

# AINI - Embodied Conversation Agent Applicable for Interactive Games

<sup>1</sup>Ong Sing Goh, <sup>2</sup>Chun Che Fung

School of Information Technology, Murdoch University, Murdoch, Western Australia  
 {<sup>1</sup>os.goh, <sup>2</sup>l.fung } @murdoch.edu.au

*Abstract:* - Recent advances in Natural Language Processing (NLP) and Artificial Intelligence (AI) in general have advanced this field in realizing the vision of a more humanoid interactive system with embodied conversation agent (ECA). This paper presented our development of the ECA called Artificial Intelligent Natural-language Identity or AINI. In this paper, we also present and discuss the use of such ECA for inclusion in interactive games. We present the technical design of our ECA and its performance.

*Key-Words:* - Embodied Conversation Agent (ECA), Artificial Intelligence (AI), Games, Artificial Intelligent Natural-Language Identity (AINI)

## 1. Introduction

During the past decade, one can observe that there are rapid advances in embodied conversational agent (ECA), spoken language technology, natural language processing and multimodal interfaces to replace chatterbots which based on pattern matching and typed-text-based. All these have stimulated the interest towards a more human-like conversational interfaces [1] [2] and [3]. Many researchers have also been observed in AI researches into natural language conversation [4], [5], [6] and the application of ECA in Games [7, 8]. They have proposed different techniques and produced several natural language conversation systems. Every year they present their work by competing for the Turing Test (TT)[9].

There are hundreds of different ECAs developed for a variety of reasons. They range from hardwired programs with simple coded patterns to systems built upon embedded learning algorithms which continuously expand their language knowledge base. Our ECA called Artificial Intelligence Natural-language Identity was created purely for fun or as part of interactive games [1]. Many ECAs are designed to provide specific information and direct the dialogue to specific topics.

On the other hand, conversational characters represent the convergence of animated interface agents and human-computer dialogue systems. As animated agents get more realistic, the user naturally expects to be able to interact with them in natural language. And as human-computer dialogue systems develop, it appears that users could interact more readily in natural language if the system is personified through an agent.

## 2. Tricks in Imitation Games

**Table 1: Tricks in Conversation Agent**

|              |   |
|--------------|---|
| <b>ELIZA</b> | <ul style="list-style-type: none"> <li>- Fostered by including substrings of the user's input in the program's output.<br/>User: You hate me.<br/>ELIZA: Does it please you to believe that I hate you?</li> <li>- Use of the Rogerian mode, which provides unimpeachable cover for the computer.</li> </ul>                |
| <b>PARRY</b> | <ul style="list-style-type: none"> <li>- Admitting ignorance.<br/>`I don't know."</li> <li>- changing the level of the conversation<br/>``Why do you ask that?,"</li> <li>- Introducing new topic, launching into a new story also called as simulates paranoid behavior.</li> <li>- launching into a new story.</li> </ul> |

|                     |   |
|---------------------|---|
| <b>Other Tricks</b> | <ul style="list-style-type: none"> <li>- Having many fragments of directed conversation stored in activation network.</li> <li>- Changing the level of the conversation</li> <li>- Humorous statements to make the program seem more human.</li> <li>- Agreeing with the user, in certain cases, the program can safely agree with the user rather than being non-committal.</li> <li>- Simulated typing, by including realistic delays between characters and imitate the rhythm of a person typing.</li> <li>- Make longer replies and its seemed more human-like that cur reply.</li> <li>- Introduce a new subject with a certain probability.</li> <li>- Fragments of directed conversation stored in activation network, "opening book."</li> <li>- Controversial statements, "People don't own cats..."</li> <li>- Agreeing with the user, rather than being non-committal.</li> <li>- Excerpting News, weather forecast, textual games, etc.</li> </ul> |
|---------------------|---|

