique implementation requirements of the AINI system. The data server layer serves as storage for permanent data required by the system, where the knowledge bases (epidemic Domain-Specific extracted by the Automated Knowledge Extraction Agent (AKEA) [11] and Open-Domain from existing Artificial Intelligence Markup Language (AIML) Loebner prize knowledge base [12]) and the conversation logs reside. These web-enabled databases are accessible via the SQL query standard for database connectivity using MySQL database.

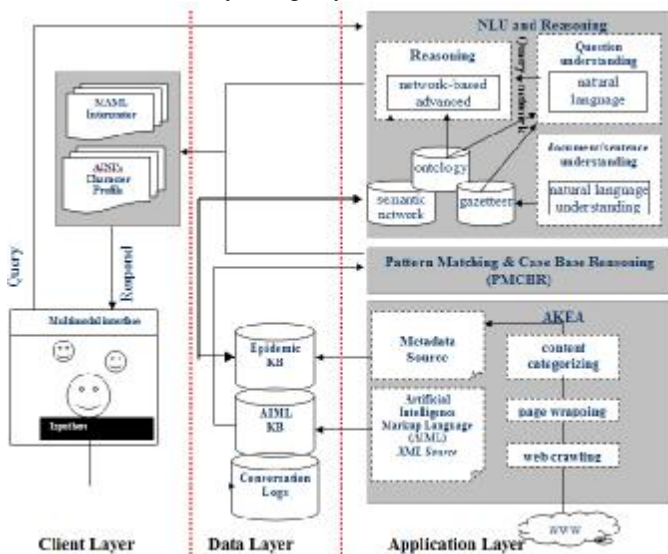


Figure 2: AINI's Architecture

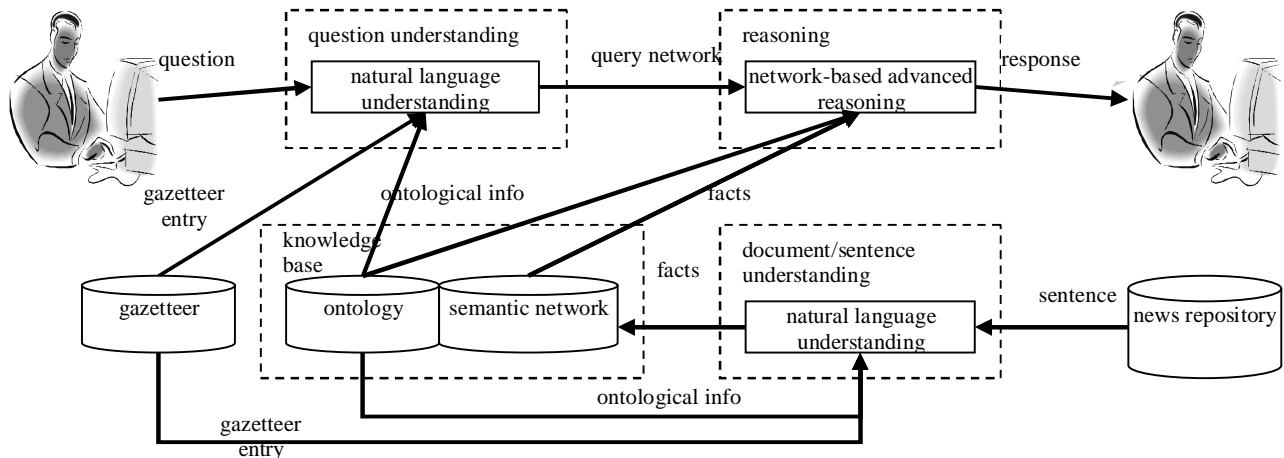
The application server layer handles the processing of logic and information requests. Here, one or more application servers are configured to compute the dialogue logic through the hybrid approach by implementing Natural Language Understanding and Reasoning (NLUR) or Pattern Matching and Case Base Reasoning (PMCBR) algorithm. The user interface resides in the thin-client layer and is completely browser based, employing Multi-modal Agent Markup Language (MAML) interpreter and Microsoft SAPI to handle the users interface. MAML is a prototype multi-modal markup language based on XML that enables animated presentation agents or avatars. It involves a talking virtual lifelike 3D agent character that is capable of involvement in a fairly meaningful conversation. The conversation engine is based on the Web-based and towards an architectural open-source practice by employing PHP, Perl scripting language, Apache Server and knowledge base stored in a MySQL server.

