IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 15, 2018, accepted February 13, 2018, date of publication February 21, 2018, date of current version March 19, 2018.

*Digital Object Identifier 10.1109/ACCESS.2018.2808369*

# A Review of Service Robots Coping With Uncertain Information in Natural Language Instructions

**M. A. VIRAJ J. MUTHUGALA, (Student Member, IEEE), AND A. G. BUDDHIKA P. JAYASEKARA, (Senior Member, IEEE)**

Intelligent Service Robotics Group, Department of Electrical Engineering, University of Moratuwa, Moratuwa 10400, Sri Lanka

Corresponding author: M. A. Viraj J. Muthugala (viraj@elect.mrt.ac.lk)

This work was supported by the University of Moratuwa under Senate Research Grant SRC/CAP/2017/03.

**ABSTRACT** Intelligent service robots are currently being developed to cater to demands in emerging areas of robotic applications, ranging from entertainment to health care. These service robots are intended to be operated by nonexpert users, and their service tasks involve direct interaction between these robots and their human users. Thus, human-friendly interactive features are generally preferred for such service robots. Humans prefer to use voice instructions, responses, and suggestions to convey ideas to their peers. However, information conveyed through natural language communication is imprecise because it tends to contain uncertain/qualitative information instead of precise quantitative information. Therefore, the ability to cope with uncertain information in natural language instructions is mandatory for human-friendly service robots. This paper presents a review of service robots and systems that can cope with uncertain information in natural language instructions. The available literature has been investigated and analyzed to identify the limitations of the existing methods and possible improvements. The identified limitations and possible improvements are presented as the outcomes of the review.

**INDEX TERMS** Uncertain information understanding, human–robot interaction, human-friendly robotics, service robotics.

## I. INTRODUCTION

An intelligent service robot is a machine that is able to gather information from its environment and use its knowledge to operate safely in a meaningful and purposive manner [1]. Recent developments in intelligent service robotics have opened up new areas of robotic applications, such as health care [2], [3], rehabilitation [4], [5], caretaking [6], [7], assistance [8], [9], education [10], [11] and entertainment [12], [13]. In particular, intelligent service robots are being developed to serve as assistive aids for elderly or disabled people [14]–[17] to address the widening gap between the supply of and demand for human caregivers, which has profound socioeconomic implications [18], [19].

The intelligent service robots used in these emerging areas of robotic applications are anticipated to interact directly with human users in domestic environments, with most users belonging to the non-expert category. Hence, human-friendly interactions between these service robots and their human users are preferred in order to provide sophisticated service [20]–[23]. Human-friendly robots should possess human-like interaction capabilities, and the ability to realize the dream of a perfect service robot obviously depends on such capabilities. The feasibility of human-human-like communication in human-robot interactions will enhance the overall interaction between a human user and his or her robot partner, which will ultimately increase user satisfaction [20], [24]. However, the development of human-friendly interactive features for service robots is complicated because it requires the incorporation of the social-cognitive features of human beings into robots [20].

Voice communication is one of the main communication modalities used by humans to convey instructions to their peers [25]. Therefore, endowing robots with human-like voice communication capabilities will enhance the overall interaction between the robots and their users. This will ultimately increase the rapport between users and their robotic assistants, allowing users to receive more sophisticated service from their robot companions [21], [26]. However, precise

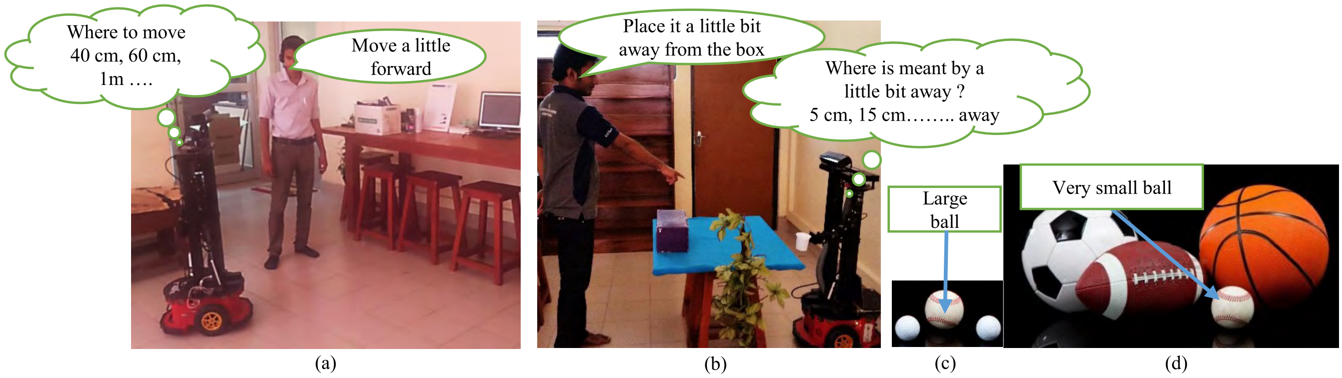


**FIGURE 1.** Example scenarios in which uncertain information is used for a purposive task and how the meanings change in different situations. (a) shows a situation in which a robot is commanded to move a little bit forward by a human user. (b) shows a situation in which a robot is asked to place a cup a little bit away from a box on a table. (c) shows a scenario in which a baseball is surrounded by two golf balls. (d) shows a situation in which the same baseball is surrounded by a football, a soccer ball and a basketball.

quantitative information is not typically conveyed through voice communication, and naturally spoken instructions and responses often contain uncertain information, lexical symbols and notions that need to be interpreted to achieve clear comprehension. For example, a human user will generally prefer to issue the command "move a little bit forward" rather than the command "move 40 centimeters forward" in a situation similar to the scenario shown in Fig. 1(a). The quantitative meaning of the term "little" is not clearly defined and depends on various factors. In this case, the quantitative distance meant by the user may be on the order of 40–80 cm. By contrast, in the situation shown in Fig. 1(b), the robot is commanded to "place it a little bit away from the box" when placing a cup on a table. The quantitative meaning of the term "little" in this situation is approximately on the order of 5–15 cm, clearly different from that in the earlier situation. Furthermore, the scenarios depicted in Fig. 1(c) and Fig. 1(d) show that the same quantitative size could be referenced using completely opposite natural language descriptors in two different situations. Fig. 1(c) shows a situation in which a baseball is surrounded by two golf balls; in this situation, the baseball will most probably be referred to as being "large" in size in a natural language instruction issued by a person. By contrast, in the scenario shown in Fig. 1(d), the same baseball is surrounded by a soccer ball, a football and a basketball. Even though the quantitative size of the baseball is the same in both situations, a person will likely refer to the baseball as being "very small" in the second scenario. Humans have the ability to interpret reasonable quantitative values for such uncertain terms. These uncertain terms are also referred to as fuzzy linguistic information or qualitative terms. Although the quantitative meanings of uncertain terms such as "little," "far," "high" and "large" depend on various factors, humans unthinkingly use such terms in voice instructions, suggestions and responses because of our remarkable cognitive capacity to understand the quantitative meanings of such terms based on the factors that affect those meanings. Therefore, it is mandatory to endow a service robot with a similar cognitive ability to understand and appropriately respond to uncertain information in voice commands and responses in order to provide human-friendly assistance to users.

To this end, this paper presents a review of existing service robots and systems that can cope with the uncertain information contained in natural language instructions and responses in the form of terms such as "little," "far," "high," "near" and "far."

### A. REVIEW PROTOCOL

The definitions of terms commonly used in this paper are given below to provide a concise and clear background for the reader.

