Toward the implementation of a Topic specific Dialogue Based Natural Language Chatbot as an Undergraduate Advisor

Conference Paper · May 2013

DOI: 10.1109/ICIEV.2013.6572650

2 authors:

Supratip Ghose

Jagat Joyti Barua
University of Information Technology and Sciences

Toward the implementation of a Topic specific Dialogue based Natural Language Chatbot as an Undergraduate Advisor

Supratip Ghose,[1,*] Jagat Joyti Barua,[2]
[1] Dept. of Computer Science and Engineering, University of Information Technology and Sciences
GA - 37/1 Pragati Sharani, Baridhara J-Block, Dhaka 1212Baridhara, Dhaka, Bangladesh
[2]Dept. of Computer Science and Engineering, University of Information Technology and Sciences
Baridhara, Dhaka, Bangladesh
*SGresearch@gmail.com

***Abstract*—In this work, we explain the design of a chat robot that is specifically tailored for providing FAQBot system for university students and with the objective of an undergraduate advisor in student information desk. The chat robot accepts natural language input from users, navigates through the Information Repository and responds with student information in natural language. In this paper, we model the Information Repository by a connected graph where the nodes contain information and links interrelates the information nodes. The design semantics includes AIML (Artificial Intelligence Mark up Language) specification language for authoring the information repository such that chat robot design separates the information repository from the natural language interface component. Correspondingly, in the experiment, we constructed three experimental systems (a pure dialog systems associated with natural language knowledge based entries, a domain knowledge systems engineered with information content and a hybrid system, combining dialog and domain knowledge). Consequently, the information repository can easily be modified and focussed on particular topic without recreating the code design. Experimental parameters and outcome suggests that topic specific dialogue coupled with conversational knowledge yield the maximum dialogue session than the general conversational dialogue.**

***Index Terms*—Chatbot, AIML, Conversational knowledge, focussed topic**

I. Introduction

The World Wide Web has grown into a rich repository of information in a distributed manner. It is a great trade-off for the information revolution that end-users are finding it challenging to locate relevant information and services quickly and easily. The web changes from static to dynamic and provides a meaningful web. We need social aware tools and social matrix and collective intelligence [16], so that, human and artefact collaborate meaningfully to lessen the burden of user searching and browsing in social platform.

Given a specific domain of interest and its audience pool, there are two important aspects of a networked knowledge transfer platform and in human computer interaction [15][17]. We have question answering system[1] as a knowledge delivery platform in which the expert delivers knowledge for the solicitation of the user. Another is knowledge acquisition in the form of online forum and social web platform in which different dialogues form a knowledge repository.

There is a growing interest in ChatBot interface that takes into account Chatbot discourse design and knowledge delivery. One platform for knowledge delivery is a lightweight dialog system Chatbot that will hold the user's attention with human-like responses. ChatBots are computer programs that interact with users in natural language although different names are adopted like virtual agent, dialogue system, chatterbot in different programming architecture. The ALICE system (Artificial Linguistic Internet Chat Entity) was created by Richard Wallace [6] uses an XML dialect called AIML (Artificial Intelligence Markup Language) to store patterns and responses upon encountering the pattern in a user dialog. Its standard distribution comes with approximately 24,000 pattern covering assorted geography, nature, and human interest facts. According to Wallace, concepts of Stimulus-Response and minimalism were designed in the AIML code. Moreover, this paper investigates how a chatbot can be programmed as a guide and how can it be used as an undergraduate Advisor for a university. Naturally, it can be extended in daily life, such as help desk tools, automatic telephone answering systems, tools, to aid in education, business and e-commerce [1][2]. Web-based chatbot system can provide an easy, natural extension to knowledge acquisition and recommender systems.

