VOLUME 2. INTERNATIONAL CONFERENCE ON APPLIED ICT, LEAD CITY UNIVERSITY, IBADAN. 56-67

An Algorithm for Resolving Natural Language Ambiguity in Human-Robot Interaction

Iroju Olaronke
Department of Computer Science, Adeyemi College of Education
Ondo, Nigeria
irojuolaronke@gmail.com

Ojerinde Oluwaseun
Department of Computer Science, Federal University of Technology
Minna, Nigeria
o.ojerinde@futminna.edu.ng

Ikono Rhoda
Department of Computer Science and Engineering, Obafemi Awolowo University,
Ile-Ife, Nigeria
rhoda_u@yahoo.com

Olaleke Janet O.
Department of Computer Science, Adeyemi College of Education
Ondo, Nigeria
shiollyjane@yahoo.com

***Abstract*—The use of social robots is fast becoming an essential part of human life. This is because social robots are becoming more ubiquitous and they are capable of performing complex tasks. For instance, social robots are now widely used to perform collaborative tasks with human beings in healthcare, education, entertainment, military as well as search and rescue missions where safety is a major concern. The interaction between human beings and social robots requires effective communication for the successful completion of a specific task. Nevertheless, effective, seamless and more flexible communication between human beings and social robots can be easily achieved through natural languages. This is because natural languages do not require complex graphical interfaces and environments, the knowledge of programming as well as extensive training. However, some words in natural languages are highly ambiguous, polysemous and their interpretations depend on the context in which they are used. Unfortunately, it is quite difficult for a social robot to capture the meaning of a word in the context in which it is used. This usually results in communication errors and thus a misinterpretation of the human intention which ultimately leads to the users' dissatisfaction. This reduces the acceptance rate of social robots in the society. Hence, this paper proposes an algorithm that explicitly specifies the semantics of words in natural languages between humans and social robots. This is with a view to resolving ambiguities and facilitating seamless communication and natural interaction between human beings and social robots.**

Keywords-human-robot interaction; natural language; ontology; social robots; spoken dialogue system

I. INTRODUCTION

In recent times, social robots are increasingly permeating the human society. Consequently, social robots are rapidly becoming ubiquitous in the human society. For instance, social robots are used as receptionists in hotels in Japan and Belgium [1].They are also used for providing assistive and rehabilitative care to patients within the hospital environments and they also serve as peers and tutors in educational centers [2]. Hence, social robots interact and work collaboratively with human beings to perform specific tasks. The interaction between human beings and social robots can be enhanced through natural language interaction. A natural language is a language that is either spoken or written by human beings for communication or it could be in gestural form such as sign language [3]. Typical examples of natural languages include English, German, Arabic, Chinese, Japanese, Spanish and French. Natural language interaction, on the other hand, is the process of communicating in natural languages. There are several benefits of natural language interaction between social robots and human beings. For instance, interacting in natural languages provides a means for social exchanges such as greetings, task-based dialogues for coordinating activities and topic-based discussions [4]. Natural language interaction between human beings and social robots increases the level of autonomy of social robots and also facilitates human-robot interaction [5]. It also makes humans and robots adequate and competent partners [6]. Furthermore, communicating with social robots in natural languages provides a flexible and intuitive means of issuing instructions to robots without complex graphical interfaces and environments, the knowledge of programming languages as well as extensive training [7]. This fact is in line with Doshi and Roy [8] who emphasized that both care-givers and care recipients rarely have experience with intelligent robots when used as a medical assistive technology. Natural language interaction provides an efficient means of making a technical system such as a social robot easily accessible to its users [9]. Hence, contemporary social robots now depend on interactions with humans in a natural way. However, natural languages are highly ambiguous, allow several interpretations and they are polysemous in nature. For instance, the same action can be described by different verbs. For example, the word leave can also mean go and depart. Furthermore, a single word can have different meanings. For example, the word plant can mean a photosynthetic organism, a manufacturing equipment or the process of sowing. It is however difficult for social robots to capture the semantics of words in the context in which they are used. Hence, this ambiguity usually results in communication errors and misunderstanding of the human's intentions. This ultimately leads to the misinterpretation of the human's intentions and users' dissatisfaction [8]. Hence, it is required that the robot response to a natural language is consistent with the human's expectations. Consequently, this study proposes a framework that resolves ambiguity during natural language interaction between humans and social robots. The natural language used for this study is English language.