- **Uncertain information:** Lexical symbols in natural language instructions that do not have definite quantitative meanings. Examples include "little," "far," "high," "large" and "near." Uncertain information may also be referred to as fuzzy predicates, fuzzy linguistic information, qualitative terms or uncertain terms. Such uncertain terms are often used in relation to spatial information, the size/length of an item, the properties of an object, etc.
- **Environment (related to adaptation entities):** The spatial properties of the surroundings of a robot that can be quantitatively evaluated, such as distances between objects, the free space in the room, and the sizes of the room and the objects in it. However, properties of objects that cannot be measured numerically (e.g., the danger of placing gasoline near an open flame, the fact that a water glass is typically placed near a lunch dish) are excluded.
- **Experience (related to adaptation entities):** Knowledge of any previous action or state and any information acquired through interaction with the environment and users except contextual knowledge (e.g., previously moved distances, knowledge acquired through user feedback).

- **Context (related to adaptation entities):** Specific knowledge about the properties of objects, tasks or situations that may not be numerically measured (e.g., the danger of placing a can of gasoline near an open flame, the easily breakable nature of glass).

The literature to be reviewed in this survey was selected by exploring major indexing databases such as IEEE Xplore, SCOPUS, ScienceDirect and Google Scholar. Manuscripts published in peer-reviewed journals or conference proceedings were considered for this review, whereas unpublished or non-peer-reviewed manuscripts such as technical reports, theses and dissertations, news items and web articles were excluded, with the exception of books and chapters that could be used to provide supporting statements or additional information such as definitions. If the same core concept was reported in two or more documents, such as a conference paper and an extended journal article, the major focus was placed on the journal article. However, if a journal article was from a publisher other than IEEE, images were obtained from corresponding conference articles published by IEEE, if available. Furthermore, only manuscripts published in English were considered. The scope of this review is limited to the literature that addresses the understanding of uncertain information in natural language voice instructions by service robots and systems. The literature related to the understanding of natural language representations in general and the uncertainties related to difficulties in voice recognition was excluded from this survey. However, a brief discussion of voice and natural language communication in human-robot interaction (HRI) is given in section II to provide the reader with insight into the understanding of uncertain information in natural language voice instructions. The current status of robots and systems that can cope with uncertain information in voice instructions is presented in section III. The limitations of the existing systems and possible improvements are discussed taxonomically in section IV. Finally, the investigation is concluded in section V.

## II. VOICE AND NATURAL LANGUAGE COMMUNICATION IN HUMAN–ROBOT INTERACTION

With the development of voice recognition and voice synthesis engines, studies are being conducted in pursuit of enhanced voice communication interfaces for robotic systems [27]–[29]. However, many of the early studies in this area primarily focused on the implementation of interfaces for voice communication between robots and humans, and these studies were limited to the basic control of a robotic system through a limited number of user instructions, such as the control of automated wheelchairs [30]–[32]. Such systems are only capable of understanding simple, single-word commands such as "go" and "stop" that are prerecorded in the memory of the system.

For a general-purpose service robot, the ability to handle only a limited number of simple instructions is not sufficient since the number of functionalities of such a robotic system is much higher [24], [33], [34]. For instance, a service robot that works in the role of a nurse must be capable of conveying empathy to a patient when communicating sensitive information related to his or her condition, whereas a domestic service robot that serves as a customer-handling agent at a service desk should be capable of adjusting its speech based on customer characteristics. In addition, such a limited system would not facilitate human-like service. Therefore, human-like voice communication abilities are desirable for service robots, especially for achieving human-like human-robot communication. To this end, service robots with human-like voice communication capabilities have been developed, and such robots are capable of obeying natural language user instructions and responding with natural language dialog phrases [25], [35]–[37].

Natural language voice instructions, responses and suggestions often include lexical symbols and notions, uncertain terms, redundant words and prepositions. Therefore, robotic systems with human-like voice communication abilities should possess the ability to understand such terms appropriately. Methodologies have been developed for computing the spatial relations conveyed by prepositions such as "behind," "at" and "near" [38]–[41]. The methods proposed in [38] and [39] are capable of distinguishing between the meanings of "at" and "near," the methods proposed in [40] and [41] are capable of grounding spatial relationships in human-robot interactions, and the method proposed in [42] can create an abstract map of the working environment based on a semantic description containing prepositions. Natural language voice instructions are often inaccurate or ambiguous, and the exact meanings of such commands depend on the context of interest. For example, consider the expression "the red ball on the table near the vase." There are two alternative interpretations of the expression: the red ball may be near the vase, or the table may be near the vase. The correct interpretation among the alternatives depends on the actual arrangement of the environment. The method proposed in [43] is capable of correctly understanding such ambiguous or inaccurate commands by considering the arrangement of the environment. Methods have been developed to enhance the voice communication between robots and humans by integrating multimodal interaction capabilities; the method proposed in [44] is capable of identifying an object referred to in a user instruction with the aid of pointing gestures performed by the user, whereas the method proposed in [45] is capable of generating gestures to be performed by a robot for object-referencing communications, and the method proposed in [46] is capable of fusing information from these multiple modalities. Knowledge acquisition and symbol grounding through multimodal human-robot interactions have also been studied [47].

The methods described above enable interaction through natural language voice instructions and responses to some extent. However, these systems still lack the ability to understand the uncertain information contained in natural language instructions, and methodologies for coping with that uncertain information have not been covered in the scope of those works. Uncertain information is often unthinkingly included in voice instructions, responses and suggestions during natural interactions, as explained in section I. The core contribution of this study is to investigate the methodologies used in robotic systems to understand the uncertain terms contained in natural language voice communication. To this end, a comprehensive exploration of such systems is presented in section III.

## III. CURRENT STATUS: ROBOTS COPING WITH UNCERTAIN INFORMATION IN NATURAL LANGUAGE INSTRUCTIONS

### A. EARLY DEVELOPMENTS AND APPROACHES

There have been many psychophysical studies on the perception of distance and related cognitive issues [48]–[50]. These studies have revealed the characteristics of distance-related human cognition, such as knowledge of relative locations, the asymmetry of cognitive distances and sources of distance knowledge. However, these concepts are limited to cognitive science, and the results of these studies have mainly been applied for understanding concepts such as cognitive distances in urban environments.

Dutta [51] proposed a concept for the representation of the spatial constraints among a set of objects given imprecise, incomplete and possibly conflicting information regarding them. Furthermore, Clementini *et al.* [52] developed a qualitative model for representing the positions of objects and performing spatial reasoning as a qualitative replacement for quantitative vector algebra. However, these concepts do not directly address the interpretation of uncertain information; instead, they are mostly limited to understanding simple qualitative representations, such as if object A is behind B, then B is in front of A. In addition, these concepts have not been implemented in real systems and have instead remained limited to mathematical modeling.

### B. SYSTEMS WITH PREDETERMINED OR FIXED INTERPRETATIONS

A method of communication between robots and humans using spatial language has been developed [53]. This system is capable of generating linguistic spatial descriptions of the surrounding environment. For example, it can generate the following dialog: "The box is behind me. The object is far." Such dialogs include distance-related uncertain terms such as "close" and "far" and direction-related uncertain terms. The system can perceive the environment through range sensors and generate spatial descriptions related to distance and direction by using the method proposed in [54]. The direction descriptors are generated by categorizing the space around the robot into 16 subdirections. The distance descriptors are generated based on the distance of the object from the robot, with the distance categorization illustrated in Fig. 2. Therefore, the direction and distance categories are fixed with hard boundaries, and the system does not possess the ability to consider fuzziness typically inherited by

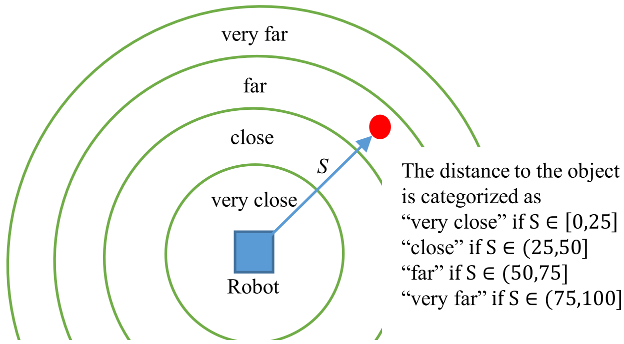


**FIGURE 2. The distance categorization performed in the system proposed in [53] and [54] to generate linguistic spatial descriptors of surrounding objects.**

linguistic descriptors. Therefore, this assumption will eventually degrade the performance of the system.

