

Query Based Intelligent Web Interaction with Real World Knowledge

Ong Sing GOH, Chun Che FUNG, Kok Wai WONG
School of Information Technology
Murdoch University, South Street, Western Australia 6150
{os.goh, l.fung, k.wong}@murdoch.edu.au

Arnold DEPICKERE
Division of Arts
Murdoch University, South Street, Western Australia 6150
a.depickere@murdoch.edu.au

Submission date: 23rd February 2007

Abstract In times of crisis, an effective communication mechanism is paramount in providing accurate and timely information to the community. In this paper we study the use of an intelligent embodied conversational agent (ECA) in the Crisis Communication Network Portal (CCNet). The proposed system, CCNet, is an integration of the intelligent conversation agent called AINI (Artificial Intelligent Natural-language Identity) and an Automated Knowledge Extraction Agent (AKEA). AKEA retrieves first hand information from trustworthy websites. In this paper, we compare the interaction of AINI software robot against two popular search engines, two question answering systems and two other conversational systems. We have found that domain-specific knowledge base has higher response satisfaction levels from the user than conversational-style responses. It can be anticipated that ECA will play an important role in increasing the acceptance and use of conversational agents.

Keywords: Automated Knowledge Extraction Agent (AKEA), Crisis Communication Network Portal (CCNet), Artificial Intelligent Natural-language Identity (AINI), intelligent embodied conversational agent (ECA), Natural Language Processing (NLP)

1 Introduction

The outbreak of diseases, virus, natural disasters and terrorists attacks have caused much miseries, fear and confusion around the world. Examples of such crisis are the Severe Acute Respiratory Syndrome (SARS), bird flu, September 11, earthquakes and the recent tsunamis. In times of crisis, a lot of people will be looking for information. This ranges from management and decision-makers, frontline special or emergency services personnel who have to deal with the situation, citizens who are directly or indirectly involved, and the general public who would like to be informed of the situation and the development.

While the media plays an increasingly important role in reporting and providing the latest information on these events, it is importance that an effective and reliable means of communication is maintained during extraordinary events. To this end, the Internet will have a crucial role as an essential communication channel. Due to the rapid advancement of network technology, availability of broadband services, mobile computing and convergence of voice communication and public broadcast into an internet based system; it is therefore foreseeable that the Internet will become the dominant communication system in the world. It is anticipated that the Internet will overtake the traditional means of telecommunication such as voice networks and the public free-to-air broadcast. Hence, a global crisis communication system based on the internet has been proposed in this paper. One of the key components of the proposal is a computer-driven, natural language “chatterbots”, capable of providing meaningful and quick responses to queries from the users in a conversational format. In this paper, the performance of the developed prototype system is compared to two popular search engines – Google and Yahoo!, two question

answering systems – START and AskJeeves and two conversational systems – ELIZA and ALICE. With more natural and meaningful responses, it is expected that the proposed ECA will provide assistance to the users during real-world crisis situations.

2 AINI’s Conversational Agent Architecture

This research project involves the establishment of a Crisis Communication Network (CCNet) portal¹. The objective is to use an embodied conversational agent (ECA) based on an architecture called Artificial Intelligent Natural-language Identity (AINI). As reported previously, the AINI architecture forms the backbone of the conversational system. It can be scaled up to any new applications. AINI engines are portable and have the ability to communicate naturally and carry on multiple independent conversations at the same time. AINI’s knowledge bases and conversational engines use plug-in principles that can quickly be augmented with specific knowledge and are portable in nature for specific purposes.

Our real-time prototype relies on distributed agent architecture designed specifically for the web and mobile technology. The software agent is based on a conversation engine using a multi-domain knowledge model and with multimodal human-computer communication interface. It also offers multilevel natural language query which communicates with one another via TCP/IP communication protocol. As a conversation agent (or chatterbot), AINI is capable to engage the user with a meaningful conversation on specific topic. In this particular application, the topic is on the possible pandemic virus, H5N1. From another perspective, AINI can be considered as a software conversation robot. It uses a form of human-computer communication approach which is a combination of natural language processing and multimodal communication. A human user can communicate with the developed system using typed natural language conversation. The embodied conversation agent system will reply in text-prompts or Text-to-Speech

