Interactive Artificial Intelligence

Sakshi Kapoor[1], Shivani Chaturvedi[2], Gayatri Mandage[3], Shraddha Bhadane[4], Amita Jajoo[5]
[1, 2, 3, 4, 5] D.Y. Patil College of Engineering, Akurdi , Pune

Abstract- *Learning and interaction are viewed as two related but distinct topics in developmental robotics. Many studies focus solely on either building a robot that can acquire new knowledge and learn to perform new tasks, or designing smooth human-robot interactions with pre-acquired knowledge and skills. The present paper focuses on linking language learning with human-robot interaction, showing how better human-robot interaction can lead to better language learning by robot. Toward this goal, we developed a real-time human-robot interaction paradigm in which a robot learner acquired lexical knowledge from a human teacher through free-flowing interaction.*

I. INTRODUCTION

Language is a central component of human intelligence which is fundamental and essential for human-human everyday communication. A basic function of language is to provide linguistic labels of objects and activities which people to refer to them in speech and share experiences in everyday communication [1]. Therefore, learning, understanding and using human languages by humanoid robot is critical for seamless human-machine interaction.

AI has always been on the pioneering end of computer science. Advanced-level computer languages, as well as computer interfaces and word-processors owe their existence to the research into artificial intelligence. The theory and insights brought about by AI research will set the trend in the future of computing. The products available today are only bits and pieces of what are soon to follow, but they are a movement towards the future of artificial intelligence. The advancements in the quest for artificial intelligence have, and will continue to affect our jobs, our education, and our lives.
Learning and interaction are viewed as two related but distinct topics in developmental robotics. Many studies focus solely on either building an Artificial Intelligence that can acquire new knowledge and learn to perform new tasks, or designing smooth human-artificial intelligence interactions with pre-acquired knowledge and skills. The present project focuses on linking language learning with human-machine interaction [4], showing how better human-artificial intelligence interaction can lead to better language learning by machine. We aim to build a real-time human-machine interaction paradigm in which a robot learner acquired lexical knowledge from a human teacher through free-flowing interaction [2,5].

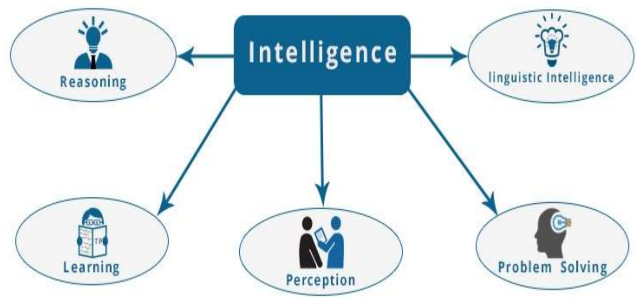

Figure 1. Components of Intelligence

1. **Reasoning**: It is the set of processes that enables us to provide basis for judgement, making decisions, and prediction.
2. **Learning** − It is the activity of gaining knowledge or skill by studying, practising, being taught, or experiencing something. Learning enhances the awareness of the subjects of the study. The ability of learning is possessed by humans, some animals, and AI-enabled systems.
3. **Problem Solving** − It is the process in which one perceives and tries to arrive at a desired solution from a present situation by taking some path, which is blocked by known or unknown hurdles. Problem solving also includes decision making, which is the process of selecting the best suitable alternative out of multiple alternatives to reach the desired goal are available.
4. **Perception** − It is the process of acquiring, interpreting, selecting, and organizing sensory information. Perception presumes sensing. In humans, perception is aided by sensory organs. In the domain of AI, perception mechanism puts the data acquired by the sensors together in a meaningful manner.
5. **Linguistic Intelligence** − It is one's ability to use, comprehend, speak, and write the verbal and written language. It is important in interpersonal communication.

II. SYSTEM ARCHITECTURE

Syntax is a form of grammar. It is concerned primarily with word order in a sentence and with the agreement of words when they are used together. So it is, in a sense, acting as a kind of 'police officer' for the way in which sentences are constructed. English is a language that has a structure known as SVO [3,6].

"The cat (subject) washes (verb) its paw (object)."

This is the correct word order and also there is agreement between the words. If there were no agreement within the sentence, it could read, "The cat washes their paw." This does not make sense. The cat may have four paws, but it is only washing one paw. For there to be agreement, the possessive 'it' has to be correct. Thus "The cats (plural) wash their (plural) paws (plural)." This is the correct use of the plural possessive (their).

At first, syntax can seem daunting and it is always difficult initially to understand what a subject, verb or object actually is. It can also be difficult to understand whether agreement between the subject, verb or object is right or wrong. There are lots of tools such as grammar checkers, programmes or worksheets to help you get to grips with syntax and to make sure that you have the right word order and that within the sentence there is always agreement between the words, tenses and so on [7]. It is true that syntax can take some time to master, but, once you understand its principles and can apply it without too much effort, then it really is worth the effort, since it will greatly improve your written English.