This paper is organized into eight sections. Section 2 presents a short overview of social robots, section 3 discusses natural language interaction between human beings and social robots while section 4 is a review of related works. Section 5 presents the design as well as the implementation of the system. Section 6 is the distinguishing feature of the algorithm while the paper is concluded in section 7.

II. OVERVIEW OF SOCIAL ROBOT

Social robots have been defined by different authors in diverse ways. For instance, Duffy [10] referred to social robots as societal robots or agents that are capable of interactive and communicative behavior. In addition, Breazeal [11] viewed social robots as sociable robots that communicate, understand and relate with human beings in a personal way. According to Breazeal [11], social robots are socially intelligent in a human-like way. Fong et al. [12] viewed social robots as embodied agents that are part of a heterogeneous group of a society of robots or humans. Fong et al. [12] emphasized that social robots are capable of recognizing each other and engaging in social interactions. Fong et al. [12] also asserts that social robots possess histories, perceive and interpret the world in terms of their own experience and they explicitly communicate with and learn from each other. Fong et al. [12] is also of the opinion that social robots express and perceive emotions, communicate with high-level dialogue, learn and recognize models of other agents. Furthermore, Bartneck and Forlizzi [13] viewed social robots as robots that interact with humans by following their behavioral norms. In addition, Hegel [14] viewed a social robot as a robot with a social interface. A social interface according to Hegel [14] is a metaphor which includes all social attributes such as the design features by which an observer judges the robot as a social interaction partner. The concept of social robot by Hegel [14] is illustrated in Fig. 1. Hegel's emphasis in Fig. 1 is that the social interface of social robots contains the social forms, social functions, and social contexts of the social robot [14]. The social form addresses the shape of the robot which supports the social interaction between humans and robots, the social function addresses the social behavior of robots while the social context addresses the situations where there is a specific social interaction between a human being and a robot. The social function of a social robot encompasses verbal interaction [14].

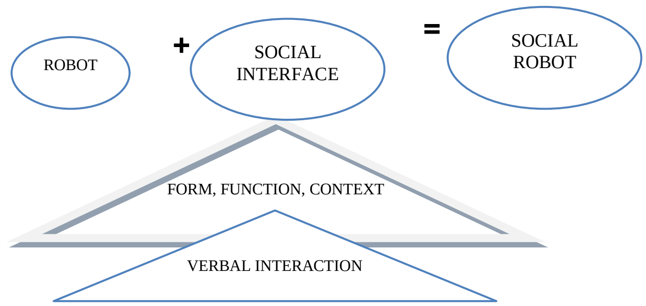

Figure 1. Hegel (2009) View of Social Robot [15]

However, from all the definitions given above by diverse authors, it can be deduced that social robots are autonomous in nature, they have the ability to establish and maintain social relationships with human beings and they have emotions. Social robots have shapes or forms; they can learn new behaviors or skills and they can communicate with human beings either by verbal interaction or non-verbal means.

III. NATURAL LANGUAGE INTERACTION BETWEEN HUMANS AND SOCIAL ROBOTS

Spoken language is one of the most intuitive forms of interaction between humans and social robots [8]. This is because social robots are used to perform collaborative tasks with people in their homes, workplaces, and outdoors [16]. According to Spiliotopoulos [17], spoken dialogue systems (SDSs) allow humans to interact with social robots through spoken dialogues in natural languages. The general architecture of spoken dialogue system is composed of six basic components as shown in Fig. 2.

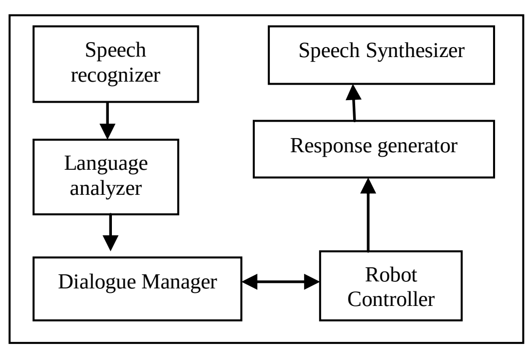

Figure. 2. The general architecture of a spoken dialogue system [17]

The components of a spoken dialogue system are discussed below:

A. Speech Recognition System

The speech recognition system is used to identify the sequence of words uttered by the human. The speech recognition system also processes the speech input and converts the speech into a written form.

