

ARTIFICIALLY INTELLIGENT VIRTUAL ASSISTANT
P.DhivyaBharathi (Assistant Professor/CSE)¹ D.J.Srihema² R.Monisha³
V.Divya⁴ B.Veenitha⁵
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
Er.PerumalManimekalai college of Engineering, Hosur-635117.

Abstract :Systems today are getting expert day by day and intend to help human in their day to day queries. Today AI is present in a variety of fields ranging from industries in manufacturing, to diagnosis in medicine technology, to customer care in public relations. There exist lots of online Artificial Intelligence (AI) assistants that help people solve their problems. So, here in the implemented system we built an AI that will solve college related query. It's like a small scale college oriented intelligent search engine. The implemented system is basically a Virtual Assistant that is strictly college oriented. The implemented system entertains the queries of a student regarding college related issues. The implemented system was constructed Using Visual Basic .Net.

Keywords - *Virtual Assistant, Bot, Chat, College, Assistant, Academic, AIML*

1. Introduction

It is very easy to interact with system even if user doesn't know how to read and write so, it becomes incentive for user to use this system by using natural language. The model receives response from user either in form of text or in form of speech. Then it

The implemented system can assist

responds to the user in both text as well as sound format. One of the most important goals is to enhance the interaction realness of these systems. The concept of "Alice" was the first chatterbot which receives question from user based on pattern recognition. Large amount of sentences were involved to train the model. catchy interfaces using human-like avatars are provided for this reason which are capable to adapt their behavior according to the conversation content. This kind of agents interact vocally with users by the use of Automatic Speech Recognition (ASR) and Text To Speech (TTS) systems. Their "emotions" change according to the sentences entered by the user. Facial movements are simulated in the 3D talking head by rational free form. As the name of implemented system suggests, the vision was to develop a virtual assistant for college students to solve college based queries and provide college related information. The term 'Virtual' may be defined as "Not physically existing as such but made by software to appear to do so." Thus, the implemented system is a virtual assistant, i.e., it is a software that uses artificial intelligence to guide a student and takes actions to effectively understand student queries and respond to them rationally.

the user by solving queries related to college.

2. EHeBby architecture

The system is composed of two major factors, as shown in figure 1, a reasoner module and a Talking Head (TH) module. The reasoner processes the user question by means of the A.L.I.C.E. (Artificial Linguistic Internet Computer Entity) engine. ALICE was implemented with pattern matching algorithm which is as simple as string matching technique. ALICE takes the input in the text form and produces output in text form like question and answer based system. Whereas another chatter bot which was build earlier known as Elizabeth requires set of input rules, keyword patterns for identifying input and output rules to produce required response is not used in this system since it is complicated. In order to produce the correct information needed for the talking head animation the emotional area allows the chatbot to elaborate information related to the produced answer and a correspondent humor level. The TH system relies on a web application where a servlet selects the basis facial meshes to be

3. Proposed Work

The architecture of bot system is divided into following modules:

i) AIML

i). AIML

Artificial intelligence markup language is an extension of XML (extensible markup language). The most important parts of an AIML documents are:

a) <aiml>

b) <category>

c) <pattern>

animated and integrated to process emotion information. On the client side, all these data are used to actually animate the head. Two answers can be combined by ALICE in the case of splitting during the normalization process and by recursive process. Pattern matching process is the general feature of ALICE chatter bot, which is simple and depth-first search algorithm is used where it results in producing no output also.

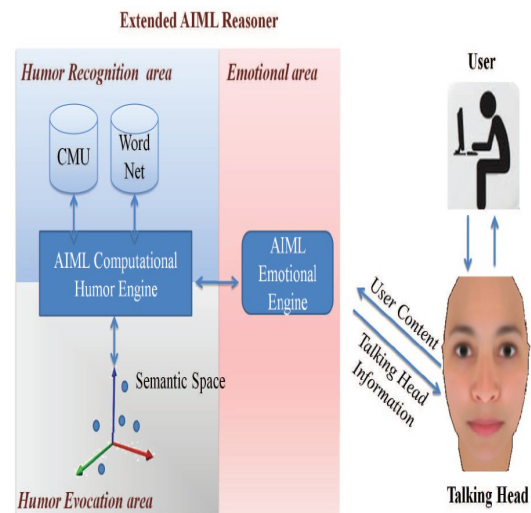


Fig. 1. EHeBby Architecture

ii) Microsoft Speech recognition

iii) Corpus

iv) Brain

v) Bot engine

d) <template>

AIML consists of aiml objects and these aiml objects consist of topics and categories which contain either parsed or unparsed data from their information is extracted.