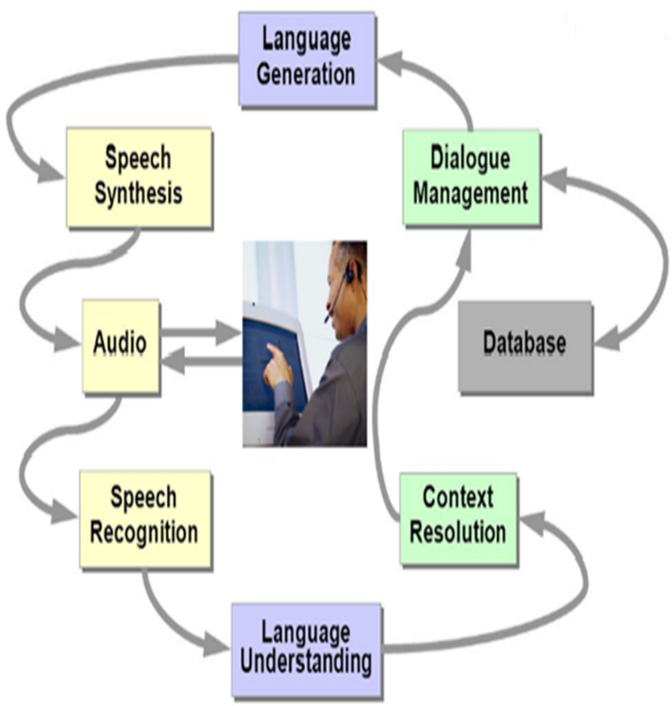

Figure 2. System Architecture

III. METHODS AND ALGORITHMS

The proposed system contains the following modules:

1. Face recognition
2. Language Processing and Storage

Face Recognition:

To recognize the unique feature of the face, the system recognizes the basic pixels and creates the base feature of the face. Further a similar possible image of the face is created using this base feature and the image is classified under various folders. The image data previously stored is compared and the contents from the image are detected and conclusions are made and conversation is done accordingly [8]. Eigen face algorithm has been used to implement face recognition. If the face is recognized, user is greeted by his name. New users are registered by taking a snapshot of their face and storing it for further interactions. Human and robot interact with the help of speech recognition and

Language processing. Human responses from all the chats are stored and the database keeps growing with every interaction. Sentences are processed to find out its object and a dictionary made in MySQL is used to store all the words and its related part of speech, example: eat-verb, chair-noun, beautiful-adjective, etc. Since the database grows with every interaction, the robot's accuracy and aptness increases with it.
Language Processing:

The AI perceived human speech in real time. A speech recognition software (Dragon naturally speaking from nuance, LLC) was first applied to convert speech into text [9]. Next, a 10-second temporal window was used to define a local context. Spoken utterances within a context were then compared to spot frequent words that were further processed in two specific ways. First, frequent words were selected as candidate words for object names and would be linked with visual input to compute word-object associations. Second, frequent words were added to a word list that the AI maintained to keep track of those words that the robot heard before. In speech production, the AI would selectively produce those words. In a way, this mechanism made the AI like a copycat – repeating what it just heard most frequently in the recent past. For example, if a human teacher happened to say "hello" to the AI in multiple times, the AI would say "hello" back to the human teacher. Thus, the AI learning system was transparent and straightforward purely driven by statistical regularities in the data without complicated inference [1]. The goal was to show how better data from interaction may lead to better statistical learning.

IV. IMPLEMENTATION

Eigenface Face Recognition:

We start with face recognition process and after that we start with speech to text conversion .Then we analyse the sentence to figure out that it is descriptive or interrogative type of sentence. To analyze what is being said in the sentence we study the English grammar rules.

Stanford NLP Package:

"A sentence can be formed from a noun followed by a verb"

Rules in this notation are phrase structure rules. They look like our earlier morphology rules. Because these rules make phrases, not words. As in our morphology unit, we can use 'tree structures' to illustrate the way these rules make larger structures. We start from taking out the verb then moving on to figuring out the subject of the sentence (noun) which specifies the 'what' 'who' of the sentence. Then there are descriptive words (adverb/adjective) which specify the 'when' 'where' 'how' of the sentence.