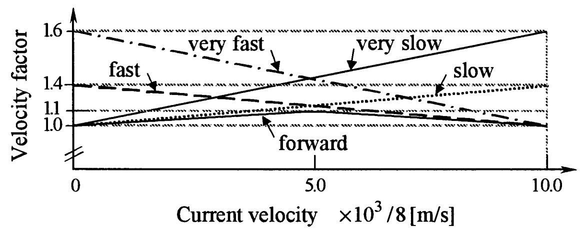


**FIGURE 3. The linear modification factors used to obtain the desired velocity when interpreting uncertain velocity instructions in the work proposed in [55]. Reprinted with permission ©IEEE 2004.**

Methodologies for controlling a robot using information-rich natural spoken user utterances have been studied with the intention of handling natural language voice instructions with fuzzy implications related to the velocity of the robot while ignoring the redundant words in natural language expressions [55], [56]. For instance, consider the command "Robot, please go very fast." In this example command, the words "Robot" and "please" have no meaning within the operational domain of the robot; only the words "go" and "very fast" are associated with the robot's functions. The proposed methodologies are capable of ignoring redundant words and interpreting the fuzzy implications of natural language voice commands to enable appropriate responses. However, they are not capable of identifying the contextual grammar, and hence, these systems cannot differentiate between the commands "Robot, go very fast" and "Robot, do not go very fast." Crisp output values for fuzzy implications such as "very fast" are generated by a fuzzy neural network. The output generated based on fuzzy linguistic information is defined by a linear modification factor based on the current state of the robot (i.e., the current velocity of the robot), as shown in Fig. 3. The desired velocity is calculated as $Desired\_Velocity = Velocity\_Factor \times Current\_Velocity$ using the corresponding velocity factor at the current speed as obtained from the graph presented in Fig. 3. The linear modification factors are defined based on the argument that phrases such as "very fast" have low significance when the machine is close to the maximum velocity, whereas the opposite is true for phrases such as "very slow." Furthermore, these linear modification factors are fixed. Hence, the output of the system for a particular state is predetermined.

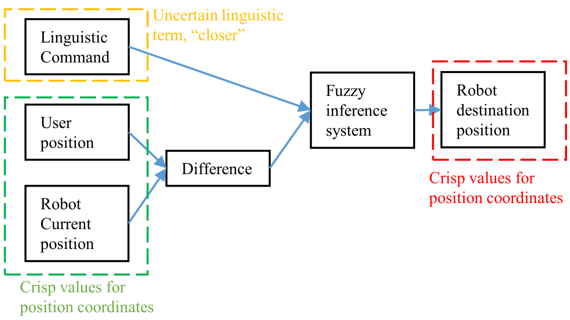


**FIGURE 4.** **Flow chart of the fuzzy command interpreter proposed in [57]. This figure is based on [57].**

A robotic aid system with a fuzzy command interpreter has been developed for feeding the physically handicapped [57]. This system is capable of assigning crisp values to fuzzy linguistic terms in user commands in accordance with the current context. A fuzzy inference system is utilized, which generates a crisp output by evaluating the difference between the robot's position and the user's position as explained in Fig. 4. The difference between the positions of the user and the robot is calculated from their coordinates with respect to a reference frame. The calculated position difference and the uncertain descriptor to be interpreted are fed into the fuzzy inference system to generate the crisp coordinates of the destination position. The system is designed based on the insight that when the distance between the robot and the user is large, the command "move closer" should cause the robot to move a large distance toward the user to reach a closer position, whereas if the distance difference is small, then the movement should cover only a small distance because the robot is already close to the user. This system provides users with a much friendlier interface through which to instruct the robot. Although the system evaluates the current context, the output is predetermined because the membership functions of the fuzzy inference system are defined as fixed entities. Furthermore, the details of the fuzzy inference system used in the proposed interpreter are not revealed.

### C. ROBOTIC SYSTEMS THAT ADAPT THEIR PERCEPTIONS IN ACCORDANCE WITH THE ENVIRONMENT

Uncertain information related to spatial information such as the sizes of objects and distances is often used in typical assistive tasks in domestic environments. The meanings of such uncertain terms obviously depend on various environmental factors. Therefore, concepts have been introduced for adapting the perceptions of robots regarding uncertain information based on spatial information from the environment of interest.

A method has been introduced for effectively evaluating fuzzy linguistic information in manipulation-related user

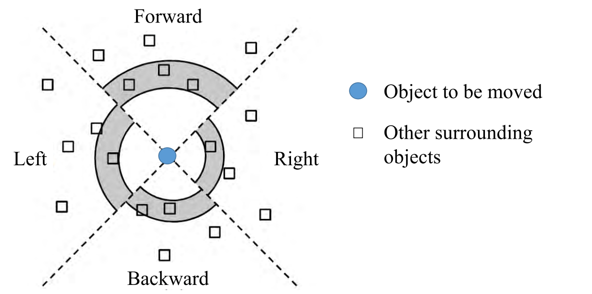


**FIGURE 5.** **Illustration of the parameter evaluation performed in the visual attention system proposed in [57]. The neighborhood region in each principal direction is indicated by the shaded area. Only the objects in these neighborhood regions are considered for evaluating the average distance; other surrounding objects are omitted. This figure is based on [58].**

instructions such as "move the red box a little to the left" based on visual attention [58]. The system is capable of assigning a quantitative distance value to the distance-related fuzzy implication in a particular user command. The target object of a user instruction is identified by directly mapping the lexical symbol to the object memory, which is done by treating the Hu's moments [59] and RGB values as the feature set, in a manner similar to the method explained in [60]. However, the system does not handle the uncertainties related to the properties of objects or uncertain information related to the movement direction. Furthermore, the possible user instructions are limited by a strict grammar model that does not permit the system to learn new patterns for user instructions. Through visual attention, the system can perceive its working environment in order to assess the spatial arrangement of the objects therein. A fuzzy inference system is used to generate a crisp distance value for a fuzzy implication by considering the average distance to the objects in the attended visual field. To calculate the average distance ($d_{avg}$), the attended visual field is divided into regions based on four principle directions, as shown in Fig. 5. Subsequently, the average distances to the surrounding objects in each neighborhood are calculated. Thereafter, $d_{avg}$ is calculated by assigning a higher priority to the region in the target moving direction than to the regions in the other directions. Then, the parameters obtained from the visual attention system are fed into the fuzzy inference system as shown in Fig. 6. To evaluate the performance of this system, experiments were conducted to investigate how the evaluated distances for different fuzzy implications vary with the arrangement of the objects in the visual field. The results of these experiments clearly indicate the ability of the system to adapt its perceptions in accordance with the spatial object arrangement. However, these results were not validated with respect to user compliance. Furthermore, the system is capable only of interpreting fuzzy implications related to motional information; it cannot evaluate uncertain information related to positional information. For example, it cannot evaluate the command "move blue box near to the red box" since it cannot evaluate the positional information related to the uncertain

