Tourist Guide with Human Machine Interaction using Spoken Commands in the Natural Language

Kiran P. Somase
Department of Computer
Engineering
DIT Pimpri

Satyaki Banerjee
Department of Computer
Engineering
DIT Pimpri

Souravi Biswas
Department of Computer
Engineering
DIT Pimpri

ABSTRACT
This paper proposes NLP based system to facilitate user interaction by verbal communication in a seamless manner. With the advancement of technology the human machine interaction becomes an unavoidable scenario. Introduction of natural communication is an addition that has to be made sooner or later in every department. Devices that accept interaction based on Natural Language, like, using voice commands, can understand basic human orders or answer questions whenever user expressions fit into the known language pattern. The purpose of this work is to introduce a system which is capable of interacting with human and is able to assist them using authentic knowledge and provide a smooth experience.

General Terms
Artificial Intelligence, Natural Language Processing

Keywords
Human-Device Interaction, Natural Language Processing

1. INTRODUCTION
The goal of proposed system is to introduce a new way of assisting the tourists using human machine interaction. Thus reducing the hassle of local guides and providing them with authentic information. To achieve the goal of the proposed system we will use seamless human machine interaction which will be achieved using voice commands. The system will be functioning with sufficient capabilities to run uninterrupted for required durations. To achieve comfort, the user is allowed to enter text as an input instead of speech, if he so desires. Inclusion of text based output if machine generated voice is not audible to user due to unavoidable external circumstances. The system is to provide reliable information to the tourist for authentic experience.

2. NATURAL LANGUAGE PROCESSING
Natural Language Processing (NLP) employs computational techniques for the purpose of learning or understanding human language content. Natural Language Processing is used when the time comes to take to the next level the interaction between humans and devices [10].

Optimally, the virtual world should understand the user regardless of the terms used to interact with it, guessing the intention behind those terms, i.e., different sentences with the same goal should trigger the same outcome. For instance, the sentences "shut down the television" and "switch off the TV" have the same intended objective. For humans, realizing that both sentences should have the same result is natural and intuitive, but not for devices. In this field, some tools have been developed.

The system processes Natural Language in two stages. The first stage requires searching synonyms and related words to create a database of terms that can expand the keywords provided by the device to describe an action. These data feeds the Cluster Manager. The second process is the analysis of the user's queries to create the Order Primitives by the Order Interpreter [4].

2.1 Device Metadata Processing
Device actions are described using natural language tags. The set of tags are likely to be provided by the manufacturer or a developer but it also may be provided by the end-user. The main problem of selecting the appropriate words is that they must fit into the terms that any potential user would use. This is a double-edge sword. On the one side, a device description containing many related tags can target a large number of potential users by default.

The reason is that habits, culture or age-segmentation affects the way users use Natural Language so it can be thought that adding more words to a description would increase the number of potential users. However, this can dramatically increase the number of false positives as device actions would be candidate for more orders than they should be. On the other side, a succinct device description that precisely describes an action may decrease the number of false positives improving user experience but may target a smaller number of potential users.

To tackle these problems, the system performs a controlled expansion of keywords by fetching synonyms and related words from online dictionaries. This process can be controlled attending to the evolution of Natural Language expressions and cultural traits. How the Natural Language evolved is studied by the system analyzing the context in terms of which words or terms are most used by the user regulating the dataset of synonyms or related words handled by the Controller. Thus, manufacturers or developers are expected to add a small set of terms in natural language that describe the action. The subsequent word expansion process will handle user diversity [9].

2.2 Analyzing User Natural Language Queries
Once word clusters have been created after device tag expansion and they have been assigned Ids, the system is ready for user orders. When a user formulates an order in Natural Language, the system will select the important words from the order and compares them with the words in the clusters generated during device metadata processing inferring which device action should be triggered.

To do so, the Order Detector recognizes a user order and sends it to the Order Interpreter. The Order Interpreter builds

an Order Primitive based on the relationships between words. Stanford University has developed a toolkit for natural language processing that provides these relationships. Adverb that serves to modify the meaning of other word. The Order Primitive contains the set of words of the input order in Natural Language filtered by the relations among them. The resulting set of words after filtering is constituted by all the words in the input order that have any of the previous relationships with other word within the input set. All these words are considered when it comes to interpreting the user order. Superfluous words, those with no relation with the rest, are discarded.

For each of these words, the Workflow Orchestrator searches in the database of clusters by querying the Cluster Manager. The Cluster Manager provides the Id of the cluster where each word is included. Thus, the resulting Order Primitive contains the filtered words (selected words) and the collection of cluster identifiers in which these words are included [2].