Some people interpret the TT as a setting in which you can "cheat". The imitation game (IG) has no rules constraining the design of the machines. Turing describes how machines could be "rigged" to overcome certain obstacles proposed by opponents of the idea that

machines can think. An obvious example is about machines making mistakes. When the machine is faced with an arithmetical challenge, in order not to give away its identity of being fast and accurate, it can pause for about 30 seconds before responding and occasionally give a wrong answer. Being able to carry out arithmetical calculations fast and accurately is generally considered intelligent behavior. However, Turing wishes to sacrifice this at the expense of human-ness. Some commentators think this is "cheating". The machine is resorting to certain "tricks" in its operations rather than imitating the human ways. However, arithmetic is a highly specific domain. Modifying the programs in this manner cannot hurt: If a machine can pass the test, it can then be re-programmed not to cheat at arithmetic. If it does not resort to this, the interrogator can ask a difficult arithmetical problem as his/her first question and decide that he/she is dealing with a machine right then and there. We believe the best way to handle this issue is considering this as "deception" rather

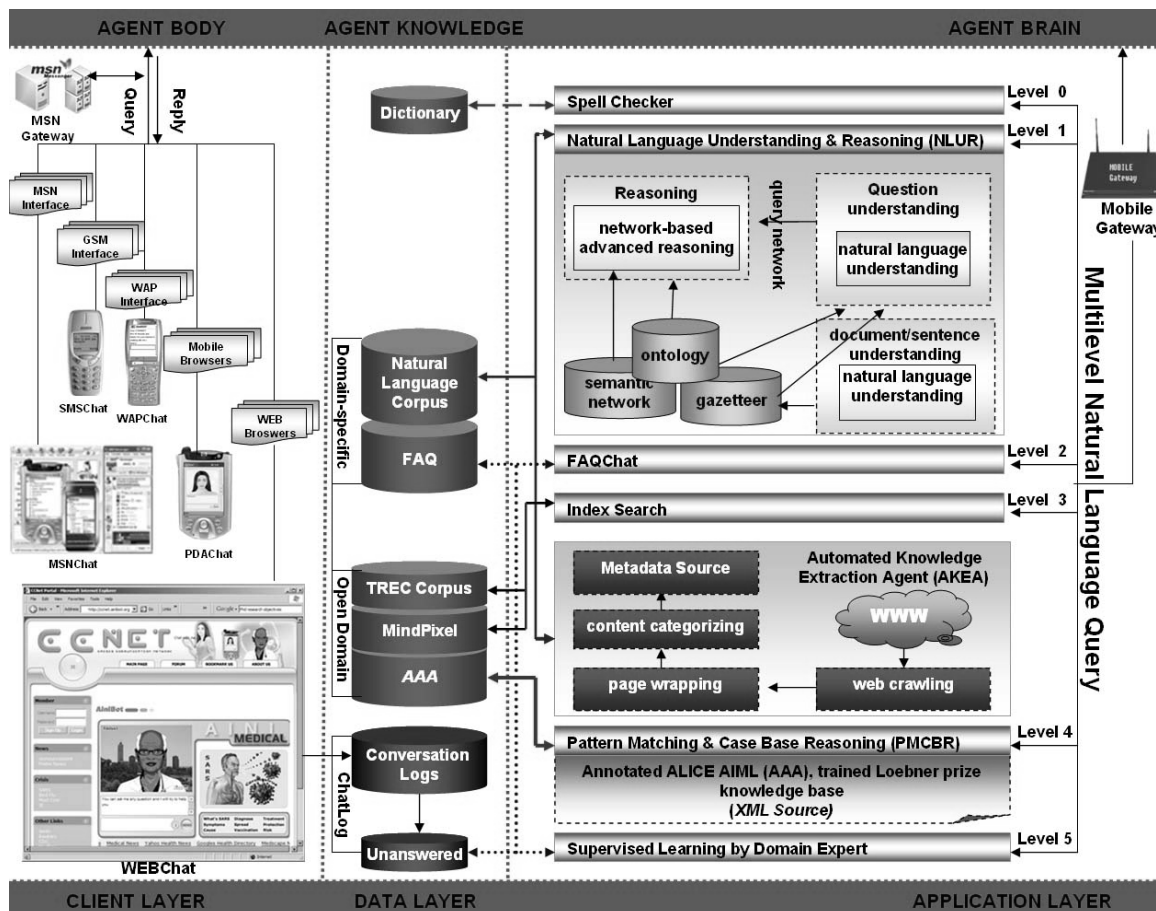


Figure. 1 AINI's ECA Architecture

than as "cheating". After all, in a way, the game is all about deception. It can be seen that Turing considers it possible that a sufficiently human-like machine (i.e., a machine that is sufficiently good at playing the IG is bound to make such mistakes as we attribute to humans.