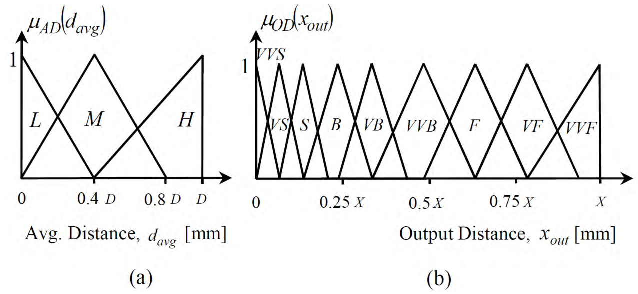


**FIGURE 6. The membership functions of the fuzzy inference system used in the system proposed in [58]. (a) shows the input membership function for the average distance to the surrounding objects; the fuzzy sets are adjusted with respect to *D*, which is the distance to the farthest object. (b) shows the output membership function of the system; the fuzzy sets in the output membership function are adjusted with respect to *X*, which is the distance to the nearest object in the target moving direction. The fuzzy predicates that can be identified by this system (i.e., "very little," "little," "medium" and "far," according to the grammar model) are fed into the system through an input membership function with singleton fuzzy sets. This figure is reproduced from [61]. Reprinted with permission ©2009 IEEE.**

term "near." This is one of the major limitations of this system. In addition to that, the system uses an overhead camera to perceive its environment; hence, its attended visual field is not human-like.

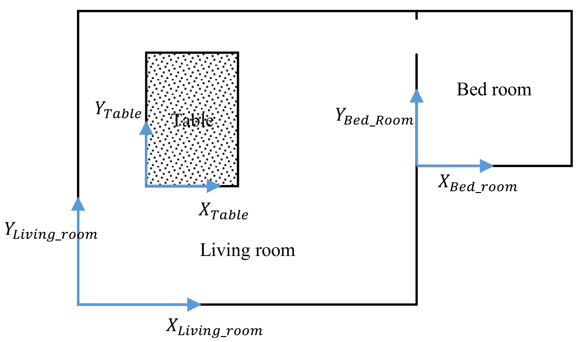


**FIGURE 7. Illustration of the concept of frames used in [62] and [63]. Each room or object is associated with its own reference frame.**

Schiffer *et al.* [62], [63] proposed a method that can be used by a service robot for qualitative spatial reasoning regarding the positional information contained in user instructions. The proposed concept has been combined with a logic programming language known as GOLOG [64] and a framework for reasoning about actions and changes known as situation calculus [63]. This enables reasoning on fuzzy fluents related to positional information by a robot operating in a domestic environment. The basis of the reasoning method is that the fuzzy information associated with positional information in a domestic environment depends on the associated frame or point of view. The assignment of frames in an example situation is illustrated in Fig. 7. As an example, "far" with respect to a large room such as a living room has a higher quantitative meaning than "far" with respect to a small room such as a bedroom, whereas "far" with respect to a table in the living room has a much smaller quantitative meaning than in either of the previous two cases. Therefore, the meanings of fuzzy terms are scaled in accordance with the frame size, which is the size of the corresponding room or the corresponding object, such as a table. Therefore, this method enables the adaptation of perceptions based on the environment. However, experimental results on the variations in the quantitative values interpreted from qualitative information have not been gathered and analyzed. The adaptation process entirely depends on the size of the frame, and other environmental factors that may influence interpretation, such as free space and object arrangement, are not considered. These are the main shortcomings of this work.

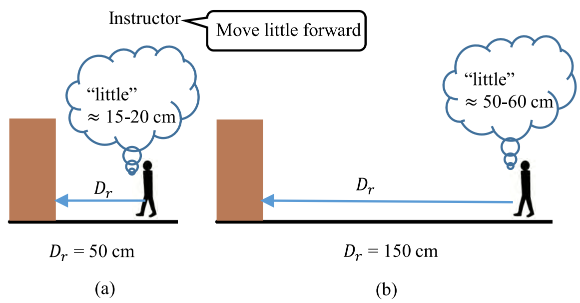


**FIGURE 8. Illustration of the main motivation for the work proposed in [65]. In both situations, the person is instructed to move a little forward. The person is standing 50 cm and 150 cm away from the wall in scenarios (a) and (b), respectively.**

According to the method proposed in [65], knowledge of the size of the room alone is not sufficient for effective interpretation of the uncertain information in navigational commands such as "move a little forward" since inside the same room, there may be different object arrangements that can affect the meaning of uncertain information, such as information related to the available free space. The authors' main argument is that the movement constraints imposed by the arrangement of the environment play a major role in modifying the meaning of distance-related uncertain information in navigational commands. This can be explained with the aid of the scenarios illustrated in Fig. 8. In both scenarios, the person is instructed to move a little forward inside the same room but starting from different initial positions. In the scenario shown in Fig. 8(a), the person is standing in front of a wall with a 50 cm gap between him and the wall. Therefore, the distance meant by the term "little" may be approximately 15-20 cm. By contrast, in the situation shown in Fig. 8(b), the person is standing 150 cm away from the wall; accordingly, the distance meant by "little" may be 50-60 cm. Furthermore, the availability of free space also influences mobility. Therefore, the proposed system utilizes environmental factors such as the room size, the available free space and the movement restrictions imposed by obstacles to adapt a robot's perception of uncertain information based on the spatial arrangement of the environment. The robot is equipped with sonar sensors with which to perceive its environment, and it calculates the required environmental factors from stored navigation maps. A Mamdani-type fuzzy

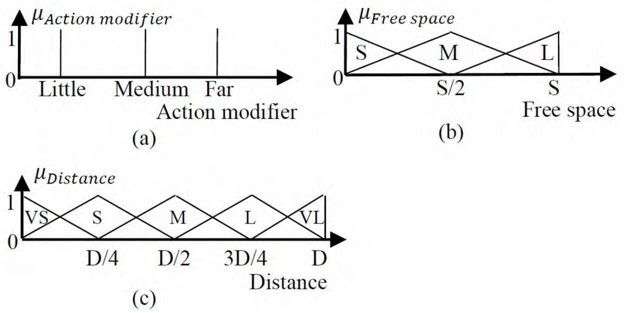


**FIGURE 9. The fuzzy membership functions of the system proposed in [65]. (a) shows the input membership function for the uncertain term in a particular command. (b) shows the input membership function for the available free space. (c) shows the output membership function, which is adjusted according to the perceptive distance *D*. This figure is based on [65].**

inference system [66] is used to generate quantitative outputs based on uncertain terms in user instructions by analyzing the environmental parameters mentioned above. The input and output membership functions of the fuzzy inference system are shown in Fig. 9. The input membership function for the free space is adjusted according to the room size ($S$). The output membership function is adjusted according to $D$, where $D = D_r$ ($D_r$ is the distance to the nearest movement obstruction in the target direction, as illustrated on Fig. 8.). Experimental results have been compared against user expectations to evaluate the performance of the proposed system. However, this system lacks the ability to handle uncertainties related to positional information, and the possible user instructions are limited by a strict grammar model that is not updated during operation. These are the main limitations of the proposed system.

The spatial information analysis ability of the system introduced above has been further improved by deploying a module called an occupied density analyzer [67], which analyzes the occupancy of surrounding objects rather than simply the distance to the nearest movement obstruction in the target direction. The perceptive distance ($D$) that modifies the output membership function (i.e., the membership function shown in Fig. 9(c)) is adapted from a mathematical function defined by analyzing the natural tendencies of human beings. Thus, the perceptive distance ($D$) is calculated as $D = f(D_r, occupied\_density\_variation)$ instead of as $D = D_r$ as in [65]. The natural human tendencies considered when formulating the mathematical model are summarized below. The definition of the occupied density as given in the paper is the area occupied by objects per unit area within a considered region, and it is calculated as shown in (1).

