Knowledge Upload Service Using Semantic Based Categorization

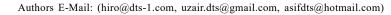**HIRONAO TAKAHASHI*, UZAIR LAKHANI**, AND ASIF RAZA***

ABSTRACT

AI (Artificial Intelligence) is one of the hot keywords in Computer Science research and development these days. Applications of AI are very diverse that include but not limited to image recognition, voice recognition, industry manufacturing, auto driving and human interaction for any service industry. Here we propose a knowledge upload system utilizing different AI techniques. The knowledge recording part of the system works in this way. First the human voice is captured using microphone. Later on the captured voice is converted into text using voice recognition techniques. Finally, the recognized text is categorized into personal or general category using semantic techniques. The knowledge retrieval part of the system on the other hand helps in retrieving the thoughts of a person whose knowledge has been uploaded regarding some particular subject. A knowledge upload service can be beneficial for each generation as the people of that generation can benefit from the thoughts of people of earlier generation. In this way precious human knowledge can be saved for the benefit of mankind. Normally AI is general which is not specific to a person but the system proposed here can be termed as a step towards PAI (Personal Artificial Intelligence).

Key Words: Knowledge Upload, Semantic Categorization, Speech Recognition, Natural Language Processing, Artificial Intelligence.

1.   INTRODUCTION

The mind of a person is one of the most important parts of his body. It is human mind that has resulted in the advancement of technology for the betterment of the human society therefore its knowledge should be preserved. The knowledge of a human mind is its precious asset. AI tries to mimic the behavior of natural intelligence of human mind. Although there is progress in the field of AI but still there is a long road ahead to cover. The importance of AI can be estimated from the fact that almost all major technology companies are working on various AI technologies these days.

In this paper we propose use different AI techniques like Speech Recognition [1] and NLP (Natural Language Processing) [2] in order to upload knowledge of different people so that later on family, friends and general public can query from the uploaded knowledge and get their answers. Additionally, the paper has proposed using semantic techniques for text categorization purpose. Text categorization can be performed by two approaches. One approach is that of machine learning based which need creation of a training dataset. On the other hand, the approach utilized in this paper is semantic based that

Authors E-Mail: (hiro@dts-1.com, uzair.dts@gmail.com, asifdts@hotmail.com)
*   Department of Computer Science, Preston University, Karachi, Pakistan.
**   Data Transmission System Inc., Karachi, Pakistan.

does not require the training dataset. The creation of the training dataset can be sometimes difficult and time consuming. Further it may require paid services of a data operator to accomplish the task. The semantic approach helps us in avoiding this dataset.

Although there are multiple ways to transfer knowledge e.g. utilizing physical actions and other alternate methods but our focus in this paper is on transferring knowledge using the voice channel only. One of the main reasons for this limitation is that the technology stack currently is not mature enough to handle other channels of knowledge. Even the voice recognition techniques are not finalized and improvements are continuously being made in them.

A knowledge upload service can be beneficial for each generation as the people of that generation can benefit from the thoughts of people of earlier generation. People talk in different languages therefore a single system cannot cover all natural languages. As English is the most widely used language we use it as our primary language. For every stage like voice recognition, NLP, knowledgebase etc. we can get good tools and technologies for English language in comparison to other languages.

This paper provides a brief overview of the proposed system architecture. It mentions different tools that can be utilized in Java in order to achieve the proposed goal. Java is a platform independent programming language that is the same source code can run on different platforms like Microsoft Windows, Apple Mac OS X, and Linux etc. The best case for the system is achieved when it passes the Turing Test while answering the queries.

It should be noted that our approach for preserving human knowledge is different from the approach that utilizes using complex machinery directly connected with the mind in order to transfer the knowledge. The negative effects of such an approach on human health cannot be neglected.

The rest of the paper is organized in this manner. Section 2 discusses the related work while the system architecture is discussed in section 3. Methodology covering the algorithms is discussed in section 4 while section 5 covers the evaluation part. Finally, the paper is concluded in section 6.

2.   RELATED WORK

The subject of knowledge upload is very close to the mind upload concept. Research work has been done related to mind upload e.g. in [3] the author has discussed the advantages that can be achieved through mind upload. The concept of the Digital Mind in the paper covers both Human Mind as well as Artificial General Intelligence.

