University Chatbot using Artificial Intelligence Markup Language

Cho Zin Oo, Cho War, Hnin Ei Ei Cho, Phyo Hay Mar Wai, May Myat Myat Khaing, Hay Man Htun

Abstract—**Chatbots are conversational systems that can do chat interactions with human automatically. It is developed to be virtual assistant, making entertainment for people, helping for answering the questions, serving as human partner in smart homes etc. Most of the chatbots utilize the algorithms of artificial intelligence (AI) in order to get the required responses. In this project, we provide the design of a University Chatbot that provides an efficient and accurate answer for any user questions about university information. This is the University Chatbot for inquiring about school information in Myanmar Language based on Artificial Intelligence Markup Language.**

Index Terms—**AIML, Natural Language Processsing, Pattern Matching, Response Generation**

I. Introduction

Conversational agents become essential by interacting of machines with the desired users to provide natural language interfaces. So, the role of chatbots in the information technology and communication is widely in used. Many chatbots are created day by day through marketing, medical, education and banking. Chatbot is also a user assistant substance that is intended to produce a communication with human through their regular language. The bot would match the input sentence from the user with that pattern existed in the knowledge base. This system is simple Myanmar chatbot using AIML but it can answer the necessary information for the users. There are many chat engines with different methods and can perform chatting. Some famous chatbots are SimSimi, Mitsuku, A.L.I.C.E, and now the machine learning chatbots like Siri, Alexa, Cortana and so on. Although modern chatbots apply the power of artificial intelligence to answer complicated questions, they still need some improvements for low resource languages.

II. Human Computer Interaction in Chatbots

Human Computer Interaction is a communication field of study focusing on the design of computer technology and the interaction between human and computers. Conversation system between a human and a computer is either chatting by typing text or speech dialogue using the voice. Thus, interaction with natural language is a feasible option for connecting machine agents and human users. Human Computer Interaction may need to consider for chatbots as the main object of design, focus on services than user interfaces, and design for interaction in networks of human and machine actors. The main parts which include human computer interaction in conversation systems design are (a) the techniques used to produce keywords, (b) pattern matching techniques used inside the chatbot and (c) the type of response. So, HCI is considered taking on human-chatbot interaction design as an area of research and practice.
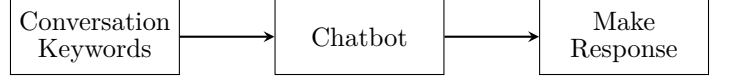

Fig. 1: Human Computer Interaction

III. Artificial Intelligence Markup Language

AIML is an XML based markup language for specifying chatbot content. It was created by the ALICE bot free software community in 1995-2000 for the people to input dialogue pattern knowledge into chatbots based on the ALICE free software technology. An AIML Interpreter is able to load and run the bot, then, it provides the bots responses in a chat session with a user. AIML consists of data objects called AIML objects, which are made up of units called topics and categories. The topic is called an optional top-level element, it has a name and a set of categories related to that topic. Categories are the basic units of knowledge in AIML. One category is a rule for matching an input and converting to an output, and consists of a pattern, which represents the user input, and a template, which responses the answer. The AIML pattern is simple and consists of words, spaces, and the wildcard symbols, _and *.

```
<?xml version="1.0" encoding="UTF-8"? >
<aiml version="2.0" >
<!AIML code goes here − >
</aiml >
```

A. AIML Categories

There are three AIML types: (1) atomic categories, (2) default categories, and (3) recursive categories.

1) Atomic categories are those with patterns that do not have wildcard symbols, _and *.
2) Default categories include wildcard symbols_and *.
3) Recursive categories are the categories with templates <srai> and <sr> tags, which represent recursive artificial intelligence and symbolic reduction. Applying a combination of wild cards and srai, the stop words of the sentences can be carefully checked out from the user input. Recursive categories involve many applications: (i) symbolic reduction which reduces the complex grammatical forms to simpler ones; (ii) divide and conquer category splits an input into two or more subparts and add the responses to one; (iii) synonyms resolution is possible to appear different words with the same meanings depending on the consisting text; (iv) keyword detection is possible to find the same response when a definite keyword is found in the user input. These are some examples of different AIML categories.

IV. Proposed System

We have implemented a Myanmar interactive chatbot for university frequently asked questions. AIML is defined
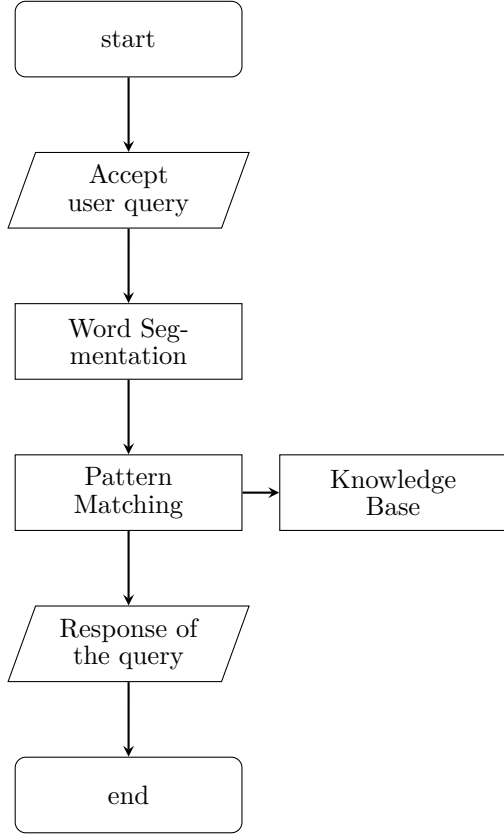
Fig. 2: Flowchart of the System

with general inquiries and messages which are replied by applying AIML formats. According to the Artificial Intelligence Markup Language, we have used different AIML tags to get the user required information from the bot.