B. Language Analyzer

The language analyzer identifies the syntactic structure of the sentences and generates a logical representation of the user's speech. The generated logical form of the user's utterances is used to analyze the structure of the sentences, how they are semantically related and also to identify a corresponding task to be executed by a robot [18]. However, the robot does not understand the semantic representation of the user's speech [19]. Hence, the semantic representation is mapped to the system's knowledge base concepts or ontology in order to interpret the meaning of the sentence.

C. Dialogue Management System

The dialogue management system is used to keep track of the interaction between the robot and its users. It also interprets the user's spoken utterances or non-verbal actions, and selects a communicative action [20]. There are three types of dialogue management techniques. These include the state based dialogue technique, the frame based dialogue technique and the plan based technique. The state based technique is the most commonly used and simplest dialogue management technique [21]. The state based technique represents each dialogue as a series of states. This technique usually requires specific information from the user at each state and also generates a response to the user [17]. Consequently, user's utterances are easier to predict thereby resulting in the development of more robust systems [17]. State based technique is however difficult to implement for complex tasks. The frame based technique however uses frames as an alternative to series of states. In the frame based technique, each frame represents a task or a subtask, and each frame has a slot that represents the information that the system needs to complete a task. According to Spiliotopoulos [17], plan based techniques identifies a user's plan and determines how that plan can be used to achieve a task.

D. Response Generator

The response generator is responsible for generating appropriate responses which are then passed on to the speech synthesizer.

E. Speech Synthesizer

The speech synthesizer is also known as the text-to-speech system. It is responsible for generating the verbal utterance that is made by the social robot.

F. Robot Controller

The robot's controller responds to the user's intent as perceived by the robot's sensors.

IV. RELATED WORKS

A lot of efforts have been made towards the design of spoken dialogue systems to facilitate natural interaction between human beings and social robots. However, most of these systems are associated with a lot of challenges. Nevertheless, this section provides a review of existing spoken dialogue systems. Spiliotopoulos [17] developed Hygeiorobot, a mobile robotic assistant for hospitals. The robot performs simple tasks in the hospital such as the delivery of messages or medicines to specific rooms. The robot interacts with the hospital staff via spoken dialogues in English Language. The spoken dialogue system allows the users to specify information that are necessary to deliver a medicine or message to a specific room or patient. The dialogue management system uses the state based technique because the development time for state based dialogue management system is shorter. One of the major limitations of the Hygeiorobot is that it is difficult to implement for complex tasks since it employs the state based technique for dialogue management. In addition, the Hygeiorobot lacks an ontology or knowledge base for specifying the meaning of the terms in the human utterances. Hence, the spoken dialogue system of Hygeiorobot does not resolve ambiguity in the spoken dialogues between the users and the robot.

She et al. [19] developed an approach which allows humans to teach a robot an action through natural language instructions. This approach consists of two basic components which include the Natural Language Processing (NLP) module and the Referential Grounding module. The NLP module consists of a semantic processor which extracts semantic information (the linguistic entities and their relations) from the users' utterances and represents it as a language graph. However, the semantic representation of the human utterance is not understandable by the robot at this point. Consequently, it is grounded to the robot's representation of perception and action which is represented as a vision graph in the referential grounding module. This approach is however very complex as it might be difficult to teach a new action. In addition, the robot may not perceive the environment perfectly; hence teaching and learning may not be successful.

Brooks et al. [22] developed a system that enhanced natural language interaction between a mobile robot and a team of humans in an urban search and rescue mission. The system consists of two means of communication. These include communication via natural language utterances and the map mode which provides information on objects or places that are significant to the robot's operation. When natural language instructions are entered into the system, the syntactic structures of the sentences are identified and semantic information is extracted from them. The system used the Bikel [23] parser and the null element restoration of Gabbard et al. [24] to parse the sentences, while Verbnet was used to extract verbs and their arguments in the parse tree. The system identifies possible matching frames for the verbs identified such that the match that expresses the most semantic role is selected. The system used the mutual knowledge based system developed by Clarks [24] for referential grounding. One of the limitations of the system is that the mutual knowledge based system contains information about the objects that the robot and the human are aware of. Hence, an error response is generated when an object cannot be found in the knowledge base. Kemke [6] described a system for natural language interaction with robots. The system comprised of five basic components which include natural language input analysis, natural language frame interpreter, the action interpreter, clarification dialogue module, and the knowledge base consultation module. The system receives verbal utterances from the user. This input is processed with the Earley natural language parser. The parser generates a structural representation of the verbal input by extracting the noun phrases, verb phrases and prepositional phrases from the user's input. The parser also accesses a lexicon which stores words relevant to a domain and their synonyms. The natural language frame interpreter uses the structural representation of the verbal input to create a case-frame or semantic representation of the input which is grounded onto a knowledge base. The function of the action interpreter is to resolve ambiguities by checking the knowledge base to see if there is a matching word. The system forms an instruction from the case-frame and translates it into an action that the robot executes. The system is however not 100% precise in terms of disambiguation or ambiguity resolution. For instance, once an ambiguity could not be resolved, the system asks the user for clarification in the clarification dialogue module.

