A Study of Decrease in Artificial Immune System towards Human Computer Interaction

S Sathya Franklin[1], Dr. Kailash Aseri[2]
[1]Research Scholar, Sunrise University Alwar.
[2]Professor, Sunrise University Alwar

Abstract

Human-Computer Interaction (HCI) deals with both the ways in which humans interact with computers and the design technologies that let humans interact with computers in novel ways. Unlike other tools with limited usage, a computer has many usages and it takes place as an open-ended dialog between the user and the computer that relates human-computer interaction to human-to-human interaction and an analogy which is crucial to theoretical considerations in the field. Though the term HCI was first coined in 1980 but it was later popularized by "The Psychology of Human-Computer Interaction" in 1983. Since then, it has been received priority by many researchers for its growth and that necessitates further research for more novel and robust designs. The current status of human-computer interaction stands at the intersection of computer science, behavioural sciences, design, media studies, and several other fields of study. The growth in human-computer interaction field has been in quality of interaction, and in different branching in its history. Machines are insignificant unless they are user friendly. The demand of user has been changing very frequently because of which technological intervention is necessary. Researchers of different fields are now concentrating on the concepts of multimodal, intelligent, adaptive and active interfaces to enable the computers to dynamically adapt content of interfaces and develop responsiveness to cognitive load and human emotion. The modified user interface offers new scopes of interaction to which previous research findings may appear to be inadequate and/or obsolete. Moreover, not all existing technologies are accessible and/or affordable by all community.

Keywords: *Human-Computer Interaction, Machines, computers, accessible, human emotion*.

1. INTRODUCTION

With the advent of highly reliable computer technology in the last few decades, the Human-Computer Interaction is definite to play major roles in the areas of artificial intelligence, cognitive science, psychology, sociology, anthropology, management science and engineering. Artificial Intelligent Algorithms have enhanced this technique to resolve many challenges of design, construction and implementation of Human Computer Interaction (HCI) systems. Different types of HCI systems are designed and implemented for same purpose. Research in this area is enormous; a complete division of computer science is dedicated to this subject to make appropriate design for differently abled people. Different internet browsers and available websites are also an example of Human Computer Interface that uses appropriate buttons and menus to make it more users friendly. A Touch Screen, Computer Mouse, program on Mac or Windows machine that comprises icons of disk drives, trashcan, folders, pull-down menus, sensors etc. are all examples of advances in Human-Computer Interface design [100]. It is an emerging area of research that can discover unexplored areas to relate better communication between users and computer. The Human-Computer Interaction seeks to determine the most efficient way to design understandable electronic messages for the user. It involves Study, Planning and Design of the interaction between users and computer. There are two types of Human-Computer Interactive Architecture: Unimodal and Multimodal System.

2. LITERATURE REVIEW

Sunhera, D. & Tejaswi, V. (2016) has implemented a speech based home automation system using Bluetooth and Global System for Mobile (GSM). The system is implemented with few speech commands that are stored within the application database. The system does not accept any command(s) that is not stored in the database. It is also reported that more speech commands could be added for further robustness of the system and therefore the reduction of memory requirement of the database would be increased.

A Control of electrical appliances using voice recognition technique is implemented by JArthi, J. E. & Jagadeeswari, M. (2014) where the system takes the necessary actions through speech commands. The rule based system accepts few speech commands to perform certain operations. It is also mentioned in the paper that the system(s) that is designed with minimum speech commands, conflicts less to choose correct commands.

Mittal, P. & Singh, N. (2013). illustrates the issues and challenges of a Speech Based Command and Control System in Mobile Phones. The boundaries of limited storage space in mobile phone devices are discussed in this paper. The paper depicts that the offline speech recognition system uses local database loaded in a system; where the online speech recognition system(s) utilizes global database that resides in their server to execute.

The same issues were reported by Forsberg, M. (2003) in his paper "Why is Speech Recognition Difficult?". The paper describes limitations on Body Language, Linguistic approach, external Noise during recognition, Continuous Speech, Channel and the Speaker Variability, Pronunciation of different users etc. It is stated that a huge amount of data and search space are the barrier in this field of research. Other ambiguities of speech recognition viz. homophone ambiguity, word boundary ambiguity, etc. were also the cause of concern. In the paper, it is concluded that in every case of speech recognition, it is impossible to distinguish between two or more than two homophones at the word level speech recognition.