3.1 AIML KB

AIML categories:

1. To manage a general conversation

2. To generate humorous sentences by means of jokes.
3. To retrieve humorous or funny sentences through the comparison between the user input
4. To recognize an humorous intent in the user sentences.

3.2 Humour recognition area

The humour recognition consists of identification, inside the user sentences, of particular humorous text features. We focus on three main humorous features: alliteration, antinomy and adult slang.

3.2.1 Alliteration-This technique is aimed at discovering possible repetitions of the beginning phonemes in subsequent words.

EXAMPLE: *Veni, Vidi, Visa: I came, I saw, I did a little shopping Infants don't enjoy infancy like adults do adultery*

3.2.2 Antinomy- This module detects the presence of antinomies in a sentence which has been developed exploiting the lexical dictionary WordNet.

EXAMPLE: *A clean desk is a sign of a cluttered desk drawer Artificial intelligence usually beats real stupidity*

A huge amount of information or corpus is required such that an accurate NLP engine can be build. After obtaining the corpus, sentence chunking is done so that each sentence can be converted to speech or voice. STT is speech to text, where sound or voice which is received core engine process these sentences depending upon the situation and it is moved forward to speech synthesizer to respond. Processed separately. They are further divided into named entity recognition

3.2.3 Adult slang recognition module

This module analyzes the presence of adult slang searching in a set of pre-classified words.

EXAMPLE: *The sex was so good that even the neighbors had a cigarette Artificial Insemination: procreation without recreation*

ii). Natural language Processing

Before speech recognition can work, there arises need of natural language processing. One need to understand how language should be interpreted as there is ambiguity in processing any language as shown in fig.2

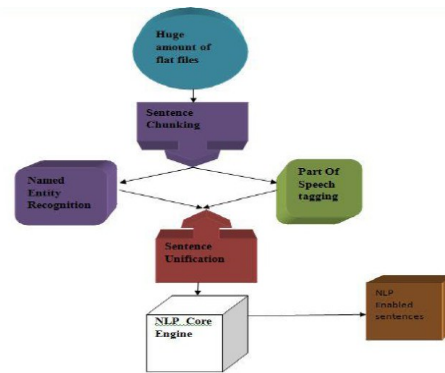


Figure 2: Processing of NLP

iii). Microsoft Speech Recognition

This speech recognition consists of TTS and STT. TTS is text to speech system where text is converted to speech or voice. STT is speech to text, where sound or voice which is received from the user is converted to text for processing of information by bot. Initially, sentence given or received is passed to NLP and after that, NLP core engine process these sentences depending upon the situation and it is moved forward to speech synthesizer to respond.

i). Corpus -To respond accurate to question or in any situation given to the bot.

ii). Brain -The brain file size is around ninety mega bytes which is compressed form of corpus

iii). Bot Engine –consists of following components

- a) Kernel b) AIML Parser c) Utility Manager
- d) Word Substitution e) Pattern Manager

3.3 Humor evocation area

The humor evocation area computes the semantic similarity between the sentence said by the user and the encoded sentence in semantic space; also it tries to answer to the user with a funny expression which is conceptually close to the user input. This procedure allows to go beyond the firm pattern-matching rules, generating the funniest answers which semantically fit the user query.

3.3.1 Semantic space creation

A semantic representation of funny sentences has been obtained mapping them in a semantic space. The semantic space has

3.4 Emotional area

The generation of emotional expressions in the Talking Head suits emotional area. There are many possible models of emotions in literature. They can be distinguished in three different categories of models. The first one refers the models describing emotions through collections of different dimensions (arousal, , unpredictability, valence, potency, intensity ...). The second one refers models based on the hypothesis that a human being is able to express only a limited set of primary emotions. All the range of the human emotions should be the result of the

been built according to a Latent Semantic Analysis (LSA) based approach described in Agostaro (2005)Agostaro (2006). we use the following similarity measure Agostaro (2006):

$$\text{sim}(v_i, v_j) = \begin{cases} \cos^2(v_i, v_j) & \text{if } \cos(v_i, v_j) \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The closer this value is to 1, the higher is the similarity grade. The geometric similarity measure between two items establishes a semantic relation between them. In particular given a vector **s**, associated to a user sentence **s**, the set CR(**s**) of vectors sub-symbolically conceptually related to the sentence **s** is given by the **q** vectors of the space whose similarity measure with respect to **s** is higher than an experimentally fixed threshold **T**.

$$CR(s) = \{v_i | \text{sim}(s, v_i) > T \text{ with } i = 1 \dots q\} \quad (2)$$

Thachatterbot can also develop its own AIML KB mapping in the suggestive area new items like jokes, riddles and so on introduced by the user during the conversation.

combination of the primary ones. The last category includes mixed models, according to which an emotion is generated by a mixture of basic emotions parametrized by a set of dimensions.