Summarily, it is observed that existing spoken dialogue systems only look for the meaning of the user's words in the ontology or knowledge base system that they deploy for semantic analysis; they did not take into consideration the interpretation of the spoken words in the context in which they are used. This can however lead to communication errors and a misinterpretation of the human intention which can eventually result to the users' dissatisfaction. Hence, the problem of ambiguity still exists in existing spoken dialogue systems. Consequently, this study proposes an algorithm that will resolve the problem of ambiguity in natural language interaction between human beings and social robots.

V. SYSTEM DESIGN AND IMPLEMENTATION

In this section, the proposed framework for disambiguating words during natural language interaction between human beings and social robots is designed. The proposed framework is as shown in Fig. 3. In addition, a chatbot that communicates with a human being based on the topic of sowing a seed in the soil was designed in this section. The system uses natural language understanding mechanisms to facilitate the flow of information between the chatbot and the user. Fig.4 shows the system's algorithm. The algorithm was implemented on Windows 7 using python programming language, an open source; object oriented programming language which is rich in natural language processing. Python was chosen for the system implementation because its syntax and semantics are transparent and it has a good string-handling functionality. Four commonly used python libraries were employed in the system design. These include sys, nltk, random and string.

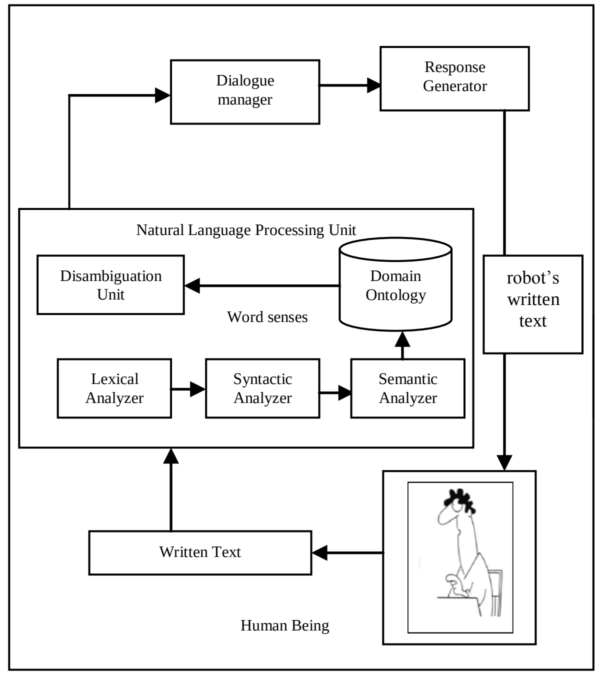

Figure. 3. The proposed framework

The sys library function represents system-specific parameters and functions. It provides access to some variables used by the interpreter. The nltk **is a** suite of libraries and programs used for symbolic and statistical natural language processing (NLP). The random library was used to generate random response given by the robot while the string library module contains functions that allows standard strings in python to be processed.