Google uses its speech engine that is accessible from any smart devices. Very few commands are used by Google Now in offline mode whereas a large amount of speech command database works online. In 2017, McGraw, I., Prabhavalkar, R., Alvarez, R., Arenas, M. G., Rao, K., Rybach, D., Alsharif, O., Sak, H., Gruenstein, A., Beaufays, F. & Parada, C. (2016) proposed a new algorithm that employs a quantized Long Short-Term Memory (LSTM) acoustic model trained with connectionist temporal classification (CTC) to directly predict phoneme targets, and further reduce its memory footprint using an SVD-based compression (Singular Value Decomposition) scheme. The Bayesian interpolation is used to minimize the memory by a single language model for both dictation and voice command domains. Currently this technique is used in newly launched Google Devices like Google Nexus, Google Pixel, etc. in an offline mode. To reduce the memory, the system uses local words such as specific information like phone contacts, messages, or from other installed apps, etc. This technique follows a user of a specific device and does not affect other devices or the main Google Acoustic Database.

Kumar, A. J., Schmidt, C. & Koehler, J. (2017). established a Question Answering System with Knowledge Graph based Speech Interface. The system utilizes an existing automatic speech recogniser by adapting acoustic and language models for the speech interface. During the research, less attention is paid to the recognition errors and language processing inaccuracies. Unsupervised data is collected for training or testing to detect the correctness or incorrectness of user's responses. The research follows Spoken Language Recognition and Understanding (SLRU) in the Automatic Speech Recognition System with the limitations of out-of-vocabulary words, multi-domain speech recognition etc.

Maas, A. L., Xie, Z., Jurafsky, D. & Ng, A. (2015) produced a qualitative difference between transcriptions by Lexicon-Free Speech Recognition (LFSR) approach and a standard Speech Recognition System using Artificial Neural Network (ANN). According to this paper Google Speech, Apple Siri and Microsoft Speech rely on Spoken Language Understanding with Sentence Level Recognition. The limitations of Word Level Speech Recognition technique still exist in these well-established speech recognition techniques as they follow the Linguistic Approach.

Schuppler, B. & Schrank, T. (2018) illustrates how a phonetic model can be used to automatically disambiguate homophones using the example of German pronouns on short or syntactically incomplete context. The Kiel Corpus of Spontaneous Speech Database for German Language is used that contains conversations from 18 speaker pairs and each conducting 7 dialogues. The model is trained with lexical features based on Deep Neural Networks (DNNs) and finally 3% more accuracy1 over homophones was achieved. The research is functional on Sentence Level Speech Recognition (SLSR) where the homophone is recognized through the trained lexicalbased acoustic model.

Mishra, R. K. & Singh, S. (2014). worked with Hindi homophone words. In the research, the secondary meanings of the ambiguous homophone word are examined. It is also examined that the sentence context is biased towards the dominant meaning. In their research, both the homophones with dominant and subordinate meaning are extracted and preferred through a supervised method in syntactical model.

The study by Chen, W. F., Chao, P. C., Chang, Y. N., Hsu, C. H. & Lee, C. Y. (2016) utilized the Event-Related Potentials (ERPs) to discriminate two types of phonological-to-orthography (P-to-O) mapping in Chinese language, namely homophone density and orthographic consistency. These results advise that orthographic information influences phonological representations by activating during on-line spoken word recognition. The homophone density effect that indexes the orthographic variation at the character level was found in a later time window to reflect the difficulty in successful retrieval and decision making. The homophone density effect has been localized to the temporal context for lexical/semantic processing.

Müller, I., Ratsch, C. & Faerber, F. (2014). has worked on reducing the words from a English Dictionary in the year of 2014. In their research they used bit compression technique to reduce an English dictionary. The maximum string length they used in their dictionary is ten. In their research they succeed to reduce 60% memory from the English dictionary.

Obrist, M., Ranasinghe, N. & Spence, C. (2017). It is also reported in an article that the Dragon Naturally Speaking Software does not distinguish between homophones. It is further mentioned that the technique uses Supervised Algorithm to recognize a proper Homophone word. In this software, if the system picks any wrong homophone word, then the user can choose his/her preferable word by using "Speak Again" phrase repetitively to obtain the proper word. In this software, no automatic technique is applied to increase the performance to select a correct homophone during speech recognition.

Fernandes, D. A. B., Freire, M. M., Fazendeiro, P. A. P. & Inácio, P. R. M. (2017). introduced the principles of artificial immune systems and surveyed several works applying such systems to reduce computer security problems. The Artificial Immune System Algorithms like Negative Selection, Clonal Selection, and Artificial Immune Network Algorithm etc. are mentioned and summarized with the security issues in cloud environments'. The novel applicability of the mentioned systems is also elaborated in this paper.