In this paper, a proposal is carried on to explain the design of a Chat Bot, specifically tailored as an undergraduate student information system that helps students in a University with admission and course information. In particular, the proposal investigates the implementation of ALICE Chatbot system as a domain specific chatterbox named University FAQbot (UFAQbot). It is assumed that a Chatbot will perform as an undergraduate advisor and act as a student counselor to frequently asked question [13] in question-answering system [18] of students. Our work will show on how a chatBot can work as domain specific information system and experiments on how the system's accuracy could be improved based on a specific domain. The implementation of this project on a University environment is particularly useful for students' looking for information regarding admission in a University, and its course curriculum. Even though most of the information is available on the web, students often like to have personal interaction with the advisor. In such an environment, a chat robot could be designed to provide academic advice. The main goal of such a system is to conveniently retrieve information without having to look or browse several web pages to fetch answers to frequently asked questions. This paper describes our use and evaluation of ALICE as a knowledge delivery and acquisition platform in the form of UFAQBot Chatbot in experiments of specific academic domain. The evaluation depends on the development of new metrics, which we describe to analyze ALICE's performance in the user sessions. In the remainder of the paper, we describe related work, including quite a few implemented systems in the Literature review. We then describe our research questions as well as our undergraduate Advisor system.

II. Related Works

Chatterbot development is reasonably well studied ever since the Turing Imitation Game (TIG) [9] was first proposed. Eliza [7] was the first famous chatbot, and ALICE [5] was another milestone. The Loebner Prize [10] and The Chatterbox Challenge [11] are both annual competitions which have their roots in TIG. However, these are typically text-only experiments, although some limited visual components are often added. This focus is on, however, whether with the text exchange alone, we can replicate human "behavior". The purpose of a chatbot system is to simulate a human conversation; the chatbot architecture integrates a language model and computational algorithm to emulate information chat communication between a human user and a computer using natural language.

With the improvement of data-mining and machine-learning techniques, better decision-making capabilities, availability of corpora, robust linguistic annotations/processing tools standards like XML and its applications, chatbots have become more practical in daily life applications such as help desk tools, information retrieval tools, automatic telephone answering systems, advertising, tools to aid in education, business and E-commerce. In E-commerce, chatbot helps in information retrieval tasks, such as for searching and browsing, as menu based navigation poses difficulties in locating the appropriate information. The dialogue system provides additional information on products and simplify decision making process to find a product that satisfy customer's requirements [1][2]. According to Dr. Wallace, perhaps, the biggest market of chatbot is Entertainment Markets, in which, we can imagine that chatbots can act as a talking book for children and provide foreign language instruction or can be a tutor in Intelligent Tutoring system. One such study used an ALICE system to help Chinese university students practice their conversational English skills. The study was qualitative in nature and used pre-existing conversational English skills [3]. The study focused more on user attitudes rather than on chatterbot efficiency. It was discovered that 62% of users chatted for 10 lines or less, and that 8.5% of the time ALICEbot has no specific pattern to match the given input and had to rely on root-level generic responses. In all of these conversational entities, one thing is common; and that is, they are having the difficulty of maintaining dialogue for sustainable period of time. Another tutoring study focused on using ALICE as a course enhancement tools with Social and Political Theory knowledge[4]. This study found that most subjects used the system as a search engine rather than as a conversation partner. It was further concluded that their system was unable to function as a stand-alone tutor. Dialog system can adequately carry out the conversations with the user and can log the conversations which can be good source for knowledge acquisition for domain specific topic. Therefore, techniques of knowledge acquisition were rightly used in [15][17] with their system AZ-ALICE chatbot that is an extension in ALICE chatbot. They tested their system in a domain specific knowledge acquisition as testbed with a number of participants and it was shown that mass knowledge acquisition improved the domain-specific chatterbot responses. They hypothesized that use of a chatterbot as a knowledge acquisition tool appears to be a stable instrument in gathering both conversation and domain-related knowledge and showed with statistics that it would show a promising future in domain restricted areas. The contribution of this paper over their system is that we have designed the code based on an architecture that maps the general dialogue system into domain specific chatterbot using existing AIML code repository. The dialogue design in our system helps the chatterbot to prolong the conversation with the user on a topic specific session with our semantic codebase design and ultimately serves as a knowledge delivery platform for a user chatting with topic focussed chatbot. The rest of the sections are designed for the reflection of this contribution. Section 3 discusses on design of system and sample dialogue with ALICE, keeping our objective on mapping in mind. Following that we present our experimental design and an analysis of the experimental results. Finally, we conclude our study and provide details of further development for this platform.

III. DESIGN OF THE SYSTEM

The design of the paper mainly involves around the semantics of Alice CodeBase. There is an "art" to writing AIML, and creating a robot character, especially when writing default responses. This required a lot of imagination and sense of humor. In Alice, BotMaster can exercise precise control over a conversational flow [12].