4. EHeBby talking head

Our talking head can speak several languages since it is conceived to be a multi-platform system. so that various implementations have been realized. The different components of our model are presented: model generation, animation technique, coarticulation, and emotion management.

The face model generation involves a special tool called FaceGen which is mainly used for the creation of 3D human heads and characters as polygon meshes. Numerical parameters are used to control the facial expressions. A variety of facial expressions can be obtained since the intensity of each emotion can be controlled by a parameter or mixed to each other.

Morphing is done for facial expressions to express the animation. It starts from geometrics called keyframes. The keyframes vertex translates from its position to occupy one of the corresponding vertex in the subsequent keyframe. A set of visemes are generated for this reason instead of modifying a single head geometric model. This approach is less efficient than animation engine in case of modifying the facial expressions like tongue position, labial protrusion and so on.

4.1 Cohen-Massaro model

$$S = \{f_1 \in [0, t_1]; f_2 \in [t_1, t_2]; \dots; f_n \in [t_{n-1}, t_n]\} \quad (4)$$

A viseme is described as “active” when ‘t’ falls into the corresponding time interval. The preceding and the following visemes are described as “adjacent visemes”. The weights are computed as follows:

$$w_i(t) = D_i(t) = a_i \cdot \exp(-\theta_i \cdot |t - \tau_i|) \quad (5)$$

where τ_i the mid point of the i-th time interval. The w_i must be normalized:

$$w'_i(t) = \frac{w_i(t)}{\sum_{j=-1}^{+1} w_{i-j}(t)} \quad (6)$$

This is a mathematical model used for coarticulation. This computes the weights to control the keyframe animation. the vertex positions of an intermediate mesh between two keyframes are determined by these weights. Dominance is the name given to the influence on the organs of articulation of the face rather than perpendicular segments and can be mathematically defined as a time dependent function.

$$D_{sp} = a \cdot \exp(-\theta |\tau|^c) \quad (3)$$

where a is the peak for $t = 0$, q and c control the function slope and t is the time variable referred to the mid point of the speech segment duration. The coarticulation can be thought as composed by two sub-phenomenons: the pre- and post-articulation. The prosodic sequence S of time intervals $[t_{i-1}, t_i[$ associated to each phoneme can be expressed as follows:

The coordinates of the interpolating viseme

vertexes $v(l)_{int}(t) \in \{V_{int}(t)\}$ will be

computed for each time instant as follows:

$$v_{int}^{(l)}(t) = \sum_{k=i-1}^{i+1} w'_k(t) \cdot v_k^{(l)}(t) \quad (7)$$

the index l indicates corresponding vertices in all the involved keyframes. Also this computation is simplified by our implementation.

4.1.1 Dominant visemes and Diphthongs

A sequence of two adjacent vowels is called diphthong. The vowels in a diphthong must be visually distinct as two separate entities. The visemes belonging to the vowels in a diphthong must not influence each other. Otherwise, both the vowel visemes would not be distinguishable due to their fusion. The dominance of a vowel with respect to a consonant is accomplished with a less steep curve than the consonant one as shown in the figure.

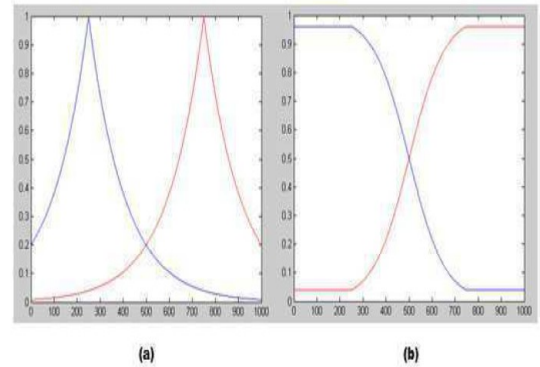
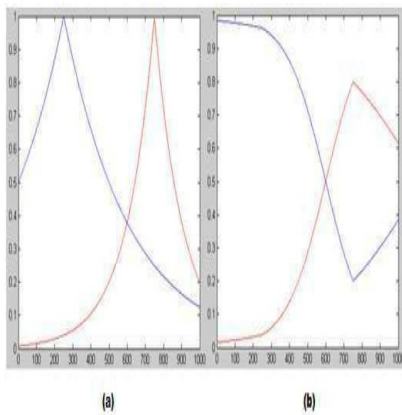


Fig 3- The dominance function for the diphthong case (a) and the weights diagram (b) for the diphthong case
Fig. 4. The same of Fig.2 for the vowel-consonant case

4.2 Audio stream synchronization

Facial movements are driven by the .pho file since the Talking Head is essentially synchronized with the audio streaming which determines the phoneme (viseme) and its duration. a variety of options to produce the prosody for the language and speech synthesizer to use is provided by Espeak.