Jaspan, H. B., Lawn, S. D., Safrit J. T. & Bekker, L. G. (2006). displayed how the Adaptive Immune System is a principal significance of vaccinology. The method from this paper is used for designing the maturity model for homophone affinity. J. J. T. Owen analysed various forms of immunodeficiency and reported the importance of maturation technique for an Artificial Immune System.

Tang, N. & Vemuri, V. R. (2005). proposed a document clustering method using AINet. Based on the immune network and affinity maturation principles, the AINet performs an evolutionary process on the raw data that removes the data redundancy and retrieves clustering results. During the implementation, 15% of its data after applying AINet clustering method was taken. This paper indicates the positivity of using the Hierarchical Clustering Method associated with Artificial Immune Network (AINet).

Rassam, M. A. & Maarof, M. A. (2012). proposed an Artificial Immune Network Clustering approach for Anomaly Intrusion Detection. In the research, the application of bio-inspired cluster approach and detected anomaly from KDD-99 dataset applying AINet Algorithm was investigated. During the investigation, only 10% of the total data for their experiment from the KDD-99 dataset1 was taken. The process of using the AINet clustering approach is influential for the current research.

Mankoff, J., Hudson, S. E. & Abowd, G. D. (2000). presented different ambiguities during speech recognition system and provided an integrated toolkit level support in their application 'Burlap' by applying pen, speech and user identification method. In their research, supervised approach was used to avoid the homophone. The system did not follow any automatic method to remove the homophone ambiguity.

3. UNIMODAL SYSTEM

HCI such as smart phones, biometric attendance system, speech interface technology, face recognition system etc., which is based on only one modality is termed as Unimodal System [3]. The major categories of Unimodal systems are illustrated below.

Visual-Based

This is the most important area in Human-Computer Interaction (HCI) research; wherein researchers try to deal with different aspects of human responses which can be recognized as a visual signal [3]. Expression Analysis, Body Movement Tracking and Gesture Recognition etc. are the major constituents of Visual based HCI.

Audio-Based

The audio-based communication between computer and user is another significant area of HCI systems which deals with the information acquired by different audio signals. The research possibilities in this area are Speech Recognition, Auditory Emotion Analysis, Speaker Recognition, Human-Made Noise/Sign Detections (Gasp, Sigh, Laugh, Cry, etc.), and Musical Interaction etc.

Sensor-Based

With the improvement in technology, the sensors are becoming tiny, cheaper, accurate, reliable, efficient, responsive, and leading to a reliable and cost effective utility to extend a natural feeling for controlling computer applications, smart TVs and game consoles. Capturing and recognizing gestures, expressions and vocal inputs led to the evolution of human-computer interfaces. The harmony of these different areas is that at least one physical sensor is used between user(s) and machine(s) to provide the interaction [3]. Very primitive and sophisticated sensors are Pen-Based Interaction, Mouse & Keyboard, Joysticks, Motion Tracking Sensors and Digitizers, Pressure Sensors, and Taste/Smell Sensors etc.

Multimodal System

Multimodal human-computer interaction states the way of interaction between virtual and physical environment through natural modes of communication, enabling a free and natural communication, interfacing users with automated systems in both input and output. There are two views on multimodal interaction, viz; Human perception and control, wherein the word modality refers to human input and output channels. The second view focuses on using two or more computer input or output modalities to build system that makes synergistic use of parallel input or output of these modalities.

The exact number of input modes, their types and the way in which they work together may vary widely from one multimodal system to another multimodal system. Multimodal interfaces integrate a diverse combination of gesture, speech, gaze, facial expressions and other non-conventional modes of inputs [3].

4. SPEECH RECOGNITION SYSTEM

The principle of Speech Recognition System (SRS) comprises of identifying the human speech or language uttered by human as sound waves and then digitizing, decoding and transmitting these incessant or analog impressions to appropriate strings or sentences. The block diagram of a Speech Recognition System (SRS) is given in Figure 1.1.