ALICE uses XML knowledge bases to match user input against a predefined response set. The shortcoming of this system is that it cannot adequately answer all of the queries given to it. ALICEbots are also able to expand their present knowledge bases through XML-base AIML (Artificial Intelligence Markup language). These would imply that ALICEbots could be given 'expert appearance' within a particular domain of knowledge[6]. This expansion has already been witnessed in the areas of foreign language fluency, and specific domain-related knowledge fields which can either be supervised by a Botmaster or unsupervised where knowledge is gathered en masse from trusted sources. For engineering purpose, in many Alicebots, it would be desirable to domain knowledge Engineering by using tags. ALICE has three types of categories. The first being Atomic category, contains a pattern representing a simple user input and a template implies robot response. We use an extensive use of default categories for the domain engineering. Default categories are those with patterns that have a wildcards " $*$ " or "_". These patterns result from a reduction process, while the robot searching for the appropriate matching. These wildcards match any input but they are differing in their alphabetical order.

The critical step of domain engineering lies through extensive extension of recursive tags. For initializing episode, we selected six AIML files and among those personalization.aiml, that.aiml, brain.aiml, knowledge.aiml are modified with UFAQBot [rep] design in the experimentation with the below mentioned design architecture.

We designed a chat robot that meets the requirements specified in the Introduction section. Figure 1 shows the components of the proposed chat robot. The main components

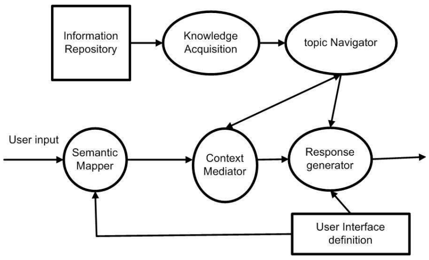

Fig. 1. Example of Alice Codebase

include a Semantic Mapper that maps a user input into semantic elements, a Context Mediator that comes into action if conflicting semantic elements are detected, a Topic Navigator that guides the chat robot through the Information repository and Response Generator that generates the natural language output. The Information repository is a set of Information Nodes. Each node has a set of patterns and templates. In addition, each Information Node has a set of links leading to other information Nodes. The data structure represents a link contains the name of the connecting Information Node, a weighted edge preference value, set of prompt texts and input patterns. Information repository is specified in AIML like language and can be easily translated to AIML. We describe this representation scheme further. The information repository IR can be formally defined as follows: $IR = (E, e_0, L)$, in which, E is a set of episodes E_1,E_2,$E_3 \ldots E_n$ and L is a set of links l_1,l_2,$l_3 \ldots E_n$, in which m and n are finite. Each link $l \in L$ starts from an episode e_i and ends in another state e_j, in which $(e_i \in (E \cup e_0)$ and $e_j \in (E \cup e_0)$ and could be represented by $(e_i, e_j), p_i j)$, in which, $p_i j(0 \leq p_i j \leq 1)$ and $\forall_j \Sigma p_i j = 1$. Here e_0 is a special episode that corresponds to the outside world (and is in fact a set of episodes defined Alice's brain). The respond generator makes use of the weighted preference values to create the robot's response. However, weighted preference value does not directly control the navigation. Navigation depends mainly on user input.

Topic Navigator is responsible for making the transition t from (IR,c_i) to (IR,c_j), in which the current episode is $c_i \in (E \cup e_0)$ and the target episode is $c_j \in (E \cup e_0)$ The target episode depends on the semantic element determined by the Semantic Mapper. Semantic Mapper directs all semantic elements reduced from user input to the Context mediator. This work aims to provide a manual approach of modifying the ALICE recursive category. This is achieved with the AIML code principal of reductionism, in which, Symbolic reduction refers to the process of simplifying complex grammatical forms into simpler ones. Usually, the atomic patterns in categories storing robot knowledge is gathered in the simplest possible terms, for example, we tend to prefer patterns like "Content in CSE370?" to ones like " What information do you know about CSE370" when storing information about a particular course. Originally, AIML interpreter try to match word by word to obtain the largest pattern matching and selects the best one. The behavior can be described in terms of the class Graphmaster which has a set of nodes call Nodemappers that map branches from each node and branches represents the first words of all patterns and for wildcards. So it is parent child relationship. If there are conflicting semantic elements, the Context mediator would resolve the conflicts possibly by further dialog with the user.