## 2.1 Pattern Matching and Case-Based Reasoning (PMCBR)

Our approach to handling conversations with the embodied intelligent agent is to have case-based rules that run on top of reasoning rules. The case-based rules, which do simple pattern matching, have the advantage of being quick and thus, being able to return near instantaneous responses to the users. This is the ultimate goal in maintaining believability in the interaction. The Conversation Engine handles this reactive component. The PMCBR Conversation Engine is based on the ALICE [13] engine. The ALICE chat engine implements the AIML, which allows dialogs between the user and agent to be easily acknowledged. Judging from the specification based on XML, we believe AIML to be the perfect markup language to our system.

## 2.2 Natural Language Understanding and Reasoning (NLUR)

The functionalities of the framework are nicely packed into two main subsystems namely natural language understanding and network-based advanced reasoning as shown in the Figure 3. There are three groups of storage structures namely news repository, knowledge base which consists of ontology and semantic network, and gazetteer. Collectively, the system interacts with the environment in three forms namely reading and understanding sentences from news articles that have been processed to populate the semantic network, receiving question from users, and providing response to users.



**Figure 3. Architecture of NLUR Knowledge base Query**

The domain knowledge and framework of AINI is constructed based on full-discourse natural language understanding and reasoning with cooperatively behavior for representing and reasoning with the meaning of natural language information from the World Wide Web.

Firstly, the framework adopts a natural language understanding approach that not only covers the necessary aspects of semantic analysis, but also include the crucial aspects in discourse analysis. As a result, such approach will ensure that the conversational system can handle both questions and information from natural language text from any information source and domain. Secondly, with the solution based on powerful and expressive representation formalism like the semantic network, facts produced by the full-discourse natural language understanding, including intrinsic properties between entities can be captured and expressed with the help of ontological information. Thirdly, with the network-based representation formalism, alternative approaches for reasoning, other than rule-based, that can fully exploit the formalism's expressiveness can be adopted in the solution. Also, with the ontological commitment supported by the network-based representation formalism, the integration of advanced reasoning features can be done.

### 2.2.1 Natural Language Understanding

This subcomponent is responsible for reading and understanding two things: questions from users and sentences of processed news articles from news repository. In more specific terms, this subsystem is to turn the English representation of the meaning of questions or news content into the network [14]. The process is carried out in four phases by four natural language processing modules namely sentence parsing, named-entity recognition, relation inference and discourse integration.

The input to this subsystem must be in the form of a single English sentence where each can either be the individual sentences of the news content from news repository or the questions from users. The output of this subsystem takes the form of a network-bound meaning representation of all the sentences in a discourse.

The sentence parsing module takes a single English sentence and using Minipar, produces the grammatical categories and relations of the sentence. These two types of information are then fed into named-entity recognition and relation inference. The named-entity recognition module implements both the noun phrase chunking and category assignment algorithm discussed in the previous section. This module first performs noun phrase chunking and later assigns a category out of the many predefined based on advice from the gazetteer and the ontology.

The tagged noun-phrases or named-entities are then passed on to relation inference and together with the output of sentence parsing, relation inference will deduce and extract useful grammatical relations. The output of named-entity recognition and relation inference for different sentences of the same discourse are gathered and used by the discourse integration module to generate a network representation of the meaning of the discourse. This discourse integration module is the one that implements the four-stage approach with light-weight anaphora resolution.

### 2.2.2 Reasoning with Cooperative Behavior

Network-based advanced reasoning subsystem is responsible of discovering the valid answer and generate unambiguous response or generate an explanation for users' questions [15]. The process is executed in five phases by five modules namely network-to-path

reduction, selective path matching, relaxation of event constraint, explanation on failure and template-based response generation.

The network-to-path reduction module collapses the query network into sets of path sequences to reduce the complexity in discovering the answer. The output of network-to-path reduction is two sets of path sequences that will be used by the selective path matching module to discover the answer from the semantic network through a series of conditional path unification. To extend beyond literal matching of path sequence, ontological information is utilized to put into consideration events that are hierarchically equivalent. This process is performed by the module relaxation of event constraint. In case of failure during the discovery for a valid answer by selective path matching, an explanation or justification is dynamically generated by the explanation on failure module as an alternative response. This process is carried out based on the context of the question and the current status of the semantic network. If answers can be validly discovered, then readable and unambiguous natural language responses are generated by the template-based response generation module.