TABLE I: AIML Tags used in System

No	Tags used for AIML categories
1	\<category>\</category>
2	\<pattern >\</pattern>
3	\<template >\</template >
4	\<srai >\</srai >
5	\<random >\</random >with \ \
6	\<set >\</set >
7	\<get >\</get >

There are three steps in our project. In the first step, the question is entered by the user. In the second step, the system performs word processing actions to match the users input to a pre-defined format and do the pattern matching between user input and the Knowledge Base. Finally, the answer is presented to the user in the third step.

A. Knowledge Base

Artificial Intelligence Markup Language is a well-known XML derived language to build chatbot knowledge base.

Users' frequent asked question sets are defined semantically the knowledge domain given to the chatbot. The questions are available from the university academic center related to five topics. We have used 300 question-answer pairs as data distribution. A well-designed knowledge base can positively impact the effectiveness of chatbots that will improve the interaction between users.

V. Experiment

Rule-based chatbot contains a faster time-to-relevance,delivering a faster impact on user interaction. This chatbot is one of rule-based chatbot and developed on AIML language for the University of Technology(Yantanarpon Cyber City), Pyin Oo Lwin. We have purposed this system to have a support for university routine.All the questions files need to be uploaded to Pandorabots server. The files include the university related questions and information that the students, teachers and parents frequently asked. The number of question-answer pairs in the system thatare utilized for different topics and type of categories are Greeting, Location and Address, Academic, Brief History, Conference, Faculty and Staff and Library.

The user needs to input the questions as Myanmar Language. The segmentation is done by using the UCSY word segmentor. The input from the user is normalized and processed on the Pandorabots server. The AIML files, which are separated into several categories and the chatbot's knowledge is uploaded into the server. After pattern matching, the inquired user can ask the university related questions about academic services and activities. Figure 3 and 4 show sample outputs of the system. Figure 5 shows mismatch question example of the system. Our chatbot can't answer the users' typing error or data which are not in training. eg. Question: ကျောင်းသားရေးရာကဘယ်နားမှာလဲ Answer: Main Building ပထမထပ်မှာပါ. But, if the user asked a question which has typing error such as ကျောင်းသားရေးယာကဘယ်နားမှာလဲ Answer: ထပ်မံကြိုးစားကြည့်ပါရှင်.

VI. TESTING AND E VALUATION

Testing can be made to measure the quality of chatbot. The steps included to conduct the chatbot experiments are inquiring the user questions related to the six topics show in Table 2 and the experimental analysis whether they have correct or wrong responses. The system still needs improvement due to some patterns that mismatch with the chatbot knowledge. We also have adopted the dialog efficiency matrix to evaluate the chatbot.

TABLE II: AIML Tags used in System

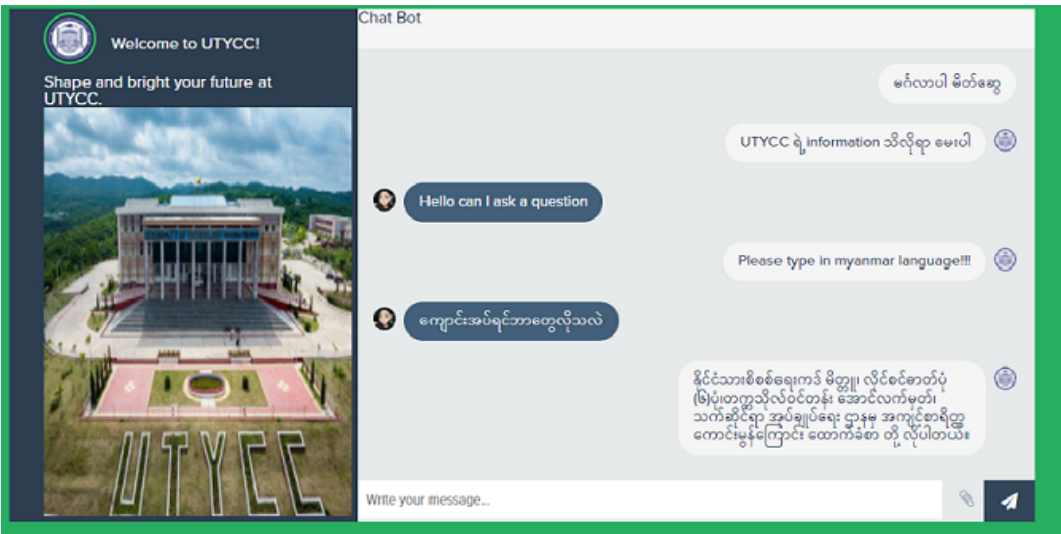| Topics	No.Q	Match	MisMatch	No.Ans
Location and Address	5	3	0	2
Academic	10	9	1	0
Brief History	10	6	2	2
Conference	5	3	1	1
Faculty and Staff	5	3	1	1
Library	5	2	3	0
Total	40	26	8	6