The chatterbots has already fools "average" questioners. If a larger collection of "tricks" sufficed, would you redefine "artificial intelligence," "average questioner," or "trick?" Perhaps the biggest obstacle to the advancing of this area is there are not many uses for fooling people besides the Turing test [10]

Fifteen years after Turing proposed the imitation game, ELIZA written by Joseph Weizenbaum [11] from MIT and PARRY[12] program from Stanford University demonstrated that "a simple computer program" could successfully play the imitation game by resorting to a few "tricks," the most obvious being to answer questions with questions[4]. The others chatterbots such as PARRY, TINYMUD and Hex also have their own tricks as depicted in the Table 1.

### 3. AINI - ECA Architecture

Artificial Intelligent Natural-language Identity or AINI architecture is shown in Figure 1. The details of the architecture has been reported in previous publications by the authors[1] Basically, the AINI engine comprises a number of knowledge modules in the Data Layer. It has the ability to communicate with three layers: the Application Layer, the Data Layer and the Client Layer. The Client Layer is capable to communicate with the user via different channels such as Web browser, Mobile Browser, WAP Browser and GSM Interface. It can carry on multiple independent conversations at the same time. AINI's knowledge bases are located within the Data Layer. There are a number of modules supporting the Application Layer which governs the manipulation and searching of the answers. The modules use plug-in principles that can quickly be augmented with domain knowledge for specific purposes.

Originally, this research project involves the establishment of an embodied

conversational agent (ECA) based on an AINI architecture [13]. The prototype system is designed specifically for the web and mobile technology as shown in the Client Layer in Figure 1. The complete software agent can be considered as a multi-domain knowledge system with multimodal human-computer communication interface. The query and answer between the user and the computer are communicated via the common protocol TCP/IP. AINI is designed to engage the user with focus on the chosen subject topic. In this particular application, the topic is on the possible pandemic virus, H5N1. AINI communicates with the user in natural language via typed messages. The system is also capable to reply in text-prompts or Text-to-Speech Synthesis together with appropriate facial-expressions on the displayed object which can be an animated avatar or a human face.

As illustrated in Figure 1, AINI employs an Internet three-tier, thin-client architecture that may be configured to work with any web application. It comprises a client layer, an application layer and a data server layer. The hybrid architecture provides features of multimodal interface (Client Layer), multilevel natural language query (Application Layer) and multiple knowledge bases (Data Layer). The process of communication and answering is as follows. 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) using the MINIPAR parser [14]. 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 corpus extracted by Automated Knowledge Extraction Agent (AKEA) [15] MINIPAR parser is capable to parses nearly 500 words per second on a Dell Precision PWS380 Server 3GH with 1GB memory.

### 4. Natural Language Query for the ECA

In ECA design, first and foremost, a good grammar unit is necessary. It seems to be a

good idea to employ more sophisticated natural language processing methods rather than the conventional pronoun transposition. It could be argued that usage of perfect grammar is not crucial since it is quite rare that humans use perfect grammar in informal communication. A good strategy for the TT is indisputably that of trying to maintain human-ness (or at least the neutrality) for as long as possible. It becomes very difficult for the machine to make the interrogator believe that it is human after he/she has his/her mind set on "unmasking" the poor thing. A promising approach is to develop programs that can learn. In fact, work has been done along this line, but not with the intention of producing computer programs to pass the TT.

Our system also differs from other approach because we implemented our chatterbots using Top-down approach for their natural language query [16]. In this multilayer natural language query, plug-in module has been proposed. Our plug-in module consists of spell-checker, Natural Language Understanding and Reasoning, FAQ Metadata and AIML engine module. This plug-in module can be integrated easily and offer a much more scalable approach for the web context. In general our approach includes the following stages:

- **Level 0** - AINI would parse the user's input by reading in sentences one-by-one, and that each sentence would be converted to a list of words to be recognized by spelling checker. Spelling checker will proposed the alternative words from the dictionary if it found any typo(s). If this step is parsed, AINI would attempt to formulate a reply to the sentence by calling next modules;
- **Level 1** - The sentence would be send to Natural Language Understanding and Reasoning Module using Minipar [17]. This module will do the sentence parsing to produce grammatical categories and grammatical relationship such as Noun Phrase and Verb Phrase. Name-entity Recognition will then identify each of the keywords to look for a match in gazetteer for their relation inference and discourse

integration. If this step is successfully parsed, the network-to-path reduction will be carried out by Network-based advanced reasoning. From the query network, the set of sequences *question* from the leaf node to the root node is obtained.