- Mobility decreases when moving toward an area in which the occupied density is high, and vice versa.
- The influence of objects in close proximity is higher than that of objects farther from the intended movement path.
- Surrounding objects in different distance fields exert different degrees of effect based on the anticipated movement distance.
- Mobility increases when moving away from an initial position in an area with a high occupied density.

$$occupied\_density = \frac{area\_occupied\_by\_objects}{total\_area\_of\_the\_region} \quad (1)$$

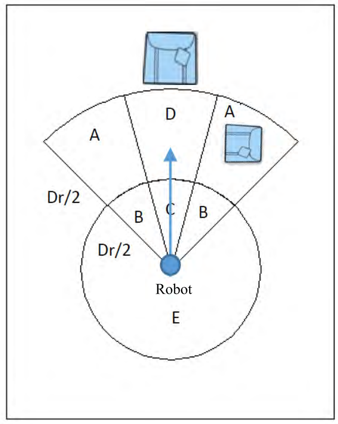


**FIGURE 10. The regions into which the surroundings of the robot are divided for analyzing the occupied density distribution in the work presented in [67]. The arrow indicates the intended direction of movement. $D_r$ is obtained as illustrated in Fig. 8. This figure has been adapted from [67].**

To analyze the occupied density distribution, the area surrounding the robot is divided into regions as shown in Fig. 10. The occupied density of each area is calculated (i.e., $OD_{\{A\}}$ for region A, $OD_{\{B\}}$ for region B, $OD_{\{C\}}$ for region C, $OD_{\{D\}}$ for region D and $OD_{\{E\}}$ for region E). Then, the combined occupied densities of regions A and D ($OD_{\{A,D\}}$) and regions B and C ($OD_{\{B,C\}}$) are calculated using (2) and (3), respectively, in which the priority constants are assigned in accordance with the natural human tendencies described above. Then, the combined occupied density of regions A, B, C and D ($OD_{\{A,B,C,D\}}$) can be obtained using (4), where $\delta_{AD}$ and $\delta_{BC}$ are priority constants that depend on the uncertain term. Finally, the perceptive distance ($D$) can be obtained from (5), where $\delta_{ABCD}$ and $\delta_E$ are experimentally defined scalar constants.

The effectiveness of the proposed method has been evaluated in a user study and proven to be higher than that of the system proposed in [65] because the proposed method can replicate more natural human tendencies that were not considered when designing the method proposed in [65]. However, the main criticism of this work seems to be the brittleness of the mathematical formula defined for adapting the perceptive distance ($D$) since the priority constants are manually coded and the regions defined for calculating the occupied density have hard boundaries, thereby disregarding the natural fuzziness inherent in spatial categorization.

$$OD_{\{A,D\}} = 0.66 \times OD_{\{D\}} + 0.33 \times OD_{\{A\}} \quad (2)$$
$$OD_{\{B,C\}} = 0.66 \times OD_{\{C\}} + 0.33 \times OD_{\{B\}} \quad (3)$$
$$OD_{\{A,B,C,D\}} = \delta_{AD} OD_{\{A,D\}} + \delta_{BC} OD_{\{B,C\}} \quad (4)$$
$$D = Dr[1 - \delta_{ABCD} OD_{\{A,B,C,D\}} + \delta_E OD_{\{E\}}] \quad (5)$$

The ability of service robots to generate uncertain terms in their own voice responses is also an important feature for the development of human-like communication abilities in robots. The method proposed in [68] is capable of synthesizing uncertain terms related to the sizes of objects for use in voice responses. The proposed concept uses visual attention to adapt the perception of uncertain information. The system has been implemented with a fuzzy inference system that can generate an uncertain term by analyzing the following environmental parameters in a particular scenario: the average size of the surrounding objects, the size of the object of interest and the size of the region of interest. Experimental results from a user study validate the applicability of the proposed system for synthesizing uncertain terms to be used in robot voice responses. However, the system is only capable of generating a predefined set of uncertain terms related to the sizes of objects.

### D. ROBOTIC SYSTEMS THAT ADAPT THEIR PERCEPTIONS IN ACCORDANCE WITH EXPERIENCE

Humans build up a knowledge base by acquiring knowledge through experience. This knowledge base can be used to form an understanding of the working environment, user expectations and context. Furthermore, such knowledge acquisition enhances the ability to interpret fuzzy linguistic information in accordance with the current environmental context and the expectations of the user. Therefore, experience is also an important factor in adapting the perception of a robot regarding uncertain information, and systems have been developed for perception adaptation based on a robot's experience.

The meaning of an uncertain term depends on the immediately previous state. For example, consider two persons driving a car. A person who has already driven 100 km may think that driving another 10 km would be a short distance, whereas a person who has driven only 15 km may think that driving another 10 km would be a long distance. Based on this phenomenon, the method proposed in [69] and [70] assumes that the quantitative meaning of an uncertain term depends on the immediately preceding movement of the robot. The proposed system is known as the fuzzy coach-player system and can be used to teach a robot certain behaviors using natural language instructions. The quantitative values associated with uncertain terms are interpreted by a fuzzy inference system that uses the immediately preceding moment of the robot as an input. The end effector movements and single joint movements of a manipulator were considered in [69] and [70], respectively, for implementation. The input and output membership functions of the fuzzy inference system used in [70] are shown in Fig. 11. The fuzzy sets in the membership functions are fixed and defined based on expert knowledge. Therefore, the adaptivity of the system to different conditions is limited, which is one of the major shortcomings of this method.

However, there are situations in which the immediately previous state alone will misrepresent the overall experience. For example, consider a situation in which a person has driven a car 80 km, 100 km, 70 km, 90 km and 2 km in 5 consecutive trips. If only the immediately previous state is considered (i.e., only the 2 km trip), the person's experience is not represented correctly. Therefore, a set of previous states should be considered to achieve an enhanced assessment of the situation. Based on this idea, the method proposed in [72] interprets uncertain information by considering a set of movements of the robot. This method uses the concept of internal rehearsal [73], namely, an internal simulation that replicates the ability of humans to internally perceive and manipulate their environment and to forecast the future [74]. A functional overview of the system is depicted in Fig. 12. The system consists of two main components: the fuzzy inference system and the internal rehearsal system. The functionality of the fuzzy inference system is similar to that of the fuzzy inference system used in [70] (i.e., the fuzzy inference system depicted in Fig. 11), although it is implemented as a fuzzy neural network. The fuzzy inference system is responsible

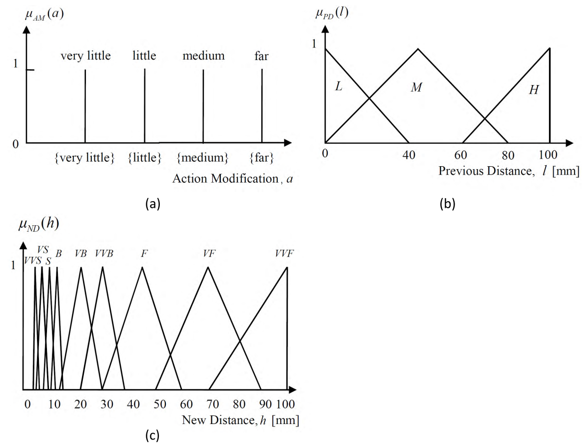


**FIGURE 11.** The membership functions of the fuzzy inference system used in [70]. (a) shows the input membership function for the action modifier (i.e., the uncertain term in a particular user instruction). (b) shows the input membership function for the robot's previous movement. (c) shows the output membership function. This figure is based on [71].