User: *Dr. Supratip is a great faculty*
FAQ bot: *Who's your favourite faculty?*
User: *Dr.Supratip*
FAQ bot: *Dr. Supratip, Oh you know what courses does he take?*

Fig. 2. User interactions with Alice

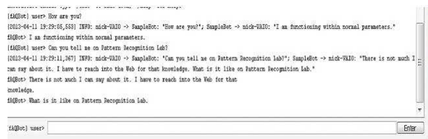

Fig. 3. Program D System interface

Once the semantic element has been successfully determined, the Context mediator submits the semantic element to the topic Navigator. Topic Navigator attempts to find a target episode corresponding to that semantic element. The AIML [6] has several tags for setting a current episode and redirecting to that topic for a user response. One such tag is $< that >$ tag which can track back the previous utterance of Robot. One such code is as followed in Figure 2:

The dialogue above can be generated by $< think >< setname = "topic" > Me < /set >< /think > < that > Whoisyourfavoritefaculty? < /that >$

Using the $< think >$ tag [6], which causes Alicebot to perform whatever it contains but hide the result from the user, the Alicebot engine will set the "topic" in its memory to "Me". This allows any categories elsewhere with an explicit "topic" value of "ME" to match better than categories with the same patterns that are not given an explicit topic. This illustrates one mechanism whereby a Botmaster can exercise precise control over a conversational flow.

The current episode is also a factor in determining the target episode. For example, in an undergraduate student advising system, the question "'who is the instructor"' maps to different instructor names depending on the course (episode) under discussion. If a semantic element is not defined in the current episode, the applicable episode can be determined by exhaustive depth-first search similar to the class hierarchy used in frame based systems [14]. If the Context mediator requires additional information from the user, it generates a query. Response Generator, on receiving the query, will construct the appropriate natural language text and forward to the user.

Finally we designed our chatbot with the following steps:

- The user inputs the sentence into the Chat UI which then passes it along to the Chat Engine.
- The chat Engine traverses a graphical representation of AIML nodes in Memory.
- Given a match is found through a procedure which is depicted in [14], UFAQBot will return the best response following the codebase design. The system Engine is designed with the program D engine [5] [12] which functions as a jetty web servlet container. A sample of conversations with Program D Engine in shown in Figure 3.

Once the transition has been successfully completed, the Topic Finder collects all links in L originating from the current episode to its neighbors and sends those links to the Response Generator. The Response Generator would then generate appropriate response by combining template strings from different information nodes based on the links and the link preference value. The higher the link preference value, the more is the importance given to that link. Therefore, by assigning low link preference values to the links leading from any of the episodes in E to e_0 , the robot can be kept focused in the academic conversation. Similarly, high link preference value must be assigned to those links originating from e_0 and reaching any of the nodes in E, so that, robots will be able to refocus even when their conversion goes out of the Question Answering context.

IV. Experiments and Discussion

The crux of this paper is to analyze the dialogue design with conversation knowledge in a domain-specific chatterbot. With this purpose, this experimental design needs to constrain the domain of knowledge and range of capabilities that the robot should require. To do this, we created three chatterbot systems that differed only in their knowledge bases; BaseBot, UFaqBot (Dom_eng) domain engineering and UFaqBot (rep) representative. BaseBot, the conversational and control chatterbot, used the Standard AIML files which consist of 23,920 knowledge based entries and assumed to hold conversation dialogue AIML entries. UFaqBot (Dom_eng) contained 350 Student information desk bot entries (categories) that augment frequently asked questions on CSE admission systems and on CSE department and coined as Information Desk dialogue AIML. We made three files for it with Admission Information, Course Information and Faculty Information. Earlier in section 2, we engineered it with our system design. In addition, it contains the same conversational knowledge as BaseBot (i.e. 23,920).

UFAQBOT (rep) representative is composed of the same 350 categories of frequently asked questions that are engineered with an augmented approach of Semantic definition which is added with approximately 4298 conversation entries out of ALICE basebot. The BaseBot categories are here as to avoid a risk of not returning any responses to participants, at the same time we wanted to test if limiting the conversation entries in BaseBot of ALICE improves the dialogue design. A very subset of rudimentary 4298 categories were selected from BaseBot and manually modified with our design so that by modifying topic changing and personality related files such ”that.aiml, personality.aiml and proj.aiml” to the context of the domain, improves the accuracy of response relating to the domain of the FAQ knowledge Bot.