- **Level 2** - If empty string is returned from the database, AINI would then go to FAQ Module using FAQ question-answering system. In this stage, we are ignoring sophisticated natural language processing or logical inference which has already performed previous module. FAQs are Frequently-Asked Questions documents, designed to capture the logical ontology of a given domain or domain-specific.
- **Level 3** - If no database match was found, AINI would then try to go to the next module call Metadata Index Search. Metadata index is information about information: more precisely, it is structured information about resources indexed. It gathers the metadata from pages on the Internet or an Intranet and lets users search the metadata stored in its index. This module relies on the application of a mix of linguistic rules and probabilistic or statistical principles.
- **Level 4** - If this too failed, AINI would then try to detect whether the sentence was a trick question. It would look for common ways judges used to trick programs which is commonly found in ELIZA and ALICE. In this case, AIML technology will respond with suitable witty replies.
- **Level 5** - If even this module failed to work (which could happen if all of its witticisms had already been used), then AINI routine would be invoked randomly to generate reply. In this stage, the result checking will still be the responsibility of the domain expert, who will be able to correct errors and to pick a proper matching response from a list of possible matches from the conversation log which is unanswered by AINI. This undoubtedly will lead to better performance and accuracy in the future conversation.

In our experiment<sup>1</sup>, we created an amusing and informative example of how chatterbots converse between them in the wild with the simulated "conversation" between Eliza<sup>2</sup>, ALICE<sup>3</sup> and AINI as depicted in the Figure 2.

```

Eliza :Hello. I am ELIZA. How can I help
       you?
AINI  :Hi there! Hi, ELIZA.
ALICE :OK I will call you ELIZA. Just keep
       talking to me as long as you can.
AINI  :My name is Aini. Gosh! Chat amongst
       yourselves.
ALICE :I'm pleased to introduce myself to
       you, Aini. Do you have any
       conditions I should know about?
Eliza :Oh, i have any conditions you
       should know about.
AINI  :Aha. Cool. I have a friend named
       E.L.V.I.S.
ALICE :Next question? That's nice. I have
       a very nice programmer.
AINI  :Do you have any hobbies? You can
       click on the button on the bottom
       right corner to get more news. I
       have a great botmaster.
Eliza : Oh, i can click on the button on
       the bottom right corner to get more
       news you have a great botmaster.
ALICE : And? Oh you can, can you?
Eliza :Oh, i can can i.
AINI  :Oh you can, can you?
ALICE :I see. No I don't think I can do
       it.
ELIZA : Why do you not think you can do
       it?

```

**Figure 2. Conversation logs between three chatterbots**

In this conversation, most of the AINI's knowledge is extracted from the AAA knowledge base. This is because, during the conversation between chatterbots, majority of the questions are based on the open-domain knowledge base, instead of domain-specific. In addition, from the conversation logs, how well does the chatterbots really remember the thread of your conversation? Many chatterbots will ask your name, and incorporate this into future responses. Some bots also attempt to identify the "topic" of conversation, and a few will be able to tell you what the topic is when you ask about it (although even this is rare). Just remember that asking your name or any other piece of information and storing it in a variable

<sup>1</sup> The experiments conducted in this paper were performed by the authors in the period of January – June 2007.

<sup>2</sup> <http://www-ai.ijs.si/ELIZA/ELIZA.html>

<sup>3</sup> <http://www.alicebot.org>

for later use doesn't constitute "understanding" what's going on in any sophisticated sense, nor is it the basis for "reasoning" in any sense of the word. Many companies try to pass off tricks like this as a kind of "memory" or "knowledge of context".