5. Some examples of interaction

5.1 Example of humorous sentences generation

The following is an example of an humorous dialogue

User: What do you think about robots?
 EHeBby: Robots will be able to buy happiness, but in condensed chip form!!

```
<category>
<pattern>WHAT DO YOU THINK ABOUT ROBOTS</pattern>
<template>Robots will be able to buy happiness, but in condensed chip form!!
</template>
```

</category>

The pattern delimits what the user can say. Every time the pattern is matched, the corresponding template is activated.

5.2 Example of humor recognition

The recognition of humorous sentences is obtained using specific tag inserted into the template, as shown in the following categories

```
<category>
<pattern>CAN I TELL YOU A JOKE</pattern>
<template>Yes you can</template>
</category>
<category>
<pattern>*</pattern>
```

```
<that>YES YOU CAN</that>
<template>
<srai><humorlevel><star/></humorlevel>
><srai>
</template>
</category>
```

If the previous answer of the chatterbot was “Yes you can” then the second category is activated. The humor level can assume three different values, low, medium and high. The following example shows the category related to a high humor levels

```
<category>
<pattern>HIGH *</pattern>
<template><think>
<prosody>
```

5.4 AIML categories for targeting

The humorous chatterbot will update its own sub-symbolic data through a targeting procedure, which will map new acquired riddles in the semantic space. By means of the ad-hoc created AIML tag `addRiddle`, the targeting is obtained as shown in the following chunk of AIML code:

```
<category>
<pattern>Listen this joke *</pattern>
<template><humorlevel><star/></humorlevel>
<think>
<addRiddle><star/></addRiddle>
</think>
</template>
</category>
```

A sentence introduced by the user as a vector in the semantic space is effected to code by the `addRiddle` tag has by means of the folding-in method. Think tag hides the entire procedure to the user. In this manner, the user will see only the chatterbot’s result to his joke..

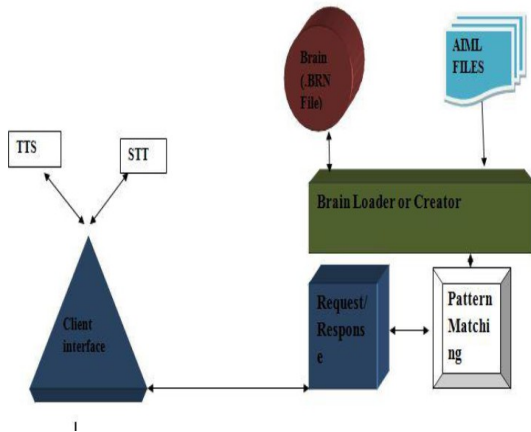
```
<star/>
</prosody>
</think>
<joy intensity="080" />
</template>
</category>
```



Fig. 5. TH reaction to a funny joke

Procedure

- 1.) A brain file is created initially, by using aiml file.
- 2.) If brain file is not present then create a brain.
- 3.) If brain file is present then load the brain file in the model.
- 4.) when the brain is loaded, bot waits for the request from user or client.
- 5.) The request is provided by the user either in the form of voice or text.
- 6.) Then the model receives request and forwards to pattern manager
- 7.) Pattern matching algorithm is applied and sent to the brain.
- 8.) The model uses the brain and get suitable response and forwards it to the user



Flow Diagram of Request and Response

6. Results

Two appropriate data sets have been created in order to validate the humor recognition. The former, called DataSet1, is composed of 100 humorous phrases extracted by ad-hoc selected web sites, characterized by the presence of a percentual of humoristic features, as shown in table 1.

Features	Percentual in DataSet1
Alliteration	45%
Direct Antinomy	10%
Extended Antinomy	6%
Indirect Antinomy	5%
Adult Slang	6%
No Feature	28%

Table 1:Humoristic Features Distribution in DataSet1

The latter, called DataSet2, is composed of 200 phrases, equally distributed between humorous and non-humorous sentences, where the non-humorous examples are always extracted

from Internet and chosen among titles of sets definitions ,proverbs and newspapers. The implemented algorithms are well performed in humor reognition, as shown in table 2.