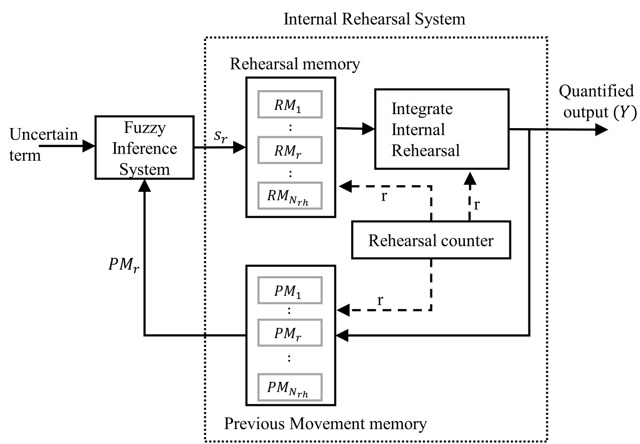


**FIGURE 12.** Flow chart of the core functionality of the internal rehearsal system proposed in [72]. This figure is based on [72].

for evaluating the quantitative meaning of an uncertain term in a particular instruction by considering the previous movements in a manner similar to [70]. However, the internal rehearsal system suggests the corresponding previous movements. The internal rehearsal system consists of a Rehearsal Memory (RM), a Previous Movement memory (PM) and a Rehearsal Counter. The RM stores the internally simulated output values provided by the fuzzy inference system (i.e., $s_r$) for the suggested previous movements ($PM_r$). The process is iterated from $r = 1$ to $r = N_{rh}$, where $r$ is the count of internal rehearsals and $N_{rh}$ is the defined threshold limit (which indicates how many previous movements are to be considered as representative of the robot's experience), without performing any real movement. Thereafter, the simulated outputs ($s_r$) in the RM are integrated as expressed in (6) to determine the required quantitative output ($Y$) for the movement, where $p_r$ is a constant that represents the probability of relevancy of the $r^{\text{th}}$ internal rehearsal for the final outcome. Therefore, $p_r$ is defined such that the probability of relevance decays over time (i.e., the value of $p_r$ exponentially decays from $r = 2$ to $r = N_{rh}$), similar to human memory. Thus, more recent previous states have a higher influence on the final output than states further in the past. The variation in the quantitative values interpreted for fuzzy linguist information with the number of internal rehearsals has been analyzed to assess the performance of the proposed concept. The proposed concept has been implemented for the end effector and posture control of a fixed manipulator. However, the system is only capable of handling a predefined set of uncertain terms, and the possible user commands are limited by a strict grammar model.

$$Y = \frac{\sum_{r=1}^{N_{rh}} s_r p_r}{\sum_{r=1}^{N_{rh}} p_r} \tag{6}$$

An adaptive fuzzy command acquisition network that processes fuzzy linguistic information in spoken language commands has been proposed [75]. The proposed system is capable of acquiring knowledge about fuzzy linguistic information based on user feedback. The concept has been implemented with a neural network to produce a system capable of learning new user commands and online learning. Hence, the possible user instructions are not restricted, and the system can acquire new knowledge while operating. However, as implemented, the size (number of nodes) of the network increases exponentially with the vocabulary size. The ability of the system to acquire knowledge of fuzzy commands has been experimentally verified. However, its ability to interpret quantitative values from fuzzy linguistic information has not been experimentally assessed. The initial training of the network requires a large data set, and users cannot provide natural language feedback to the system. These are the main shortcomings of the proposed concept. Furthermore, the proposed system has not been implemented in a real robotic system, although it is expected to be applicable for voice-controlled robots, online information retrieval systems, etc.

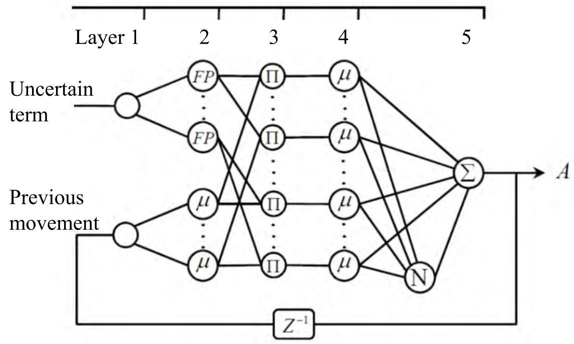


**FIGURE 13. The structure of the fuzzy neural network used in [76] for interpreting uncertain information. Note that only the important components for the interpretation process are included here, and hence, the numbering of the layers is different from that in the original publication. This figure has been adapted from [78].**

Jayasekara *et al.* [76] proposed a method for adapting the perception of fuzzy linguistic information based on user feedback. The proposed concept was implemented with a fuzzy neural network. The important components and layers of the fuzzy neural network are shown in Fig. 13. The first layer contains two types of nodes for acquiring two types of inputs: the uncertain term to be interpreted and the previous movement of the robot. The second layer act as the fuzzification layer, and the same input membership functions used in [70] and [72] are also used here, with slightly modified fuzzy sets. The third layer represents the rules by taking algebraic products between the outputs of the second layer in T-norm form, and the output of the $i^{\text{th}}$ node in this layer represent the firing strength of the $i^{\text{th}}$ rule ($\mu_i$). The fourth layer links each fuzzy antecedent to its consequent, with each node $i$ representing a triangular fuzzy set with center $a_i$ and width $b_i$. The parameters in this layer (i.e., $a_i$ and $b_i$) are initialized with values somewhat similar to the output membership function of the fuzzy inference system presented in [70] and [72] (however, a uniform distribution of fuzzy sets over the universe of discourse is considered here). The fifth layer is the defuzzification layer, and the defuzzified output ($A$) is obtained via (7) using sum-product composition for Mamdani fuzzy systems [77], where $N_R$ is the number of rules.

$$A = \frac{\sum_{i=1}^{i=N_R} a_i b_i \mu_i}{\sum_{i=1}^{i=N_R} b_i \mu_i} \tag{7}$$

The connection weights of the fifth layer of the network (i.e., $a_i$ and $b_i$) are adjusted based on user feedback. This enables more natural communication and ultimately enhances the interaction between the user and the robot. To assign quantitative values to the feedback terms, a module called a vocal cue evaluation system is deployed. This module was developed based on a fuzzy inference system that assumes that the quantitative meanings of feedback terms depend on the immediately previous state of the robot. The connection weights are modified though backpropagation based on the quantitative error identified via user feedback.

The performance of the system has been further improved by considering the willingness of the user [78], which can be used as a parameter to identify the user's motivation to change the perception of the robot regarding a particular uncertain term. This parameter is evaluated by considering a series of user feedback instances. The performance improvement of the system due to the consideration of the user's willingness for adaptation has been experimentally investigated by defining a performance index called the user satisfactory level. The satisfactory level is the ratio between the number of feedback instances of the "good" type received and the total number of feedback instances. The proposed system is capable of adapting the robot's perception of uncertain information toward that of the user. The system has been implemented to control the end effector of a fixed robotic manipulator. The possible user commands and feedback terms are bounded by a strict rule set. Furthermore, the system cannot evaluate subconscious body movements of the user that could be used as feedback, such as facial expressions; instead, explicit feedback must be given to adapt the robot's perception, which places a burden on the user.

None of the systems mentioned in this section can perceive the environment through sensors; hence, these systems cannot adapt their perceptions in response to changes in the environment. Therefore, they are not suitable for use in a dynamic environment or for mobile tasks since their experience is only applicable for a particular environment. This is the major limitation of systems that adapt based solely on experience.