3. SYSTEM ARCHITECTURE
The input from the user would be taken and supplied to the Web Interface via the Dual microphone. The dual microphone(mic) is present for cancellation of noise. The required data to be supplied to the users would be searched for in the database. Then this searched data would be supplied to the users. The system architecture is given below:
Explanation:

For general,

Let 'I' be the input,

'P' be the process

'O' be the output

where,

Input I=Speech(By user)

Process P=NLP(Natural Language Processing)

Output O=Speech(By device)

For 'ith' no of inputs,

M={{I1,I2,I3,I4},{P1,P2,P3,P4},{O1,O2,O3,O4}}

where,

M:{{no of input},{no of process},{no of output}}

For input I1,

Let I1=speak Ajanta & Ellora Caves, P1=NLP, O1=information about Ajanta & Ellora Caves in audio format.

For input I2,

Let I2=speak Painting Monalisa, P2=NLP, O2=information about Painting Monalisa in audio format.

For input I3,

Let I3=speak Shaniwar Wada, P3=NLP, O3=information about Shaniwar Wada in audio format.

Let I4= speak Banerghatta National Park, P4=NLP, O4=information about Banerghatta National Park in audio format.

3.1 Conclusion and Future Scope
The idea is being implemented as a web based app, so that the tourists do not need to download it. It would not create a problem for storage as it would not occupy the storage in the phone. It would be a reliable and inexpensive source of information for the users. Therefore, the tourists won't be tricked to pay unreasonable money to the guides and the tourists will also not be tricked in any other way. This can include the feature of extracting information directly from the internet if the user so desires in later updates.
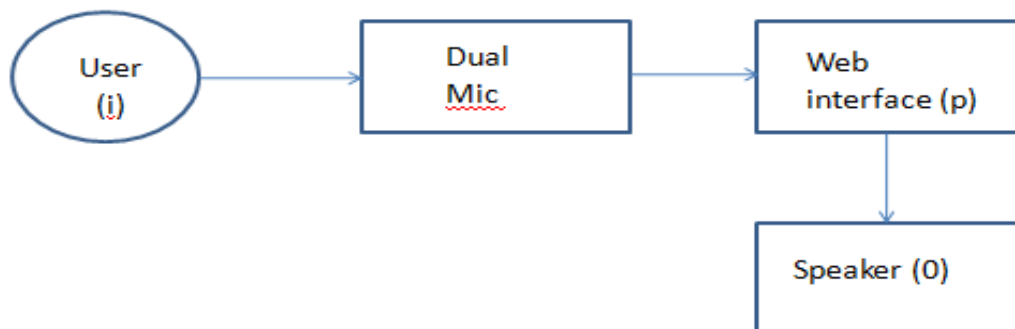

Fig 1: System Description

5. REFERENCES
[1] Seamless Human-Device Interaction in the Internet of Things. Eugenio Rubio-Drosdov, Daniel Díaz-Sánchez, Senior Member, IEEE, Florina Almenárez, Member, IEEE, Patricia Arias-Cabarcos, Member, IEEE, and Andrés Marín, Member, IEEE.

[2] E. Rubio-Drosdov, D. Diaz-Sanchez, P. Arias-Cabarcos, F. Almenarez, and A. Marin, "Towards a seamless human interaction in IoT," in Consumer Electronics (ISCE), 2015 IEEE International Symposium on, pp. 1-2,

IEEE, Jun. 2015.

[3] G. W. Furnas, T. K. Landauer, L. M. Gomez, and S. T. Dumais, "The vocabulary problem in human-system communication," Communications of the ACM, vol. 30, no. 11, pp. 964-971, Nov. 1987.

[4] T. K. Landauer, K. M. Galotti, and S. Hartwell, "Natural command names and initial learning: a study of text-editing terms,"Communications of the ACM, vol. 26, no. 7, pp. 495-503, Jul. 1983.

[5] M. R. Huq, N. T. T. Tuyen, Y. K. Lee, B. S. Jeong, and S. Lee, "Modeling an Ontology for Managing Contexts in Smart Meeting Space," in SWWS 2007: Proceedings of the 2007 International Conference on Semantic Web and Web Services, Las Vegas, Nevada, USA, Jun. 25-28, 2007, pp. 96-102.

[6] P. Bailis, J. Yang, V. J. Reddi, and Y. Zhu, "Research for practice: seb security and mobile web computing," in Communications of the ACM,vol. 60, no. 1, pp. 50-53, Jan. 2017.

[7] C. Treude, and M. A. Storey, "How tagging helps bridge the gap between social and technical aspects in software development," in Software Engineering (ICSE) 2009. IEEE 31st International Conference on, pp. 12-22, IEEE., May. 2009

[8] W. Jones, A. J. Phuwanartnurak, R. Gill, and H. Bruce, "Don´t take my folders away!: organizing personal information to get things done," in CHI´05 Extended Abstracts on Human Factors in Computing Systems, ACM, Portland, OR, USA, Apr. 02-07, 2005, pp. 1505-1508.

[9] C. D. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. J. Bethard, and D. McClosky, "The Stanford CoreNLP Natural Language Processing Toolkit," in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Baltimore, Maryland, USA, Jun. 22-27, 2014, pp. 55-60.

[10] E. Loper, and S. Bird, "NLTK: the natural language toolkit," in Proc. ETMTNLP'02, Philadelphia, Pennsylvania, USA, Jul. 2002, pp. 69-72.