Similarly, the authors in [4] propose connecting two uploaded minds in order to get more advantages. The authors have discussed the benefits of this approach as well as the barriers in achieving it.

A much cited paper in the mind upload domain is [16] which provides the philosophical analysis of the mind uploading concept. This paper is an excerpt from "The Singularity: A Philosophical Analysis", Published in The Journal of Consciousness Studies. In the paper along with other things the author has raised various questions related to the uploading concept like consciousness, personal identification etc. That is whether after uploading consciousness and personal identification can be achieved.
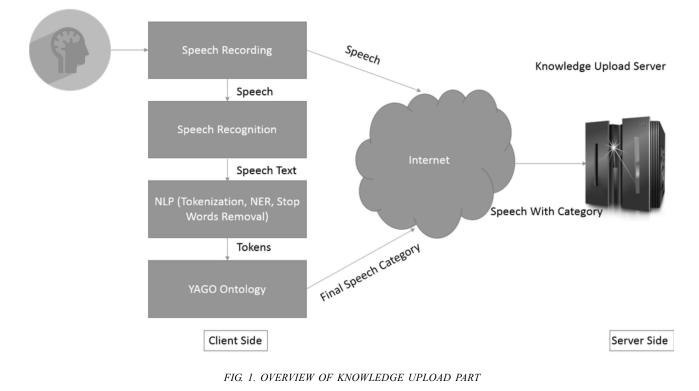
Sandberg and Bostrom [15] have prepared a report that discusses the roadmap for Whole Brain Emulation. Initially the authors have provided a reasoning why this work should be done. The authors have discussed topics like neural simulation, body simulation, and environment simulation etc. The sub parts of environment include vision, hearing, smell and taste etc. Our approach for knowledge upload differs from them as we are utilizing the voice recognition and NLP technologies which are not utilized by the reported authors.

In fact, we could not find detailed research work where the technologies mentioned in the paper are utilized for the purpose of knowledge upload therefore we can say that our approach regarding knowledge upload is different from the approaches utilized by the research community.

3.   SYSTEM ARCHITECTURE

Our proposed system architecture consists of two parts. The first part is termed as 'Knowledge Recording' while the second part can be termed as 'Knowledge Retrieval'. The first part will be utilized by the person whose knowledge is to be uploaded while the second part will be used by persons who want to utilize the already uploaded knowledge by asking different questions in order to get good answers to their questions.Figs. 1-2 provide an overview of the proposed system for part one and part two respectively.

In part one or knowledge recording module different speeches of the person whose knowledge is to be uploaded will be recorded on different occasions. Fig. 3 in a screenshot of the knowledge upload demo application. For each recorded speech the next step is the conversion of speech into text using speech recognition library. There

FIG. 1. OVERVIEW OF KNOWLEDGE UPLOAD PART

are multiple options for performing speech to text conversion. If we want to utilize some web service, then Google Speech API [5] can perform speech to text conversion with good accuracy in Java. Other options include iSpeech [6], IBM Watson Platform [7], HPE Speech Recognition [8] etc. On the other hand, if we have enough

time then we can utilize CMU Sphinx [9] which is an open source library for speech recognition in Java. For good accuracy we will need to train a model for CMU Sphinx which will require some time. Gaida et. al. [10] have compared different open source speech recognition toolkits.
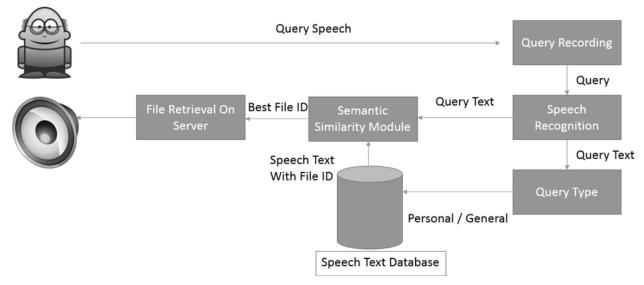

FIG. 2. OVERVIEW OF KNOWLEDGE RETRIEVAL PART

FIG. 3. KNOWLEDGE UPLOAD DEMO APPLICATION

Once the text for the speech is obtained then we can start semantic categorization of the text in order to determine the speech category. The first important step in the text categorization part is of NLP. The important parts of NLP include tokenization of text, NER (Named Entities Recognition) and stop words removal. We need to tokenize the text as we have to perform semantic enrichment that would further help in performing milestone based categorization. Three class NER involves determining the names of persons, organizations and places. Finally, we need to remove those tokens that have no effect on the category of the text. There are different libraries available for NLP in Java like Stanford NLP [11], GATE (General Architecture for Text Engineering) [12] etc.