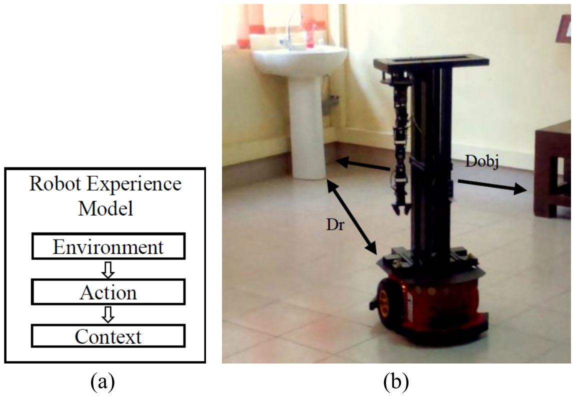


**FIGURE 14. (a) shows the structure of the Robot Experience Model (REM) [79]. (b) illustrates the environmental parameters used by the fuzzy inference system for interpreting positional information in [79]. These figures have been adapted from [79].**

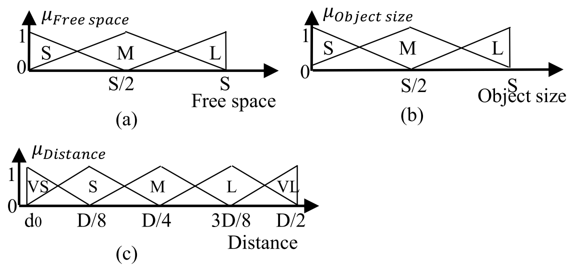


**FIGURE 15. The membership functions of the fuzzy inference system used in [79] for interpreting uncertain information related to positional information. (a) shows the input membership function for the available free space, which is adjusted according to the room size (*S*). (b) shows the input membership function for the object size, which is also adjusted according to the room size (*S*). (c) shows the output membership function, which is adjusted according to the perceptive distance (*D*). d0 is used to maintain a safe clearance. These figures have been adapted from [79].**

### *E. ROBOTIC SYSTEMS THAT ADAPT THEIR PERCEPTION IN ACCORDANCE WITH BOTH THE ENVIRONMENT AND THEIR EXPERIENCE*

The meaning of uncertain terms depends on both the environment and previous experience. Therefore, consideration of only one aspect is not sufficient for effective perception adaptation. Consequently, methods have been developed for adapting a robot's perception of uncertain information based on both environmental factors and experience.

Muthugala and Jayasekara [79] introduced the concept of the Robot Experience Model (REM) to enhance the effectiveness of understanding uncertain information. The REM is a hierarchical structure (as shown in Fig, 14(a)) that organizes a robot's knowledge of its environment, actions and context. The knowledge of the REM is used to identify the required set of robot actions to satisfy a given user command based on the environment. The proposed system includes two submodules for the interpretation of uncertain information related to motional and positional information. The submodule required for interpretation is selected based on the type of robot action to be performed. Uncertainties in motional commands are interpreted by a fuzzy inference system similar to the one proposed in [65] (i.e., the system shown in Fig. 9).

Uncertain information related to positional information is interpreted by another fuzzy inference system that evaluates the available free space, the room size, the size of the reference object and the distance between two objects (illustrated in Fig. 14(b) as *Dobj*) or the distance between the robot and the reference object (illustrated in Fig. 14(b) as *Dr*) as the basis for the interpretation. The input and output membership functions of the fuzzy inference system are depicted in Fig. 15. Here, it is assumed that the meanings of lexical symbols representing positional information such as "close" and "near" are the same. Hence, the quantitative distance does not change with such uncertain positional information. The output membership function is adjusted according to the perceptive distance ($D$), which is defined as $D = min(Dr, Dobj)$. The possible user commands are not restricted by a grammar model, and hence, flexible user commands can be used. The system is also capable of learning new lexical symbols and identifying changes in the environment by acquiring knowledge through interactive discussion, as explained in [80]. The proposed system is capable of navigating a mobile robot inside a domestic environment

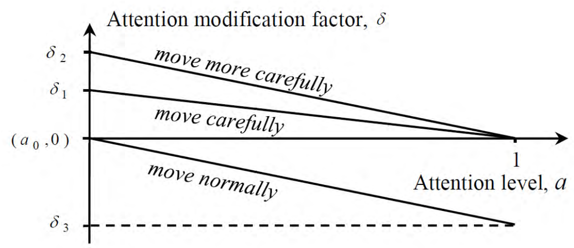


**FIGURE 16. The linear functions used in [61] and [82] to modify the attention level. Reprinted with permission ©2009 IEEE.**

based on user instructions that contain uncertain information. To validate the performance of the system, the movements of a robot were compared against user expectations. The context layer of the REM is inactive, and hence, the robot's perception cannot be adapted in accordance with its context. Furthermore, the system is not capable of adapting the robot's perception towards the expectations of the user. These are the main limitations of the proposed system.

### F. ROBOTIC SYSTEMS THAT ADAPT IN RESPONSE TO INFLUENTIAL USER INSTRUCTIONS

A robot's attention can be modified based on an external stimulus such as a voice command [81], and hence, attentive instructions such as "move carefully" influence the quantitative meanings of fuzzy implications in user commands. Therefore, Jayasekara *et al.* [61] and Izumi *et al.* [82] proposed the use of an attentive modification factor to modify the perception of fuzzy linguistic information when an attentive instruction is given. The proposed concept has been utilized in combination with the uncertain information evaluation method proposed in [58] and [76]. The user can use a set of predefined attentive instructions to influence the robot's perception by adjusting the attentive modification factor. The attentive modification factor is a linear function (as shown in Fig. 16) that depends on different attentive instructions such as "move more carefully" and "move carefully," and this function is defined such that it can replicate natural human behaviors, such as the fading of the effect of an attentive instruction with successive operations.

## IV. LIMITATIONS OF THE EXISTING SYSTEMS AND POSSIBLE IMPROVEMENTS

The existing systems have limitations, and their performance in understanding uncertain information is far below the capabilities of humans. Hence, the existing methods should be improved to enhance human-robot interactions. The limitations of the existing methods for understanding uncertain information have been analyzed, and possible improvements are suggested based on the following three aspects: scope, interaction and adaptation. Uncertain information can be related to various entities, such as spatial information, time, and event count, and these related entities define the scope of the uncertain information to be interpreted. The way in which the interactions between users and robots occur and the way in which the perception of uncertain information is adapted are regarded as the aspects of interaction and adaptation, respectively, in this analysis. The current status of the methods used for understanding uncertain information and the identified possible improvements are summarized taxonomically in Table 1.

### A. SCOPE

Humans unthinkingly include uncertain terms related to many different entities in voice instructions and suggestions. However, most existing systems are limited to handling only uncertain information related to distances in the environment [63], [70], [76], [79], the speed of movements [55], [75], the directions of objects [53], the sizes of objects [60], [68] and the joint angles of manipulators [69], [72]. Uncertain terms related to other aspects such as time, event counts and processing tasks have not been addressed. Therefore, it would be interesting to extend the capabilities of the existing systems to incorporate the ability to understand uncertain information related to such entities. However, this would be a challenging task since the factors that affect the meanings of such uncertain information would need to be identified, as previous studies have not revealed this information.

### B. INTERACTION

The present systems that have been developed for interpreting uncertain information are capable only of interacting with humans through voice communication. Hence, their interactions are unimodal, and these systems are not capable of gathering information conveyed through interaction modalities such as hand gestures, facial expressions or body movements. The information conveyed through modalities other than voice communication can be used as a supportive aid to enhance the understanding of uncertain information contained in voice instructions. Furthermore, facial expressions and subconscious body movements can be used as a substitute for voice feedback in systems that adapt based on user feedback, such as that presented in [76]. This will ultimately reduce the overhead burden on the user and hence improve interaction.