### **3 Domain Knowledge Model in AINI's Conversational System**

Another significant difference between this research and other conversational agents is the domain knowledge model. Dahlbäck and Jönsson [16] stressed that, the Domain Model represents the structure of the knowledge which comprises a subset of general world knowledge. In our research, the domain model is the taxonomy of knowledge related to the topic of the presentation, or XML-like metadata model. This will reduce the workload of the author to predict every input typed by the user, and instead, to allow the author to put more effort on scripting conversation within a specified domain or conversation Domain-Specific.

We believe that the ultimate conversational human-computer interface uses and requires different kinds of approaches. Therefore, we have been working to develop a domain knowledge model for building conversation and interactive systems. For example, according to S. Kshirsagar and N. Magnenat-Thalmann [17], having a small conversation about the weather requires a lot less resources than a philosophical discussion about the meaning of life. In our research, we defined our conversation system as a collective of specific conversation units; every unit handles a specific conversation between user and computer. In our case,

Domain-Specific is from the epidemic online document from selected websites.

Domain is one of the dimensions that determines the focus or direction of a conversational system. An Open-Domain will practice techniques based on probabilistic measures and has a wider range of information source. For a system that focuses on certain domains, it is more likely that the techniques are more logic-based and well-founded, with relatively limited sources as compared to an Open-Domain. A domain-oriented conversational system deals with questions under a Domain-Specific and can be seen as a richer approach because natural language processing systems can exploit domain knowledge and ontologies. Advanced reasoning such as providing explanation for answers, generalizing questions, etc is not possible in Open-Domain systems. Open-Domain question answering systems need to deal with questions about nearly everything and it is very difficult to rely on ontological information due to the absence of wide and yet detailed world knowledge. On the other hand, these systems have much more data to exploit in the process of extracting the answers.

In our architecture, we have implemented a multi-domain model: an Open-Domain knowledge base which is converted from the AIML knowledge base and a Domain-Specific knowledge base which is the epidemic online document extracted by (Automated Knowledge Extraction Agent) AKEA. However, if the user speaks out of the presentation topic, we define this domain category as the Unknown-Domain. This is to determine whether the user is chatting within the domain of the presentation topic or the user is conversing differently from the domain knowledge model presented. By doing this, we have rectified the trait of the conversation agent or software robot, from a diverse conversation to a specific presentation topic. The web knowledge base is continuously updated with facts extracted from online *epidemic news* using information extraction (IE) and knowledge representation by AKEA. IE is the task of extracting relevant fragments of text from larger documents, to allow the fragments to be processed further in some automated way, for example to answer a user's query. The ontology and gazetteer will be implemented as domain-dependent modular components, allowing future improvements to achieve openness in the domain.

The development of AKEA has been reported in [11] which was designed to establish the knowledge base for a global crisis communication system called CCNet. CCNet was proposed during the height of the SARS epidemic in 2003. It was aimed at providing up-to-date information to its users via a conversational software robot called AINI. The purpose of AINI is to deliver

essential information from trusted sources and it is able to interact with its users by animated characters. The idea is to rely on a human-like communication approach thereby providing a sense of comfort and familiarity. The functionalities of AINI have been reported in the past and development on AINI is ongoing [18]. It is foreseeable that the combination of AINI and AKEA will produce a more natural means of communication and computing in the near future.

#### 4 The Role of Domain Knowledge in Conversational System

AINI's domain knowledge model usually incorporates several knowledge domains, thus merging the expertise of one or more experts. A "sales" domain knowledge for instance, would contain expertise on improving sales, but it would also incorporate with an Open-Domain knowledge. Multiple domain knowledge, merged into AINI's single domain knowledge would give the users the best conversation.

We pre-defined the Open-Domain and Domain-Specific in the data layer, while the Unknown-Domain is taken from a conversation log. We are also free to modify the Unknown-Domain set from the administrator page so that the other two domains can improve their knowledge bases automatically. This improvement transition is known as semi-automatic machine learning. Based on the type of input provided by the user, the agent's response state moves smoothly from one domain knowledge base to another. According to K. Mori, et. al. [19], these two intermediate states transition called "Reluctant" and "Concede".