The proposed framework is made of four components. These include the communication unit, the Natural Language Processing unit, the dialogue manager and the response generator. The communication unit facilitates conversation between the robot and the human by allowing the human to specify information. The system allows the user to communicate with the robot by specifying his request in written text in English language. In this instance, the user specifies the word "sow the seed in the soil". Any information that is not related to the topic is discarded by the robot, and the response "I don't understand you" is generated by the system as depicted in Fig. 4. The basic goal of the natural language processing unit is to process the information exchanged between the human being and the social robot. The natural language processing unit consists of four components. These include the lexical analyzer, the syntactic analyzer, the semantic analyzer and the disambiguation unit. The lexical analyzer reads the input text character by character and produces tokens which are the basic lexical units of the human language. This process is known as tokenization. The English stopword corpus was used to remove all English stopword such as full stop and comma from the tokenized words. This study made use of Porter stemmer to reduce the tokenized words to their stem in order to reduce the processing time.

```
import random
import sys
import string
import nltk
//This part of the algorithm specifies the keyword that must be contained in the users' input
key_word=["*. sow, seed, soil  "]
//This section generates a random response
response=["ok", "alright"]
random_response=random.choice(reponse)
//This part of the algorithm ensures that only words with the keywords
//are recognized and preprocessed
while True:
userinput=raw_input("      ")
print(userinput)
if word. userinput in keyword:
//This part of the algorithm preprocesses the user's input
text_tokenize = nltk.word_tokenize(userinput)
from nltk.corpus import stopwords
stopwords.words('english')
filtered_words = [w for w in text_tokenize if not w in stopwords.words('english')]
print filtered_words
porter_word = nltk.PorterStemmer()
stem_word=[porter_word.stem(t) for t in filtered_words]
print stem_word
pos_word= nltk.pos_tag(stem_word)
print pos_word
//This section defines a grammar for the sentence
grammar = "NP: {<NN><NN>}"
//NPC is noun phrase chuncking
NPC = cp.parse(pos_word)
print NPC
K=NPC
//K is the chunked word
// This section finds the meaning or synsets of the //parsed words in Wordnet
From nltk.corpus import wordnet
syns=wordnet.synset ("K")
print (syns[0]. definition())
//This section disambiguates the word
//Y is the disambiguated word
Y= syns ∩ userinput
if Y ≥ 1   then
print 'word is related to userinput'
else
print 'word is not related to userinput'
else
print 'Please, I don't understand you'
```

Figure 4. An algorithm for resolving natural language ambiguity in human-robot interaction

Part of speech tags were also assigned to the stemmed words. The rule based noun phrase chunking was used for assigning tags to the stemmed word**.** During syntactic analysis, a noun phrase chunker was created by defining a chunk grammar, which consists of rules that indicate how the sentences in the human utterances were chunked. The part-of-speech tags provided information for the noun phrase chunks**.** The actual grammar used by the dialogue component was created using a noun phrase chunker. The rule of the chunker states that a noun phrase chunk should be formed whenever the chunker finds a noun or noun phrase that is composed of two consecutive nouns.

Grammar → <NN> <NN>
where
NN→ Noun

Fig. 5 shows the result of the tokenization, stop word removal, stemming and part of speech tagging.

```
Python 2.6.2 (r262:71605, Apr 14 2009, 22:40:02) [MSC v.1500 32 bit (Intel)] on win32
Type "copyright", "credits" or "license()" for more information.


****************************************************************
   Personal firewall software may warn about the connection IDLE
   makes to its subprocess using this computer's internal loopback
   interface.  This connection is not visible on any external
   interface and no data is sent to or received from the Internet.
****************************************************************

IDLE 2.6.2
>>> ================ RESTART ============
>>>
Enter Your Request
sow the seed in the soil.

Tokenization Process On going
['sow', 'the', 'seed', 'in', 'the', 'soil', '.']


Stopword Removal
Result of Stopword Removal for text
['sow', 'seed', 'soil', '.']


Result of Stemming
['sow', 'seed', 'soil', '.']


Part of Speech Tagging

Result of POS
[('sow', 'NN'), ('seed', 'VBD'), ('soil', 'NN'), ('.', '.')]
Result of Rule Based Noun Phrase Chunking
(S sow/NN seed/VBD soil/NN ./.)
>>>
```

Figure. 5. Result of tokenization, stop word removal, stemming, part of speech tagging and noun phrase chunking

The syntactic analyzer takes the rules of the grammar, compares it against the input sentence and generates a parse tree as shown in Fig. 6.