For each token we make query against YAGO2s [13] knowledgebase which is an upper level ontology. Through this query we can determine the category of the token. At this point we will have category for each token of the text. Finally, the category which has the maximum count can be considered as the final category for the speech text. The final category can be further classified into 'Personal' or 'General' categories as an individual speech can be at a broader level either personal like on some family related topic or general like political discussion. Next the categorized speech of the person will be uploaded to some server machine using the Internet. This completes part one of the proposed architecture.

In part two or knowledge retrieval module different people like family members, friends or general public would ask different questions from the uploaded knowledge of a person in his absence in order to get this person's opinion regarding these questions. In order to achieve this the questioner will record his query first. Later on using the procedure already discussed this query will be converted into text. Next we will determine whether the query is personal or general purpose. Suppose the query is general purpose then we will retrieve every general purpose speech text of the person whose knowledge is already uploaded and then apply semantic similarity on the speech text and the questioner query text. The speech text for which we have the maximum similarity score will be returned as the answer. Similar pattern will follow for personal type query. This completes part two of the proposed architecture. Fig. 4 in a screenshot of the knowledge retrieval demo application.
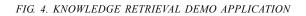
4. METHODOLOGY

We implemented our complete system in Java programming language which is one of the platform independent programming languages. First we name different packages along with the technologies that we utilized in our system (Table 1).

In order to simulate a testing environment, we made a collection of text for 13 different categories. The purpose of this text collection is to convert it into voice using the text to voice functionality on the Internet and to test the semantic categorization module of the system. Table 2 lists the categories names along with the internal numbering assigned to that category. We propose 80 plus categories for the semantic categorization system but Table 2 lists only some of the categories that are more relevant with the proposed system.

Next we list the main Algorithms only. Similarity can be of two types lexical or semantical in the broader sense. According to the lexical similarity "dog" can be similar to "dog" only. On the other hand, under semantic similarity a "dog" can be similar to a "cat" as both are of type animal. Therefore, in some cases we can get better results using semantic similarity approach described in Algorithm-I.

FIG. 4. KNOWLEDGE RETRIEVAL DEMO APPLICATION

TABLE 1. PACKAGE NAMES WITH TECHNOLOGIES UTILIZED

Package Name	Technologies
Categorizer	Stanford NLP, Java Utility, Java IO, Rita WordNet, Apache Jena, JSON
Database	Java SQL
File Upload	Java IO, Java Net
GUI Design	Java AWT, Java Swing
Sound Recording	Java IO, Java Sound
Speech to Text	HPE Speech Recognition
Audio Player	Java Sound, Java IO
File Retrieval	Java SQL, Java Utility
Semantic Similarity	Java IO, Apache Lucene, Stanford NLP, Java Utility, Apache Jena, Jersey, JSON, jsoup, JAWS

TABLE 2. CATEGORY NAMES

Category Number	Category Name
001	Market Rates
002	Insurance
003	Online Trading
004	Financial Product
005	Religion
006	Sports
008	Travel
014	Health
016	Health Animal
017	Employment
049	Illegal Drugs
082	Bank

ALGORITHM-I: SEMANTIC SIMILARITY

1: **procedure** SEMANTIC_SIMILARITY (sentence1, sentence2)

2: SemanticDataList1 ← GetSemanticDataList(sentence1)

3: SemanticDataList2 ← GetSemanticDataList(sentence2)

4: **if**(setence1.length > sentence2.length)

5:   totalFrequency ← SemanticDataList1.size()*4

6: **else**

7:   totalFrequency ← SemanticDataList2.size()*4

8: **endif**

9: **for** each Data1 in SemanticDataList1

10:   **for** each Data2 in SemanticDataList2

11:    **if**mapping(Data1.label, Data2.label) > 0

12:     obtainFrequency ← obtainFrequency + 4

13:     break

14:    **else if** mapping (Data1.type, Data2.type) > 0

15:     obtainFrequency ← obtainFrequency + 2

16:     break

17:    **else if**mapping(Data1.domain, Data2.domain) > 0

18:     obtainFrequency ← obtainFrequency + 2

19:     break

20:    **end if**

21:   **end for**

22: **end for**

23: **end procedure**

Algorithm-II retrieves the semantic data for a sentence. The first step for this is to determine the nouns in the sentence. Later on for each noun we get semantic data from the domain specific knowledgebase. The functionality of Algorithm-II is utilized for semantic similarity.