Even though the conversation agent lets the user to direct the conversation in the Unknown-Domain, the agent will continue to remind and recall the user to the current topic of the presentation. This is to convey and direct the users' attention to move back to its original Open-Domain or Domain-Specific state. However, the priority will be on the Domain-Specific. Therefore, the conversation agent will always give higher priority to Domain-Specific in an attempt to keep the user focused on the topic of the presentation. An example of domain knowledge model conversation between Sean and AINI is shown in Figure 4 on the SARS epidemic crisis.

From the conversation log, the transition state in dialog pair A1-A4 used Open-Domain and dialog pair B1-B4 used Domain-Specific. However, in dialog pair B5-B6, the user had control of the conversation although the agent reminded the user of the topic of the current presentation. The transition state in dialog pair B6 is under the Unknown-Domain state and the dialog set is

```
User_A1 : Good Morning
AINI_A1 : Morning! How are you today?
User_A2 : What is your name?
AINI_A2 : My name is Aini. What is your name?
User_A3 : Sean
AINI_A3 : Nice to meet you Sean! Are you a man or a woman?
Sean_A4 : Man
AINI_A4 : Thanks for telling me your gender, Sean. Most of my clients
are men.
Sean_B1 : You still remember when was SARS first carried out into the
world?
AINI_B1 : Yes. SARS was first carried out into the world at large on 21
February, at Metropole Hotel, Hong Kong.
Sean_B2 : How to diagnosis SARS?
AINI_B2 : Presently there is no test available yet for SARS.
Sean_B3 : Wow, tell me more.
AINI_B3 : Diagnosis is based on clinical illness, history of recent travel
to high risk areas or history of close contact with persons with
SARS. More information at
http://www.pcdom.org.my/pubinfo/pdf/SARS%20Health%20Alert\_EN\_BM.pdf
Sean_B4 : When did Singapore mark on the SARS-free anniversary?
AINI_B4 : Monday, 24 May 2004.
Sean_B5 : Ok, why not we talk about Lymphogranuloma
venereum(LGV)
AINI_B5 : I'm not sure how this related to SARS? Why don't you ask
me about what is the symptoms of SARS?
Sean_B6 : No, why can't we talk about LGV?
AINI_B6 : OK. I would do a search for it.
Sean_B7 : where does meningitis occur?
AINI_B7 : Here you go...(AINI will automatically "push" following
URLs to the user's browser.)
http://www.who.int/mediacentre/releases/2004/pr25/en/
A rare strain of meningitis, which re-emerged recently in
Burkina Faso, would have left health authorities helpless just
two years ago.
Sean_B8 : Thanks for your information. Can you show me how
meningitis spread?
AINI_B8 : OK. I will ask my friend to explain how meningitis spread.
AINI will automatically "pushe learning object page
embedded with Flash and Multiagent character to explain
how meningitis spread to the user's browser.)
```

Figure 4: Conversation log on SARS Epidemic Crisis

taken from a pre-defined dialog set from the Open-Domain. The dialog contains statements such as "I would do a search for it.", "Did I misunderstand your meaning?", "That's an interesting question.", "I'll come back to that in a minute" etc and these statements will be monitored and submitted into the Unknown-Domain data layer. The author is free to modify and enhance the Unknown-Domain set to become a Open-Domain or Domain-Specific. It depends on the context or topic domain presentation. However, in dialog pair B7 and B8, AINI reused the Domain-Specific and automatically "pushed" relevant URLs to the user's browser for the following conversation.

Bear in mind the price to pay for this empirical approach is the difficulty to scale-up such systems or to port them to new application domains, in the absence of a principled approach. As a result, these are often

supported by extensive quantitative development, large databases.

## 5 Conclusion

Based on this experiment, the domain knowledge model shows interesting behavior in the natural conversation agent. The key assumption is that important queries do not necessarily turn up the answers that they can be found in a number of different domains. In this paper, we only worked on selected epidemic crisis websites which perform knowledge extraction for Domain-Specific knowledge on the server. Although we simulate the proxy conversation log that contains of client requests, there is a possibility that new simulation result from other traces is different from the result referred in this paper.

It is believed that this research provides a well-engineered platform for experimentation with various Web-enabled question answering techniques by employing a conversation software agent. The implementation is currently under ongoing development. In the future, we will endeavor to continually refine the existing technology and to develop new knowledge base models that can be applied to new domains.