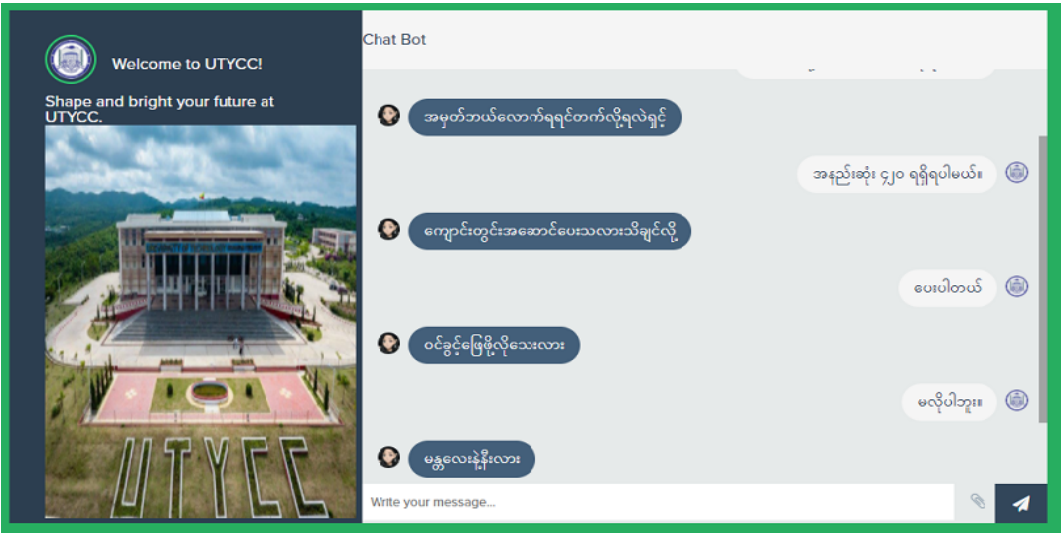Fig. 3: Sample Outputs of the System

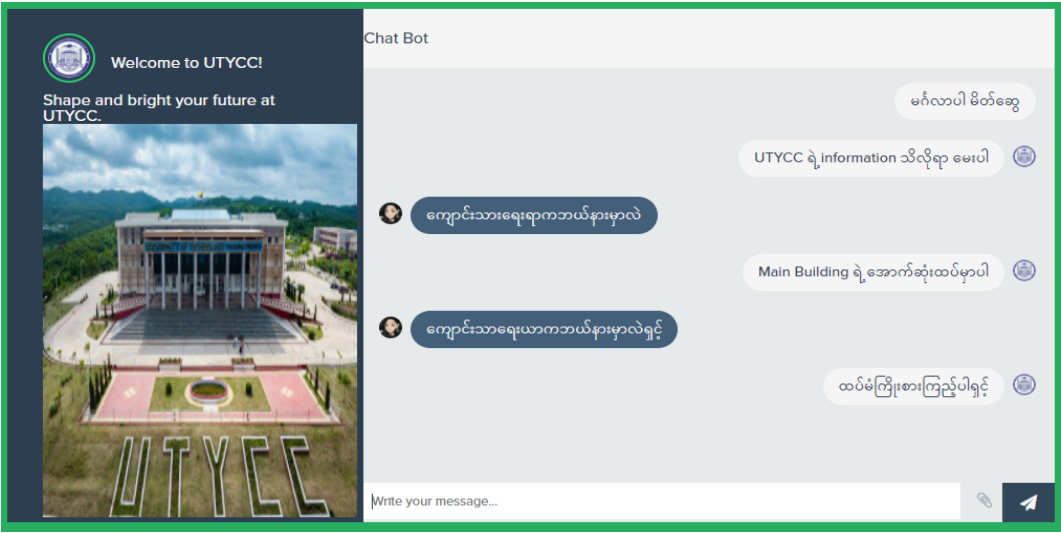
Fig. 4: Sample Outputs of the System

Fig. 5: Mismatch Question Example

VII. Conclusion

Chatbots can interact with people in effective ways. There are many chatbots in English and other languages by using different algorithms and models but there is little chatbot using Myanmar language. This is one of the University Chatbots using Myanmar Language to fulfill the information gaps between the university and its related users. Now, we have implemented a chatbot for the University of Technology(Yantanarpon Cyber City), Pyin Oo Lwin. This is simple chatbot using Artificial Intelligence Markup Language. The user can ask the useful questions about the university related the academic sectors through chatbot. The bot will help people to save time and get the information every time. We still need some improvements for the bot.