ALGORITHM-II: GET SEMANTIC DATA LIST

1: **procedure**GetSemanticDataList(sentence)

2: nounList ← GetNounList(sentence)

3: **for** each noun in nounList

4:   semanticData ← getDomainSpecificTriples(noun)

5:   semanticDataList.add(semanticData)

6: **end for**

7: **end procedure**

The mapping function in **Algorithm-III** first tests whether any token is null. In such a case the mapping is zero. On the other hand, if both tokens are lexically same then in that case the mapping is 1. Similarly, if the semantic data retrieved from Wordnet for each token is lexically same then in that case also the mapping is 1. The functionality of Algorithm-III is utilized for semantic similarity.

ALGORITHM-III: TOKENS MAPPING FUNCTION

1: **procedure** MAPPING (token1, token2)

2: **if** token1.length = 0

3:   return 0

4: **else if** token2.length = 0

5:   return 0

6: **end if**

7: **if** token1 = token2

8:   return 1

9: **end if**

10: token1SenseList ← GetWordnetSenses(token1)

11: token2SenseList ← GetWordnetSenses(token2)

12: **for** each token1Sense in token1SenseList

13:   **for** each token2Sense in token2SenseList

14:    **if**(token1Sense = token2Sense)

15:     return 1

16:    **end if**

17:   **end for**

18: **end for**

19: **end procedure**

5.   EVALUATION

In this section we will try to evaluate the two modules described earlier. First we tested the semantic categorization part of the knowledge recording module.

We utilized YAGO2s as a semantic knowledgebase for enrichment of tokens. YAGO2s is available for download in TTL (turtle) format. We need to convert the knowledgebase from TTL to TDB (Tuples Database) format.

Next using the text to speech functionality available on the Internet [14] we made a voice file against each category's text we collected earlier as described in the methodology section. Next we used these voice files for performing speech recognition and ultimately to categorize the voice file based on the text retrieved. We perform these two steps in order to simulate the knowledge recording part of the module. Below are the results for the semantic categorization part of the first module. We consider the result of the categorization engine correct if the correct category is returned in the top two category results returned by the categorization engine.

From Table 3 we can observe that the categorization part of the knowledge recording module can achieve 84.6% accuracy. Some of the important points to consider were the following:

- Input voice should be in good quality so as to properly perform voice recognition and ultimately to categorize it

- The background noise should be at minimum level

- In addition to YAGO2s we also utilized domain specific knowledgebase for improving categorization accuracy

Next we tested the Semantic Similarity part of the Knowledge Retrieval module. We tried to estimate the similarity score between query text and speech text.

The results of the evaluation for semantic similarity are shown in Table 4. From Table 4 we can observe that the semantic similarity part of the Knowledge Retrieval module has an accuracy of about 46.2%. We need to revise the semantic similarity algorithm for improving the accuracy. This can be one of the important future tasks. Following can be some of the important points to consider:

- The question is normally short and if there is slight mistake in recording or performing voice recognition for the query then the similarity score will be low therefore this should be considered

- The algorithm for the mapping function can also be modified. One such possibility can be to utilize the senses from sources other than Wordnet also and test the semantic similarity accuracy

TABLE 3. SEMANTIC CATEGORIZATION RESULTS WITH HPE SPEECH RECOGNITION (KNOWLEDGE RECORDING MODULE TESTING SIMULATION)

Voice Files Utilized	Correct Categorizations	Incorrect Categorizations
13	11	2

TABLE 4. SEMANTIC SIMILARITY EVALUATION RESULTS (PART OF KNOWLEDGE RETRIEVAL MODULE)