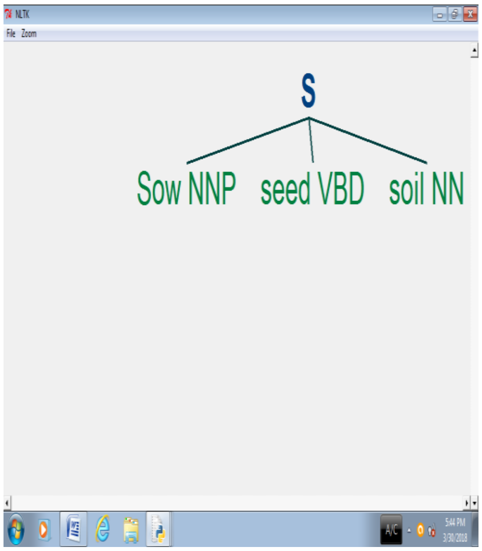

Fig. 6. The parse tree generated from the system

The semantic analyzer assigned meaning to the structure created by the syntactic analyzer using Wordnet as the domain ontology. Wordnet synonym sets (synset) was used to capture the meaning of the terms in the parse tree. For instance, the term seed have the following meanings in wordnet as shown in Fig. 7. Seed#1: seeded player; Seed #2 : source; Seed#3 : semen and Seed#4 : sow. Hence, the need to disambiguate the word so that it will be interpreted in the context in which it is used in the sentence.

The framework adapts the Kilgarriff and Rosensweig or Simplified Lesk algorithm to disambiguate the meanings of the words in the sentences. The Kilgarriff and Rosensweig algorithm measures the overlap between the meanings of a word and its context in the text. The algorithm also identifies the correct meaning for a word in the sentence at a time. In Fig. 4, the system finds if there is an overlap between the meanings of the chunked words and the sentence uttered by the human and afterwards the system returns the number of the overlap Y. A relationship is said to exist between the meanings of the words and the sentence if the number of the overlap is greater than or equals to 1. At this point, the meaning of the word is mapped to the sentence and this shows the context in which the word is used. This is to facilitate seamless communication and natural interaction between human beings and social robots. For instance to disambiguate the word "seed" using the disambiguation unit, the synsets of seed in wordnet are considered and the meaning of word seed with the highest overlap is chosen as the context in which the word is used. In Fig. 7, there is an overlap between Synset ('sow.v.01') and the user's input "sow the seed in the soil". The communication dialog manages the interaction between the human and the social robot while the response generator generates a randon response by the robot to the human in a written form. This is as depicted in Fig. 4.

```
Python 2.6.2 (r262:71605, Apr 14 2009, 22:40:02) [MSC
v.1500 32 bit (Intel)] on win32
Type "copyright", "credits" or "license()" for more
information.


****************************************************************
****************
   Personal firewall software may warn about the
connection IDLE
   makes to its subprocess using this computer's internal
loopback
    interface.  This connection is not visible on any external
       interface and no data is sent to or received from the
                               Internet.

****************************************************************
                        ****************

                               IDLE 2.6.2
>>> ============RESTART ============
>>>
[Synset('seed.n.01'), Synset('seed.n.02'),
Synset('seeded_player.n.01'), Synset('source.n.03'),
Synset('semen.n.01'), Synset('seed.v.01'),
Synset('seed.v.02'), Synset('seed.v.03'), Synset('sow.v.01'),
Synset('seed.v.05'), Synset('seed.v.06'), Synset('seed.v.07'),
Synset('seed.v.08')]

Synset('sow.v.01') is related to the user's request

>>>
```

Figure 7. Synsets of seed in wordnet

VI. Distinguishing Feature of the Proposed Framework

Words in natural languages are usually polysemous in nature and as a result it becomes a challenge for Information and Communication Technology (ICT) tools such as computers and technical systems like robots to capture the semantics of words because of the high level of inconsistencies or discrepancies in the meaning and the interpretation of the same term. Consequently, it becomes pertinent to introduce a disambiguation unit into the proposed framework. Hence, one of the distinctive features that distinguish the proposed framework from existing systems is that the proposed framework takes into consideration the disambiguation of words in the context in which they are used in the sentences uttered by the human user. However, the existing systems only look for the meaning of the words in the background knowledge or ontology without considering that a word might have diverse meanings in the ontology.

VII. CONCLUSION

This study takes a critical look at how ambiguity can be resolved when human beings are communicating with social robots with a Natural Language. This is with a view to avoiding communication errors and misinterpretation of the same term that is associated with several meanings and interpretations. Therefore, the misconception of a human's intention during human-robot interaction will be circumvented. Hence, the study proposes an algorithm that facilitates the effective, seamless and flexible communication between human beings and social robots by providing the interpretation of a term in the context in which it is used by a human being during human-robot interaction. This will ultimately increase the usability and the acceptance rate of social robots as they are rapidly becoming an essential part of human life.