In the last conversation, you could see how well does the chatterbots really understand "anaphora"? Anaphora is references to previously-named things. In English, pronouns like "I", "me" "you", "he", "she", "it", and so on usually refer to previously named people or things. Some chatterbots companies will pretend that their software can really identify the target of an anaphoric reference. But such claims can be proven false, and can easily be shown to be keyword-based as well. You need only pursue a dialog a few turns to show that a chatterbots really has no clue what you're talking about, beyond a rudimentary index of the "current topic". You can try all the tricks you want; some of them will get you witty answers, a few will even be answered "correctly" (until you twist the trick a bit more), and most will just get you a plain stupid answer.

## 5. Conclusion

This paper has given an overview of the imitation game, Turing test and ECA. This paper also provided an analysis into the intelligence of ECA and shows the feasibility of our own ECA, AINI. From the discussion, it can be seen that with the top down design approach, our ECA can converse reasonably well as compare to ELIZA and ALICE. This has indicated that our AINI has potential to be used as a natural interface for most interactive media applications like game.

### References:

- [1] O. S. Goh, K. W. Wong, C. C. Fung, and A. Depickere, "Towards A More Natural and Intelligent Interface with Embodied Conversation Agent," in *International conference on Game research and*

- development*, Perth, Western Australia, 2006, pp. 177 - 183.
- [2] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, *Embodied Conversational Agents*. Cambridge, MA: MIT Press, 2000.
- [3] S. L. Oviatt, C. Darves, and R. Coulston, "Toward Adaptive Conversational Interfaces: Modelling Speech Convergence with Animated Personas," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 11, September 2004.
- [4] J. Weizenbaum, "ELIZA - A computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 9, pp. 36-45, 1966.
- [5] Alice, "Artificial Linguistic Internet Computer Entity." vol. 10 April 2005, 2005.
- [6] J. Hutchens and M. D. Alder, "Introducing MegaHAL," in *The Human-Computer Communication Workshop*, 1998, pp. 271-274.
- [7] H. Jenkins and J. Cassell, "From Quake Grrls to Desperate Housewives: A Decade of Gender and Computer Games," in *Beyond Barbie and Mortal Kombat: New Perspectives on Gender and Computer Games.*, Y. Kafai, C. Heeter, J. Denner, and J. Sun, Eds. Cambridge, MA: MIT Press, in press.
- [8] J. Cassell and H. Jenkins, *From Barbie to Mortal Kombat: Gender and Computer Games*. Cambridge, MA: MIT Press., 1998.
- [9] H. G. Loebner, "Home page of the Loebner prize—the first Turing test," Loebner Prize, 2007.
- [10] M. L. Mauldin, "Chatterbots, Tnymuds, And The Turing Test: Entering The Loebner Prize Competition," in *AAAI-94*, 1994.
- [11] J. Weizenbaum, "ELIZA—a computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 9, pp. 36-45, 1996.
- [12] K. M. Colby, *Artificial Paranoia: A Computer Simulation of Paranoid Processes*: Pergamon Press Inc., 1975.
- [13] O. S. Goh, C. C. Fung, A. Depickere, and K. W. Wong, "An Analysis of Man-Machine Interaction in Instant Messenger," in *Advances in Communication Systems and Electrical Engineering*. vol. 4, X. Huang, Y.-S. Chen, and S.-I. Ao, Eds. US: Springer, 2008, pp. 197-210.
- [14] D. Lin, "Dependency-based Evaluation of MINIPAR," in *Workshop on the Evaluation of Parsing Systems*, Granada, Spain, 1998.
- [15] O. S. Goh and C. C. Fung, "Automated Knowledge Extraction from Internet for a Crisis Communication Portal," in *First International Conference on Natural Computation* Changsha, China: Lecture Notes in Computer Science (LNCS), 2005, pp. 1226-1235.
- [16] O. S. Goh, C. C. Fung, K. W. Wong, and A. Depickere, "Multilevel Natural Language Query Approach for Conversational Agent System," *International Journal of Computer Science*, vol. 33, pp. 7-13, 2007.
- [17] R. C. Berwick, in *Principle-Based Parsing: Computation and Psycholinguistics*, S. P. Abney and C. Tenny, Eds.: Kluwer Academic Publishers., 1992.