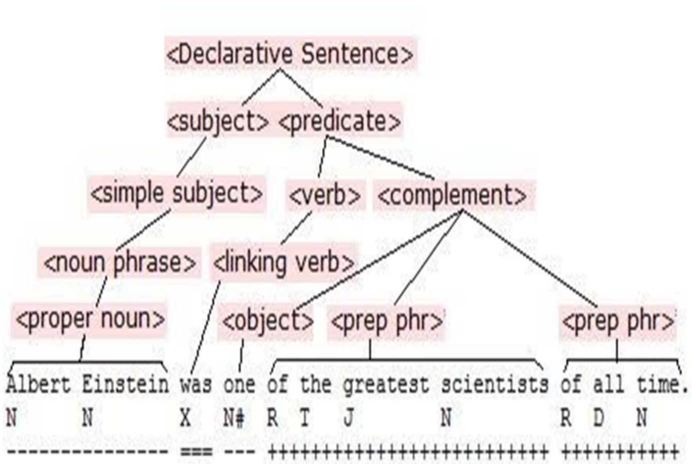

Figure 3. Classification of a sentence

In Language processing we make use of Stanford NLP Library available online with provides us with a set of tools to examine a sentence's intent, its object and sentiment [10, 11]. Taking out the subject of the discussion is an important part while figuring out the conversation's focal point. Along with that, figuring out the sentiment of the user is equally important in order to figure out the difference between, "It was a rough day today" and "It was a nice day today". The AI must be smart enough to analyze the sentences to determine which kind of reply befitted the best, "Oh it's sad to hear that!" or "That's awesome!". This is generally done by sentiment analysis provided by the NLP by distinguishing the behavior as, positive, very positive, negative, very negative and null.

V. CONCLUSION

AI is at the center of a new enterprise to build computational models of intelligence. The main assumption is that intelligence can be represented in terms of symbol structures and symbolic operations which can be programmed in a digital computer. There is much debate as to whether such an appropriately programmed computer would be a mind, or would merely simulate one, but AI researchers need not wait for the conclusion to that debate, nor for the hypothetical computer that could model all of human intelligence. Aspects of intelligent behavior, such as solving problems, making inferences, learning, and understanding language, have already been coded as computer programs, and within very limited domains, such as identifying diseases of soybean plants, AI programs can outperform human experts. Now the great challenge of AI is to find ways of representing the commonsense knowledge and experience that enable people to carry out everyday activities such as holding a wide-ranging conversation, or finding their way along a busy street. Conventional digital computers may be capable of running such programs, or we may need to develop new machines that can support the complexity of human thought.

One of the strongest demonstrations of linking interaction and learning is to show that even if the same learning algorithm is given without any changes to process additional feedbacks from human users, the learning system can improve its performance through social interaction. That is, social behaviors from a machine learner make a human teacher provide better and more teaching signals, and by doing so, better and more training data through social interaction lead to better statistical learning. We can view this as active learning through interaction, in which the machine learner generated behaviors to reveal its current learning states, and those behaviors gave human teachers first-hand information about what the robot or machine learner needed next for successful learning.

REFERENCES

[1] H. H. Clark and M. A. Krych, "Speaking while monitoring addressees for understanding," Journal of Memory and Language, vol. 50, pp. 62-81, 2004.

[2] "Learning to Interact and Interacting to Learn: Active Statistical Learning in Human-Robot Interaction" by Chen Yu, Tian Xu, Yiwen Zhong, Seth Foster and Hui Zhang in 2014 International Joint Conference on Neural Networks (IJCNN) July 6-11, 2014, Beijing, China.

[3] M. Scheutz, P. Schermerhorn, and J. Kramer, "The utility of affect expression in natural language interactions in joint human-robot tasks," in Proceedings of the ACM

SIGCHI conference on Human-robot interaction, 2006, pp. 226-233.

[4] A. L. Thomaz and C. Breazeal, "Teachable robots: Understanding human teaching behavior to build more effective robot learners," Artificial Intelligence, vol. 172, pp. 716-737, 2008.

[5] A.-L. Vollmer, K. S. Lohan, K. Fischer, Y. Nagai, K. Pitsch, J. Fritsch, et al., "People modify their tutoring behavior in robot-directed interaction for action learning," in IEEE International Conference on Development and Learning, 2009, pp. 1-6.

[6] Arnold M.Zwicky Robert N. Kantor, “A Survey of Syntax” Stanford University, Language Development Gramr and Semantics, Cal 1980.

[7] Literature Survey On Syntax Parser For English Language Using Grammar Rules. Shraddha Anantpure, Hinal Jain, Neha Alhat, Smita Bhor, Prof. Shanti Guru Computer Engineering, DYPCOE, Akurdi, University of Pune, India.

[8] A survey of Face Detection, Extraction and Recognition. Yongzhong Lu, Jingli Zhou, Shengsheng Yu National Storage System Laboratory School of Software Engineering Huazhong University of Science and Technology Wuhan, 430074, P. R. China

[9] A Survey Paper on Automatic Speech Recognition by Machine , Dinesh Kumar Dansena Department of Computer Science and Engineering Raipur Institute of Technology Raipur, India ,Yogesh Rathore (Assistant Professor), Department of Computer Science and Engineering Raipur Institute of Technology Raipur, India.

[10] Survey on Various Natural Language Processing Toolkits, L.B. Krithika and Kalyana Vasanth Akondi School of Information Technology and Engineering, VIT University, Vellore-632014, Tamil Nadu, India.

[11] The Stanford Core NLP Natural Language Processing Toolkit Christopher D. Manning Linguistics & Computer Science Stanford University manning@stanford.edu, Mihai Surdeanu SISTA University of Arizona msurdeanu@email.arizona.edu, John Bauer Dept of Computer Science Stanford University horatio@stanford.edu.