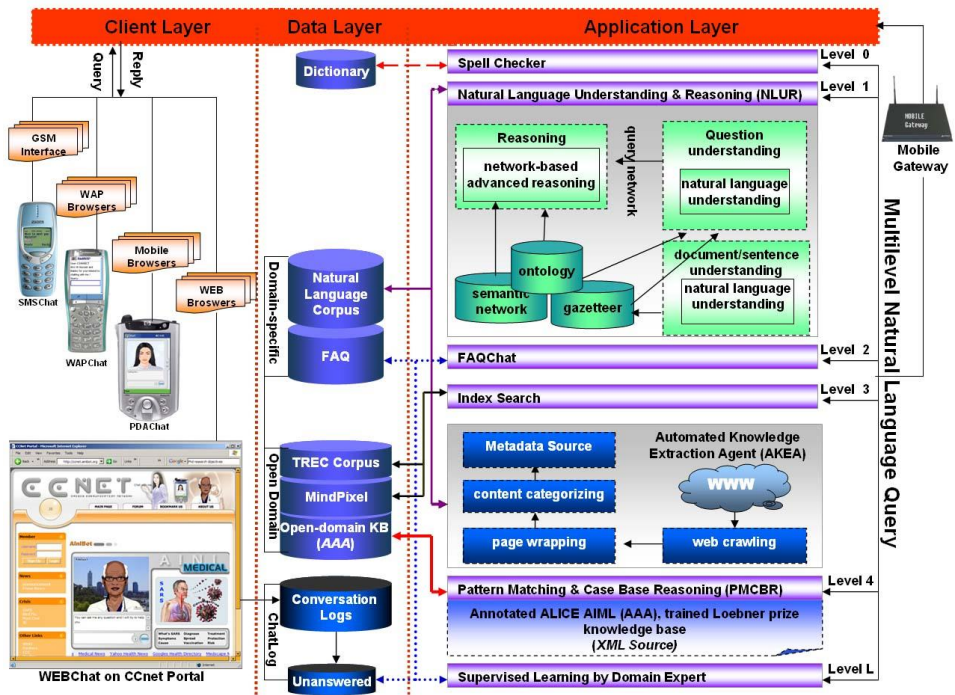


Fig. 1 AINI’s Conversational Agent Architecture

¹ The experimental portal is located at <http://ainibot.murdoch.edu.au/ccnet>

Synthesis together with appropriate facial-expressions.

As shown in Fig. 1, AINI adopts a hybrid architecture that combines the utility of multidomain knowledge bases, multimodal interface and multilevel natural language query. Given a question, AINI first performs a question analysis by extracting pertinent information to be used in query formulation, such as the Noun Phrases (NPs) and Verb Phrases (VPs) by deploying MINIPAR parser [1] in our newly built full parsing Natural Language Understanding and Reasoning (NLUR). MINIPAR is a broad-coverage parser for the English language. An evaluation with the SUSANNE corpus shows that MINIPAR achieves about 88% precision and 80% recall with respect to dependency relationships. In our experiment by using Pandemic corpus extracted by Automated Knowledge Extraction Agent (AKEA)[2], NLUR parsers is capable to parses nearly 500 words per second on a Centrino 2GH with 1GH memory.

AINI employs an Internet three-tier, thin-client architecture that may be configured to work with any web application[3]. It is comprised of a data server layer, application layer and client layer. This Internet-specific architecture offers a flexible solution to the unique implementation requirements of the AINI system.

3. Domain Metric Knowledge Model

In this paper, Bird Flu pandemic is the focus of the domain-specific research. Research and information on H5N1 pandemic have become increasingly important as the pandemic will have dire global implications. Wall Street Journal Online [4] predicted that this pandemic could be worse than the one in 1918 which killed at least 20 million people and the World Health Organization estimated that the maximum figure could be as high as 50 million [5].

The AINI's domain metric knowledge model incorporates several knowledge domains. This is analogous to the expertise knowledge from multiple experts. For example, a *sales* domain knowledge should contain expertise or knowledge on how to improve sales. However, the system should also incorporate Open-Domain knowledge to handle general