Total Queries	Correct Similarity Score	Incorrect Similarity Score
13	6	7

6. CONCLUSION

This paper has proposed a semantic categorization based knowledge upload service. This service consists of two parts i.e. Knowledge Recording and Knowledge Retrieval. Using the knowledge recording part different speeches of an individual are recorded and ultimately classified into personal or general categories. The categorization is done using semantic techniques that do not require a training dataset like machine learning based categorization approaches. Later on people who want to utilize the already uploaded knowledge can easily record their question using voice on a subject and retrieve the response. The advantage of our proposed system is that it does not require any special hardware for knowledge upload. Contrary to our proposed system other systems require sophisticated hardware that is connected with human brain. This may be harmful to human mind in the long run. Through the proposed service in this paper precious human knowledge can be saved for the benefit of mankind.

7. FUTURE WORK

In future we want to add support for Japanese language in the system. Japanese is the mother tongue of the main author. Our current principal supported language is English. Availability of tools for Japanese language is lesser in comparison with English language. We need to find out high accuracy libraries in Java for speech recognition and NLP of Japanese language. In addition to the above we can use linked data for semantic based text categorization. Linked open data includes upper level ontologies connected with domain specific knowledgebase that help in covering more domains. Extensive system evaluation is also one of the important future tasks. Additionally, the overall system accuracy need to be improved.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Forsberg, M., "Why is Speech Recognition Difficult?", Technical Report, Chalmers University of Technology, Sweden, 2003.

[2]   Nadkarni, P.M., Ohno-Machado, L., and Chapman, W.W., "Natural Language Processing: An Introduction", Journal of the American Medical Informatics Association, Volume 18, pp. 544-551, 2011.

[3]   Sotala, K., "Advantages of Artificial Intelligences, Uploads, and Digital Minds", International Journal of Machine Consciousness, Volume 4, pp. 275-291, 2012.

[4]   Sotala, K., and Valpola, H., "Coalescing Minds: Brain Uploading-Related Group Mind Scenarios", International Journal of Machine Consciousness, Volume 4, pp. 293-312, 2012.

[5]   Cloud Speech API, https://cloud.google.com/speech/ (Visited: 23rd November, 2018)

[6]   Text to Speech | TTS SDK | Speech Recognition (ASR), https://www.ispeech.org/(Visited: 23rd November, 2018)

[7]   Watson Speech to Text, https://www.ibm.com/watson/services/speech-to-text (Visited: 23rd November, 2018)

[8]   Speech Recognition |Haven OnDemand, https://dev.havenondemand.com/apis/recognizespeech (Visited: 23rd November, 2018)

[9]   Lamere, P., Kwok, P., Gouvêa, E., Raj, B., Singh, R., Walker, W., Warmuth, M., and Wolf, P., "The CMU Sphinx-4 Speech Recognition System", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Volume 1, pp. 2–5, Hong Kong, 2003.

[10]   Gaida, C., Lange, P., Petrick, R., Proba, P., Malatawy, A., and Suendermann-Oeft, D., "Comparing Open-Source Speech Recognition Toolkits", Organization of Alberta Students in Speech, Alberta, 2014.

[11]   Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S.J., and McClosky, D., "The Stanford Core NLP Natural Language Processing Toolkit", Proceedings of 52nd Annual Meeting of the Association of Computational Linguistics, pp. 55-60, USA, 2014.

[12]   Bontcheva, K., Cunningham, H., Tablan, V., Maynard, D., and Hamza, O., "Using GATE as an Environment for Teaching NLP", Proceedings of Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics, Volume 1, pp. 54-62, Philadelphia, 2002.

[13]   Suchanek, F., Hoffart, J., Kuzey, E., and Lewis-Kelham, E., "YAGO2s: Modular High-Quality Information Extraction with an Application to Flight Planning", Business Technologies and Web, Volume 214, pp. 515-518, 2013.

[14]   Free Text to Speech – Free Online TTS Service, http://www.fromtexttospeech.com/(Visited: 23rd November, 2018)

[15]   Sandberg, A., and Bostrom, N., "Whole Brain Emulation: A Roadmap", Future of Humanity Institute, Oxford University, 2008.

[16]   Chalmers, D., "Mind Uploading: A Philosophical Analysis", The Journal of Consciousness Studies, Volume 17, pp. 7-65, 2010.