UFAQbot returns in our experimentation 65% of the queries given to it. From this position, we can test the effects of two different degrees of conversational knowledge coupled with full domain knowledge. For this purpose, A total of 31, 39 and 45 subjects (students) produced the number of interactions as shown in Table I. Moreover, it shows the breakdown of categories between each of the systems.

A. Performance Metrics

The implementation of this project in a university environment is indeed something very new. This FAQbot can be used to assist the undergraduate Advisor in giving advice to students. It is hoped that this Faqbot will lessen the current workload of the undergraduate Advisor. In this project, the Botmaster was the author of the survey and a Botmaster is the master of the chatbot. Our responsibilities included reading the dialogues, analyzing the responses, and creating new replies for the patterns. We intend the serve the UITS university undergraduate advisor with our Chatbot and implement it as a FAQBot that complements the Academic Registry of UITS University to advise undergraduate and graduate student. To measure the two chatterbox system’s performance, We conducted a survey among a group of final year students of information Technology course. We asked students to interact with the chatterbox UFAQBOT(rep) and UFAQBOT (Dom_eng) and told them to stay focused on the context, in our case Admission Info, Course Info, Faculty Info. They were told to ask equal number of questions on each subtopic and then marked each of the systems response as either satisfactory or unsatisfactory, and at the end of the survey, the sum of all satisfactory responses were taken into accounts with respect to Admission Info, Course Info and Faculty Info.We introduced two evaluation subjective parameters; Satisfactory and Unsatisfactory. Satisfactory responses were based on how appropriate the system responded in context to the question asked by the user; this value was subjective decision and determined by the user, depending on how accurately the system responded to the query.

Unsatisfactory responses were just the opposite of satisfactory and occurs if the response by the system was out of the context. Correspondingly,the user determined the response as inaccurate or irrelevant to the query, when they were asked to do so. separately. The result is shown in II and in III. To gauge the various system performances, we measured the following evaluation variables; Topic switching rate, Correction rate.

Topic switching rate (TSR) is defined by the number of times the dialogue is switched from on topic to off-topic divided by the number of conversations in a dialogue session with a user. Off-topic includes nonsense response and any response undesired to user. This is shown in Table IV.

TABLE I
DISTRIBUTION OF AIML CATEGORIES

Interactions for System Type	No of AIML Categories		
	Standard	UFAQBot	Total
3771 BaseBot	23,920	0	23,920
4552 UFAQBOT(Dom_Eng)	23,920	320	24,270
4005 UFAQBOT (rep)	4298	305	4698

TABLE II
STATISTICS OF DATA COLLECTED FROM THE SURVEY WHEN STUDENTS INTERACTED WITH THE UFAQBOT(REP) CHATTERBOX

Conversation Context	Satisfactory	Unsatisfactory
Admission info	70%	30%
Course info	80%	20%
Faculty info	60%	40%

TABLE III
STATISTICS OF DATA COLLECTED FROM THE SURVEY WHEN STUDENTS INTERACTED WITH THE UFAQBOT(DOM_ENG) CHATTERBOX.

Conversation Context	Satisfactory	Unsatisfactory
Admission info	60%	40%
Course info	60%	40%
Faculty info	50%	50%

TABLE IV
TOPIC SWITCHING RATE

System Name	No. of Conversation (session)	TSR
BaseBot	51	68.3%
UFAQBOT (Rep)	51	22.8%
UFAQBOT (Dom_Eng)	51	19.4%

TABLE V
CORRECTION RATE FOR DIALOGUES

System Name	Correction Rate		
	Conversational	Information	RS
BaseBot	17.4%	38.4%	20
UFAQBOT(Dom_Eng)	16.1%	22.8%	60
UFAQBOT (rep)	20.3%	21.1%	70

Correction rate is defined as a percent of system responses that were corrected by the user, divided by the total number of interactions typed into the system. Note that, the act of correction requires time and effort on the users part and thus the user may elect to bypass possible corrections he or she judges to be less pressing. Table V shows the result. Response Satisfaction (RS) is a subjective score ranked by users to determine the overall satisfaction after testing with users.