DataSet	Correct Classification	No Recognition
DataSet1	66%	34%
DataSet2	73%	27%

Table 2: Humor Recognition in the analyzed DataSets

The results will be considered satisfactory. Moreover if the humor recognition area cannot identify the humorous level of sentences given by the user, the dialogue can continue in a funny way. In fact, the chatterbot exploits the semantic space, which allows retrieving entertaining sentences so that the dialogue can continue in a funny approach.

7. Conclusion

Chatbot is proved to be a boon in any industry where one can look for assistance

without need of human. Bot itself acts as human mind to provide correct information. It can be better as a teacher in education. It can provide appropriate details of medicines to be consumed for particular disease in healthcare industries and for traveler it can be a guide assistance to provide him shortest path for travelling particular with minimum expenses. The whole architecture relies on a suitable AIML-based chatterbot, and an

Reference:

1. Mihalcea R. and C. Strapparava. (2006) Learning to laugh (automatically): Computational Models for Humor Recognition. *Computer Intelligence*, Volume 22, 2006
- MultiWordNet (2010): <http://multiwordnet.itc.it/english/home.php> <http://www.oneliners-and-proverbs.com/> and <http://www.bdwebguide.com/jokes/1linejokes-1.htm>.
2. Waters K, Levergood T M. An automatic lip-synchronization algorithm for synthetic faces. s.l. : MULTIMEDIA '94 Proceedings of the second ACM international conference on Multimedia ISBN:0-89791-686-7
3. Weizenbaum, J. ELIZA – A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 10(8):36–45 (1966).
4. Cliff, C. and Atwell, A. Leeds Unix knowledge expert: a domain-dependent expert system generated with domain-independent tools. BCS-SGES: British Computer Society Specialist Group on Expert Systems journal, 19:49–51 (1987).
5. Wallace, R., Tomabachi, H., and Aimles, D. Chatterbots go native: Considerations for

animation engine for the talking head. The chatterbot reasoner module is based on an extended AIML architecture dealt with using suitable tags. The whole system has been tested on the humor recognition task with satisfactory results. However, our system is currently under development and much work has to be done in order to improve the whole architecture.

an eco-system fostering the development of artificial life forms in a human world. Published (2003), online: <http://www.pandorabot.com/pandora/pics/chatterbotsgonative.doc>.

6. Bran, E. Chabot's in der Kundenkommunikation (Chabot's in customer communication). Springer, Berlin (2003).

7. Abu Shower, B. and Atwell, A Using the corpus of spoken Afrikaans to generate an Afrikaans Chabot SALALS Journal: Southern African Linguistics and Applied Language Studies, 21:283–294 (2003).

8. Abu Shower B. and Atwell A. Different measurement metrics to evaluate a Chabot system Proceedings of the NAACL'07 Workshop: Bridging the Gap: Academic and Industrial Research in Dialog Technologies. Pp.89-96, ACL(2007).

9. Abu Shower B. and Atwal A. Machine Learning from dialogue corpora to generate chatbots. *Expert Update journal*, Vol. 6, No 3, pp 25-30 (2003).

10. Steven Birds, Ewan Klein and Edward Lapper. *Natural Language Processing with Python* by O'Reilly Media, Inc. (2009).

11. Kalra P, Mangili A., Magnetat-Thalman N, Thalman D. Simulation of facial muscle actions

based on rational free form deformations.
SCA '06 Proceedings of the 2006 ACM
SIGGRAPH Eurographics symposium on
Computer animation ISBN:3-905673-34-7

12.Löfqvist, A. (1990) Speech as audible gestures. In W.J. Hardcastle and A. Marchal (Eds.) *Speech Production and Speech Modeling*. Dordrecht: Kluwer Academic Publishers, 289-322.

13.Lee Y, Terzopoulos D, Waters K. Realistic modeling for facial animation. *Proc. ACM*

SIGGRAPH'95 Conference, Los Angeles, CA, August, 1995, in *Computer Graphics Proceedings, Annual Conference Series*, 1995, 55-62.

14.Liu K, Ostermann J. Realistic Talking Head for Human-Car-Entertainment Services. IMA 2008 Informationssysteme für mobile Anwendungen, GZVB e.V. (Hrsg.), pp. 108-118, Braunschweig, Germany

15. Jonas Sjöbergh and Kenji Araki. A Very Modular Humor Enabled Chat-Bot for Japanese. *Pacling 2009*