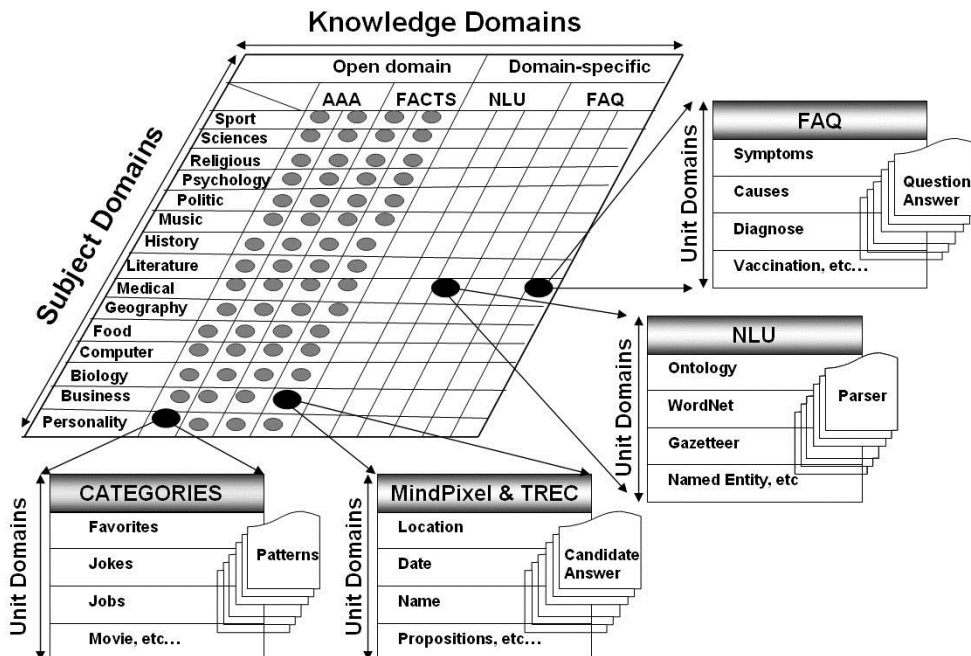


Fig. 2 AINI's Domain Metric Knowledge Model

or generic questions. By including multiple domain knowledge bases with AINI's single



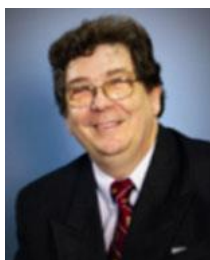
Ong Sing GOH: Currently he is a researcher at School of Information Technology, Murdoch University. He was a founding chairman for Artificial Intelligent and Modeling Group and associate professor of Faculty of Information Technology and Communication, Technical University of Malaysia Melaka. He received his B.A.Ed (Hons) in 1990 and M.Sc. (1996) in Computational Linguistics, University of Manchester Institute of Science and Technology, UK. His main research interest is in the development of intelligent agent, natural language processing and speech technology to facilitate graceful human-computer interactions, conversational robot and mobile services. He is a member of IEEE Computer Society, Member of Malaysia Invention and Design Society (MINDS), Senior Member, Malaysia Senior Scientists' Association (MSSA), and Member of International Association of Engineers.



Chun Che Fung: He received the B.Sc.(1st Class Hon.) and M.Eng degrees from the University of Wales in 1981 and 1982 respectively. He was awarded his Ph.D degree from the University of Western Australia in 1994. He has been an Associate Professor of the School of Information Technology, Murdoch University since 2003. Prior to his present appointment, he worked at the School of Electrical and Computer Engineering, Curtin University from 1988 to 2002, and the Department of Electronic and Communication Engineering, Singapore Polytechnic from 1982 to 1988. His research interests are computational intelligence techniques and intelligent systems applications. He is a Senior Member of IEEE and a Member of ACS, IEAust and ACM.



Kok Wai Wong: He received his Bachelor (Hons.) of Engineering (Computer System Engineering) and Ph.D. from Curtin University of Technology in 1994 and 2000 respectively. He is currently working as an Associate Professor with the School of Information Technology at Murdoch University in Western Australia. He is a Senior Member of IEEE, a member of ACM and ACS. He is also involved in the editorial board for the Journal of Advanced Computational Intelligence and Intelligent Informatics, ACM Computer in Entertainment, International Journal of Intelligent Information Technologies, and International Journal of Information Technology and Web Engineering. His current research interests include intelligent data analysis, game and digital media technology, and data mining.



Professor Arnold Depickere: He has received an MBA (Info Sys) from Maastricht School of Management, Maastricht, Netherlands. Prior to joining Murdoch University, Arnold had a combination of 30 years of experience in both the IT industry and IT education; 9 years of which was in Europe followed by 21 years in the Far East. He is currently holds the position of Executive Dean for the Division of Arts; encompassing 65% of the University.