## References

- [1] The Google Timeline, <http://www.google.com/corporate/timeline.html>, Last access on 15 March 2005
- [2] J. Dowding, J. Gawron, D. Appelt, J. Bear, L. Cherny, R. Moore, and D. Moran, "Gemini: A natural language system for spoken-language understanding," presented at 31st Annual Meeting of the Association for Computational Linguistics, Columbus, Ohio, USA, 1993.
- [3] J. F. Allen, "The TRAINS project: a case study in building a conversational planning agent," *Journal of Experimental and Theoretical AI (JETAI)*, vol. 7, pp. 48, 1995.
- [4] S. Seneff, E. Hurley, R. Lau, C. Pao, P. Schmid, and V. Zue, "Galaxy-II: A reference architecture for conversational system development," presented at 5th International Conference on Spoken Language Processing, ICSLP-98, Sydney, Australia, 1998.
- [5] J. Cassell, T. Bickmore, M. Billinghurst, L. Campbell, K. Chang, H. Vilhjalmsson, and H. Yan, "Embodiment in Conversational Interfaces, REA," presented at SIGCHI conference on Human factors in computing systems: the CHI is the limit, Pittsburgh, Pennsylvania, 1999.
- [6] N. Braun, "Interaction Approach for Digital Video Based Storytelling," presented at International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, Czech Republic, 2000.
- [7] P. Tarau and E. Figa, "Knowledge-based conversational agents and virtual storytelling," presented at The 2004 ACM symposium on Applied computing, Nicosia, Cyprus, 2004.
- [8] L. Razmerita, T. Nabeth, A. Angehrn, and C. Roda, "InCA: An Intelligent Cognitive Agent-based Framework for Adaptive and Interactive Learning," presented at Conference on Cognition and Exploratory Learning in Digital Age, IADIS CELDA 2004, Lisbon, Portugal, 2004.
- [9] MBR, *1st Malaysian Robotic Interactive Programme, The Malaysia Book of Records (MBR), A Tribute to Tun Dr. Mahathir Mohamad*, Gold ed: R&D Communication Sdn Bhd, 2004.
- [10] G. O. Sing, C. C. Fung, and L. M. Ph'ng, "Intelligent Agents for an Internet-based Global Crisis Communication System," *Journal of Technology Management And Entrepreneurship*, vol. 2, 2005.
- [11] G. O. Sing and C. C. Fung, "Automated Knowledge Extraction from Internet for a Crisis Communication Portal," presented at First International Conference on Natural Computation, Changsha, China, 2005.
- [12] Loebner Prize Gold Medal, <http://www.loebner.net/Prize/loebner-prize.html>, Last access on 6 July 2005
- [13] Artificial Linguistic Internet Computer Entity, <http://www.alicebot.org>, Last access on 10 April 2005
- [14] W. Wong, G. O. Sing, and M. M. Yusof, "Syntax Preprocessing in Cyberlaw Web Knowledge Base Construction," presented at International Conference on Computational Intelligence for Modeling, Control and Automation (CIMCA 2004), Gold Coast, Australia, 2004.
- [15] W. Wong and G. O. Sing, "Evaluation of Response Quality for Heterogeneous Question Answering," presented at IEEE/WIC/ACM International Conference on Web Intelligence (WI 2005), Compiègne, France, 2005.
- [16] N. Dahlbäck and A. Jönsson, "Integrating domain specific focusing in dialogue models," presented at Eurospeech'97, Rhodes, Greece, 1997.
- [17] S. Kshirsagar and N. Magnenat-Thalmann, "A multilayer personality model," presented at 2nd International Symposium on Smart Graphics, 2002.
- [18] G. O. Sing and C. C. Fung, "Intelligent Agent Technology in E-Commerce," in *Intelligent Data Engineering and Automated Learning, LNCS.*, vol. 2690, J. Liu, Y. Cheung, and H. Yin, Eds.: Springer-Verlag, 2003, pp. 10-17.
- [19] K. Mori, A. Jatowt, and M. Ishizuka, "Enhancing Conversational Flexibility in Multimodal Interactions with Embodied Lifelike Agents," presented at International Conference on Intelligent User Interfaces, Miami, Florida, USA, 2003.