V. CONCLUSIONS AND FUTURE WORK

The Design of University Faqbot is aimed for implementing topic specific FaqBot so that it helps students as an academic advisor. In order to do this, the objective was in the experiment is to design the parameter which meets the hypothesis that domain knowledge is more effective in a ChatBot environment to obtain domain specific knowledge than conversation knowledge alone. We let the users to test the systems and we evaluated the dialogue history in a particular session and all sessions by number of interactions. It was found that in topic specific knowledge Bot yields the best result for topic switching ratio and correction rate over BaseBot which confirms our hypothesis. Moreover, UFAQbot(rep) shows the gain in satisfactory result over UFAQbot (Dom_eng), which proves that our architecture, if were applied to conversation dialog modification, would yield maximum dialogue session with user over a certain topic. Further, we found that domain-specific knowledge coupled with conversational knowledge based on our architecture design yields that best result of 19.4% and 21.1% over a Chatbot that have simple conversation dialogue and topic specific dialogue file. However, UFAQbot(rep) is made with only 4298 conversational dialogues with BASEbot and it's not tested whether increasing dialogues will decrease the performance. Another limitation is that the subjective parameter for response satisfaction does not reflect the true response satisfaction. The contribution of this system is that we addressed the problem of evaluating a low-level dialog system's ability to bestow domain knowledge in a very systematic way. In future, more statistical parameters on data mining and user profiling would be of use. Future ALICE research based on UFAQbot should focus on specific categories of knowledge that participants are most likely to correct as well as which knowledge categories have higher Response Satisfaction scores. Finally, measuring the quality of user-suggested knowledge would be worthwhile.

REFERENCES

[1] J. Chai and J.Lin, "The role of natural language conversational interface in online sales: a case study," *International Journal of Speech Technology.*, vol. 4, pp. 285295, Nov. 2001.

[2] J. Chai, Horvath V., Nicolov, N., Stys, N., K., Zadrozny, W., Melville, P.,"Natural Lanuage Assistant: A dialogue system for online product recommendation," *AI Magazine*; *ProQuest Science Jounals.J* Summer 2002; 23, 2, pp 63-75

[3] J. Jia, "The Study of the Application of a keywords-based Chatbot System on the Teaching of Foreign Languages." [4]

[4] R. Moore, G. Gibbs, "Emile:Using a Chatbot Conversation to enhance the learning of Social Theory," Univ. of Huddersfield, Huddersfield, England, (2002)Ausburg, Germany, University of Ausburg. 2002.

[5] Wallace R.A.L.I.C.E. The Artificial Linguistic Internet Computer Entity.Pandorabots. [Online]. Available: [Online].Available:http://www.pandorabots.com/pandora/

[6] Wallace, R., "The Elements of AIML style" A.L.I.C.E. Artificial Intelligence Foundation, Inc.

[7] Eliza- a friend you could never have before. [Online]. Available: ¡http://www-ai.ijs.si/eliza/eliza.html¿

[8] Turing A.M.: Computing machinery and intelligence. Mind 59, 433-460 (1950).

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

[10] Loebner H. (2006). Hugh Loebner's Home Page [Online]. Available:http://www.loebner.net

[11] The Chatterbox Challege. http://www.chatterbocxchallenge.com

[12] AIML Foundation:[online]. Available: http://www.alicebot.org

[13] A. S. Bayan, B.A., Atwell, E. and Roberts, A., "FAQCHAT as in information retrieval system," proceedings of the 2nd Language and Technology Conference, Poznan, Poland ,April, 21.-23, 2005.

[14] A. S. Bayan. A. Eric, "A comparison between Alice and Elizabeth Chatbot Systems," research report series, university of Leeds, December, 2002.

[15] R. P. Schumaker, Ginsburg, M., Chen H., Liu Y.,"An evaluation of the chat and knowledge delivery components of a low-level dialog system: The AZ-ALICE experiment," *Decision Support Systems*, 42(4):2236-2246 (2007).

[16] Nardi, B., Editor., "Context and Consciousness: Activity Theory and Human-Computer Interaction," Cambridge: MIT Press, (1996).

[17] R. P. Schumaker, Liu Y., Ginsburg, M., Chen H., " Evaluating Mass Knowledge acquisition in the Alice Chatterbot: The AZ-ALICE dialog System." *Int J. of Human Computer Studies*, 64: 1132-1140 (2006)

[18] Molla, D., vicedeo J., "Question answering in restricted domain: an overview," *Computational linguistic*, 33(1): 41-61.