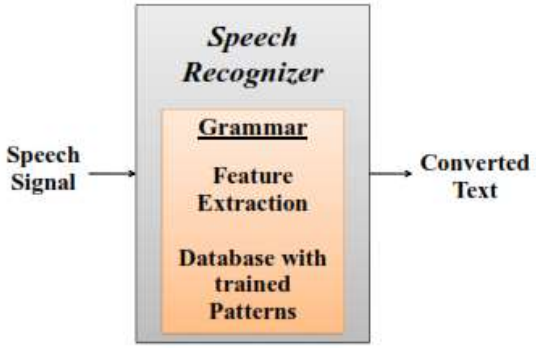

Figure 1.1 Block Diagram of SRS

The various approaches of a Speech Recognition System are:

Acoustic Phonetic Approach

Acoustic Phonetic Approach consists of two variables, wherein the first one is the concealed status of the phonemes stored in the computer database and the second one is perceptible incidence segment of the digital gesture. Every phoneme has its own probability and the segment is coordinated with the phoneme according to the probability. The corresponding phonemes are then unruffled collectively to form the accurate texts according to the lexicon system of the given language. A nondeterministic probability model, viz. Hidden Markov Model (HMM) is used in the Acoustic phonetic approach [79] which is extended by this approach to carry out the probabilistic analysis.

5. CONCLUSION

The main contribution of this paper dissertation is on the advancement of Speech Recognition System. The study has been concentrated towards the minimization of two limitations of speech recognition system viz.; memory requirement and the homophone ambiguity in word level speech recognition system. Since the word level recognition plays an important role wherein the homophone ambiguity still exist, the ability of selecting an exact homophone word by the existing speech recognition systems needs to be improved. The String and Synonym Matching Algorithm (SSMA) has been implemented to reduce the memory requirement of system database from any existing Speech Recognition System (SRS). On the other hand, considering the unique property of Immunological System that learns to improve self-performance based on its preference, the Artificial Immune System (AIS) is applied to reduce the homophone ambiguity from word level Speech Recognition System (SRS). Supervised classification has been attempted in the event of user-system interaction, while unsupervised classification is used to depict the elimination of user interaction with the system. To reduce the high probability of obtaining the incorrect homophone word, the Affinity Maturity Model (AMM) has been employed with the maturation model of the dataset at the early stage.

6. REFERENCES

1. Sunhera, D. & Tejaswi, V. (2016). Implementation of speech based home automation system using Bluetooth and GSM. IEEE International Conference on Signal Processing, Communication, Power and Embedded System (SCOPUS), pp.807-813.
2. Arthi, J. E. & Jagadeeswari, M. (2014). Control of Electrical Appliances through Voice Commands. IOSR Journal of Electrical and Electronics Engineering, e-ISSN: 2278-1676, p-ISSN: 2320-3331, Vol.9, Issue.1, Ver.V, pp.3-18.
3. Mittal, P. & Singh, N. (2013). Speech Based Command and Control System for Mobile Phones: Issues and Challenges. Second IEEE International Conference on Computational Intelligence & Communication Technology (CICT), pp.729- 732, doi:10.1109/CICT.2016.150.
4. Forsberg, M. (2003). Why is Speech Recognition Difficult?. [Available At: http://www.speech.kth.se/~rolf/gslt_papers/MarkusForsberg.pdf] [Access Date: Dec 15, 2015].
5. McGraw, I., Prabhavalkar, R., Alvarez, R., Arenas, M. G., Rao, K., Rybach, D., Alsharif, O., Sak, H., Gruenstein, A., Beaufays, F. & Parada, C. (2016), Personalized Speech Recognition on Mobile Devices. 2016 IEEE Conf. on Acoustic, Speech and Signal Processing (ICASSP), Mar 20-25, 2016, ISSN: 2379-190X, doi: 10.1109/ICASSP.2016.7472820.
6. Kumar, A. J., Schmidt, C. & Koehler, J. (2017). A knowledge graph based speech interface for question answering systems. Netmedia, Fraunhofer IAIS, Germany and Enterprise Information Systems, University of Bonn, Germany, Vol.92, pp.1-12.
7. Maas, A. L., Xie, Z., Jurafsky, D. & Ng, A. (2015). Lexicon-Free Conversational Speech Recognition with Neural Networks. Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp.345-354, doi:10.3115/v1/N15-1038.
8. Schuppler, B. & Schrank, T. (2018). On the use of acoustic features for automatic disambiguation of homophones in spontaneous German. Signal Processing and Speech Communication Laboratory, Graz University of Technology, Elsevier, Inffeldgasse 16c, 8010 Graz, Austria.
9. Mishra, R. K. & Singh, S. (2014). Activation of shape and semantic information during ambiguous homophone processing: eye tracking evidence from Hindi. Cognitive Processing, Springer, Vol.15, Issue.4, pp.451-465, doi:10.1007/s10339-014-0622-4.