College Enquiry Chatbot Using A.L.I.C.E

Balbir Singh Bani, Ajay Pratap Singh

Abstract— **In this paper, a proposal is carried on to explain the design of a chat bot specifically tailored as an application which is going to help new students to solve all the problems they face and the questions which arises in their mind during and after the admission . In particular, the proposal investigates the implementation of ALICE chat bot system as an application named as college enquiry chat bot. A keywords-based human-computer dialog system makes it possible that the user could chat with the computer using a natural language, i.e. in English.**

*Index Terms—***About four key words or phrases in alphabetical order, separated by commas.**

I. INTRODUCTION

A.L.I.C.E. (Artificial Linguistic Internet Computer Entity) is an award-winning free natural language artificial intelligence chat robot. The software used to create A.L.I.C.E. is available as **free** ("open source") **Alicebot and AIML** software.

Try **talking to A.L.I.C.E.** just like a real person, but remember you are really chatting with a machine! A.L.I.C.E.'s Alicebot engine utilizes AIML (Artificial Intelligence Markup Language) to form responses to your questions and inputs.

Unlike other commercial chat robot software costing thousands of dollars, the Alicebot engine and AIML are freely available under the terms of the **GNU General Public License** (used by GNU/Linux and thousands of other software projects). The A.L.I.C.E. project includes hundreds of contributors from around the world.

We worked with A.L.I.C.E (Artificial Linguistic Internet Computer Entity) to develop our chatterbot system.

A. What is Domain-specific knowledge system?

Domain specific knowledge system is a set of knowledge base consisting of AIML files. In our knowledge system we included AdmissionInfo.aiml and collegeguide.aiml, these files consists of different patterns and templates that relates to environment and generates responses to them, this along with modified conversational knowledge base in ALICE helps our chatterbot to stay focused only on chat conversational topic

B. Why A.L.I.C.E.?

ALICE open source chatbot developed by Dr.Wallace, which is based on natural language understanding and pattern matching; also the architecture of the "chatbot engine" and the "language knowledge model" are clearly separated, which gives us the opportunity to easily inject a newly developed knowledge model such as our FAQ knowledge model.

In this paper we will further give overview of ALICE and how the knowledge base of ALICE is implemented using AIML files, we will also look at some related work with ALICE and learn how to keep the system focused on a specific domain.

C. What is A.I.M.L.?

AIML stands for **Artificial Intelligence Markup Language**, but it is just simple XML.AIML is an XML based markup language meant to create artificial intelligent applications. AIML makes it possible to create human interfaces while keeping the implementation simple to program, easy to understand and highly maintainable. AIML was developed by the Alicebot free software community and Dr. Richard S. Wallace during 1995-2000. AIML is used to create or customize Alicebot which is a chat-box application based on A.L.I.C.E. (Artificial Linguistic Internet Computer Entity) free software.

Following are the important tags which are commonly used in AIML documents.

S.No.	AIML Tag / Description
1	**\<aiml\>** Defines the beginning and end of an AIML document.
2	**\<category\>** Defines the **unit of knowledge** in Alicebot's knowledge base.
3	**\<pattern\>** Defines the pattern to match what a user may input to an Alicebot.
4	**\<template\>** Defines the response of an Alicebot to user's input.

EXAMPLE

```
<aiml version = "1.0.1" encoding = "UTF-8"?>
   <category>
      <pattern> HELLO ALICE </pattern>
        <template> Hello User!
         </template>
          </category>
           </aiml>
```

RESULT
User: Hello Alice
Bot: Hello User.

Balbir Singh Bani,
Ajay Pratap Singh,

II. COMPARISON

A. ALICE used a simple pattern template to represent input and output, and also using simple pattern matching algorithm. Between Elizabeth uses Input rules, keyword patterns and output rules to generate a response.

B. The recursive techniques used in ALICE is considered as a power point of the system, it is used for simplifying the input by calling match categories recursively. Contradictory, the nature of some rules in Elizabeth may cause cycling or iteration, which is solved by applying the rule only once if it is applicable for the active text more than 10 times in succession.