REFERENCES

[1] Tractica White Paper, "Utilizing humanoid robots for customer engagement: benefits and challenges, use cases and industry, verticals, and business considerations," *Softbank Robotics*, pp. 1-31, 2016.

[2] O. Mubin, C.J. Stevens, S. Shahid. A.I. Mahmud, and J.J.Dong, "A review of the applicability of robots in education," Journal of Technology in Education and Learning, pp. 1-7, 2013.

[3] O.G. Iroju, and J.O. Olaleke "A Systematic Review of Natural Language Processing in Healthcare", International Journal of Information Technology and Computer Science, pp. 44-45, 2015.

[4] S. Domingo, "Natural language processing for human-robot interaction" Universidad Politécnica de Madrid, 2015.

[5] T. Kollar, S. Tellex, D. Roy, and N. Roy, "Grounding verbs of motion in natural language commands to robots," 12th International Symposium on Experimental Robotics, pp. 31-47, 2014.

[6] C. Kemke. "From Saying to Doing – Natural language interaction with artificial agents and robots," in Human-Robot Interaction, Itech Education and Publishing, Vienna, Austria, 2007.

[7] F. Duvallet, "Natural language direction following for robots in unstructured unknown environments," Unpublished PhD Thesis, The Robotics Institute School of Computer Science Carnegie Mellon University Pittsburgh, 2015.

[8] F. Doshi, and N. Roy, "Spoken language interaction with model uncertainty: An adaptive human-robot interaction system," Connection Science, pp. 1-21, 2008

[9] T. Laengle, T. C. Lueth, E.Stopp, G. Herzog and G. Kamstrup, "KANTRA - A natural language interface for intelligent robots, Proceedings of the 4th International Conference on Intelligent Autonomous Systems," pp. 1-14, 1995.

[10] B.R. Duffy, *The social robot,* Ph.D Thesis, Department of Computer Science, University College Dublin.

[11] C. Breazeal, *Designing Sociable Robots*, Cambridge: MIT Press, 2002.

[12] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," Robotics and Autonomous Systems, pp. 143-166, 2003.

[13] C. Bartneck, J. Forlizzi, "A design-centered framework for social human-robot interaction," Ro-Man, pp. 591–594, 2004.

[14] F. Hegel, C. Muhl, B. Wrede, M. Hielscher-Fastabend, and G. Sagerer, G, "Understanding social robots," Advances in Computer-Human Interactions, pp. 169-174, 2009.

[15] Kruijff-Korbayová, I., Natural language interaction for social robots, Adaptive Strategies or Sustainable Long Term Social Interaction, 2013.

[16] F. Duvallet, Natural language direction following for robots in unstructured unknown environments, Unpublished PhD Thesis, The Robotics Institute School of Computer Science Carnegie Mellon University Pittsburgh.

[17] D. Spiliotopoulos, I. Androutsopoulos, and C. D. Spyropoulos, "Human-robot interaction based on spoken natural language dialogue", Intelligent Robots and Systems, 2000.

[18] A. Perzylo, S. Griffiths, R. Lafrenz and A. Knoll, "Generating grammars for natural language understanding from knowledge about actions and objects", *Robotics and Biomimetics*, 2015.

[19] L. She, Y. Cheng, J. Y. Chai, Y. Jia,S. Yang, and N. Xi, "Teaching robots new actions through natural language instructions", The 23rd IEEE International Symposium on Robot and Human Interactive Communication, pp.25-29, 2014.

[20] T. Belpaeme, P. Baxter, R. Read, R. Wood, H. Cuay´ahuitl et al., "Multimodal child-robot interaction: Building social bonds," *Journal of Human-Robot Interaction*, pp.33-53, 2012.

[21] M. McTear, Modeling spoken dialogues with state transition diagrams: Experiences of the CSLU toolkit, *Proceedings of the International Conference on Spoken Language Processing*, pp. 1223-1226, 1998.

[22] D.J. Brooks, C. Lignos, C. Finucane, M. S. Medvedev, I. Perera et al., "Make it so: Continuous, flexible natural language interaction with an autonomous robot," *Association for the Advancement of Artificial Intelligence*, 2012.

[23] D. Bikel, "Intricacies of Collins' parsing model," *Computational Linguistics*, pp. 479–51, 2004.

[24] R. Gabbard, M. Marcus, and S. Kulick, "Fully parsing the penn treebank., Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics,184–191, 2006.