C. In ALICE there is the ability to combine two answers in the case of splitting happened within Normalization Process, or the partitioning caused by the recursive process. The recursive process provide a way to partition the sentence to two sentences then combine their results which is not available by Elizabeth.

D. The most important and strong issue in ALICE is the pattern matching algorithm, which is easy and depend on depth first search. This algorithm try to find the longest pattern matching between Elizabeth gives the response according to the first Keyword pattern matched.

E. Both systems can change personal pronouns, a lot of complicated appear in Elizabeth related to writing some rules in upper case and others in lower, which may cause a lot of errors and give unsuitable answers. Also both systems allowed memorization for the previous input, output for further using. But Elizabeth allows other actions to occur while the conversation is under progress, which is called the dynamic process like adding, modifying and deleting script commands.

F. If the same input repeat during the conversation, Elizabeth try to give different answers by using different random selection responses from the respond list. ALICE sometimes gives you the same answers or different ones.

G. Elizabeth has the ability to give the derivation structure for a sentence using the grammatical analysis which is not provided by ALICE.

H. Elizabeth allows the user to create his ownscript files and it incorporates analysis tables for all steps in matching, helping the user to understand how this answer is generated.

III. CONCLUSION

Both systems are chatbot system that are adopted from Eliza program and each one has advantages and disadvantages. Alice stored a huge corpus text and Elizabeth provides the grammatical analysis for sentences, both of them give a good introduction to Artificial Intelligence, Natural Language Understanding and Pattern Matching. The main inference point is it will be easier to build a machine learning for ALICE since it uses simple patterns, templates to represent input and output. Elizabeth uses more complex rules for which you need to write input transformation, Output transformation and keyword patterns to represent a user input and Elizabeth answer. You can do that by simple pattern template using ALICE. Another important point used by ALICE is the ability to partition the user input of splitting it in two sentences and then combines the answers. This is an important issue in language processing. Elizabeth does not provide this facility. According to Elizabeth structure it will be difficult to do splitting.

REFERENCES

[1] R.S. Russell, Language Use, Personality and True Conversational Interfaces. AI and CS, University of Edinburgh, Edinburgh, 2002.

[2] Jia, "The Study of the Application of a Keywords-based Chatbot System on the Teaching of Foreign Languages," University of Augsburg, Augsburg, Germany 2002.

[3] AIML Based Voice Enabled Artificial Intelligent Chatterbot, International Journal of u- and e- Service, Science and Technology Vol.8, No.2 (2015), pp.375-384

[4] M. Awad and L. Khan, "Web Navigation Prediction Using Multiple Evidence Combination and Domain Knowledge," IEEE Transactions on Systems, Man and Cybernetics, Part A, vol. 37, pp. 1054-1062, 2007.

[5] R. P. Schumaker and H. Chen, "Leveraging Question Answer Technology to Address Terrorism Inquiry," Decision Support Systems, vol. 43, pp. 1419-1430, 2007

[6] S. Potter, "A Survey of Knowledge Acquisition from Natural Language," in TMA of Knowledge Acquisition from Natural Language, vol. 2003. Edinburgh, 2001

[7] Weizenbaum, J. 1966. ELIZA- A Computer Program for the Study of Natural Language Communication between Man and Machine, CACM 9(7), 36-43

[8] http://www.chatbots.org/virtual_assistant/

[9] https://developer.pandorabots.com/docs

[10] IEKA.com help centre Bot. http://www.ieka.com

[11] Krusschwitz, U., De Roeck, A., Scott, P., Steel, S., Turner, R., and Webb, N. (2000). Extracting semi structured data lessons learnt. In proceedings of the 2nd international conference on natural language processing (NLP2000), pages 406-417.

[12] Wallace, R.S. 2000. Symbolic Reductions in AIML, http://alicebot.org/srai.html