The inclusion of uncertain terms in vocal responses generated by robots can enhance their human-like communication abilities. However, most present systems are only capable of interpreting uncertain information in user instructions, and few studies have addressed the generation of uncertain terms to be used in robot vocal responses. The system proposed in [53] assigns fixed meanings to uncertain terms used in responses, whereas the method proposed in [68] is capable of synthesizing uncertain terms related only to the sizes of objects by adapting the robot's perception based on visual attention. Therefore, the capabilities of the existing systems are not sufficient in this regard, and studies should be conducted to develop methods of effectively generating uncertain terms to be used in the vocal responses of robots.

**TABLE 1. Summary of the current status of the methods used for understanding uncertain information and the possible future improvements.**

| | Current Status | Possible Improvements |
|---|---|---|
| Scope | ⊗ Understanding is limited to uncertain information related to<br>• Distances in the environment (e.g., [57], [58], [63], [67], [70], [72], [76] and [79])<br>• Speed of movements (e.g., [55] and [75])<br>• Directional information (e.g., [53])<br>• Object sizes (e.g., [60] and [68])<br>• Joint angles (e.g., [69] and [72]) | ⊗ Extend the capabilities to the understanding of uncertain information related to other aspects such as time, event counts and process/task-related information |
| Interaction | ⊗ All systems are capable only of interacting through voice communication between the robot and the user<br>⊗ Very few works have attempted to synthesize uncertain information to be used in robot vocal responses (e.g., [53] and [68])<br>⊗ All other works have focused on addressing the issue of interpreting the uncertain information in user instructions | ⊗ Extend the capabilities to the evaluation of information conveyed non-verbally to adapt the perception of uncertain information<br>⊗ Develop methods of adaptively synthesizing uncertain information to be used in robot vocal responses |
| Adaptation | ⊗ The systems are capable of adapting their perception of uncertain information based on<br>• The environment<br>∘ Frame size (e.g., [63])<br>∘ Average distance between objects (e.g., [58])<br>∘ Room size, free space and arrangement of obstacles (e.g., [65] and [67])<br>• Experience<br>∘ Immediately previous movement (e.g., [69] and [70])<br>∘ Set of previous movements (e.g., [72] )<br>∘ User feedback (e.g., [75] and [76])<br>• Experience and environment<br>∘ Size of reference object, room size, free space and arrangement of objects (e.g., [79])<br>• Influential user commands (e.g., [61] and [82])<br>⊗ Fuzzy logic (type I) and fuzzy neural networks are often used<br>• Fuzzy logic (e.g., [57], [58], [68], [69] and [79])<br>• Fuzzy neural networks (e.g., [72], [75] and [78])<br>⊗ Performance evaluations have been conducted through user studies<br>• User satisfactory level (e.g., [78])<br>• User feedback (e.g., [79])<br>• Human response (e.g., [67] and [68]) | ⊗ Consider multiple bases for concurrent adaptation of perception, e.g.,<br>• Environment and user feedback<br>• Environment and contextual knowledge of different objects and tasks<br>⊗ Consider more environmental factors for adaptation of perception, e.g.,<br>• Heights of objects<br>• Shapes of objects<br>⊗ Perceive the environment in a human-like manner to improve the perception effectiveness, e.g.,<br>• Use a human-like vision system instead of overhead cameras<br>⊗ Consider the focusing of human attention to extract and identify key environmental parameters<br>⊗ Adapt perception by considering specific knowledge of a particular context, e.g.,<br>• Common properties of the arrangement of a lunch table<br>• Danger of hot items or flames<br>⊗ Investigate the possibility of using fuzzy type II systems for interpreting uncertain information<br>⊗ Introduce an objective performance measurement index |

Note that only key publications are given as examples.

### C. ADAPTATION

According to the analysis of the existing literature, various methods are used in existing systems to adapt a robot's perception of uncertain information. These methods rely on different types of information and different artificial intelligence (AI) techniques. Therefore, the limitations of the existing systems and the possible improvements to the adaptation methods are analyzed separately with respect to these two aspects, i.e., the adaptation entities and the artificial intelligence techniques. Furthermore, the performance evaluation methods used in the existing approaches are also discussed.

#### 1) ADAPTATION ENTITIES

Existing methods are capable of adapting a robot's perception based on different factors (or entities) that effect the meaning of uncertain information. For example, the methods proposed in [58], [63], and [65] are capable of perception adaptation based on the environment, whereas the methods proposed in [70], [72], and [76] are capable of perception adaptation based on experience, and the method proposed in [79] is capable of utilizing both experience and environmental factors for perception adaptation.

Systems that adapt their perceptions based on the environment are capable of adaptation based on the spatial factors of the environment. For instance, the method proposed in [63] uses the room size, the method proposed in [58] uses the average distance between objects, and the method proposed in [65] uses the room size, the available free space and the arrangement of the obstacles in the environment. In most studies, a robot perceives its environment through sonar sensors and navigation maps; visual feedback is used only in the system proposed in [58]. Therefore, the environmental perception capabilities of the existing methods are limited to a few aspects of the environment. For example, Muthugala and Jayasekara [79] proposed a navigation system that considers the sizes of landmark objects when assigning quantitative values to uncertain positional information. However, the system cannot determine the actual size of a landmark object because it uses only the footprint area of the object without considering the height. This ultimately hampers the performance of the system. The system that utilizes visual information to perceive its environment [58] does not possess stereoscopic vision; instead, it uses an overhead camera. Therefore, the view of the camera is not similar to that of a human. However, the meaning of uncertain information may differ depending on the field/angle of view. Hence, this system is at a disadvantage for interpreting uncertain information since it does not have human-like environmental perception. To improve the effectiveness of robots in understanding uncertain information, a human-like visual attentive mechanism should be incorporated into robots in future work. Furthermore, other environmental factors that influence the meaning of uncertain information are not utilized by the existing methods. However, no comprehensive studies have been conducted to identify all of the environmental factors that affect the meaning of uncertain terms and how their influence manifests. Therefore, investigations need to be performed to identify the influential environmental factors and how they can be used to adapt a robot's perception of uncertain information.

#### 2) AI TECHNIQUES

Most systems that can understand uncertain terms in user commands utilize fuzzy inference systems to assign quantitative values to those uncertain terms (e.g., [57], [58], [63], and [79]). To achieve a learning ability, fuzzy neural networks are utilized in systems that can adapt their perceptions of uncertain terms based on user feedback (e.g., [75], [76], and [78]). Fuzzy logic systems are most often used for this purpose due to their ability to effectively model the knowledge of human beings in robotic systems without knowledge of the underlying dynamics [66], [77] or most of the behaviors related to the human-robot interaction domain [83], [84].

The fuzzy inference systems utilized in these systems are fuzzy type I systems. However, Mendel [85] showed that interval type II fuzzy sets can better represent linguistic uncertainties since the membership grade of an interval type II fuzzy set is an interval instead of a crisp value. Therefore, interval type II fuzzy sets can be used to improve the understanding of uncertain information by robots. Furthermore, the recent development of computationally effective algorithms for implementing type II fuzzy inference systems [86] offers the possibility of using general fuzzy type II techniques for improved performance. Therefore, an interesting direction for future work would be to model such systems using general type II fuzzy sets or interval type II fuzzy sets in order to investigate the resulting performance gain, despite the implementation and computational complexity.

As explained in section IV-C.1, the perception of uncertain information should be adapted in accordance with the context. To identify the relevant context, a fuzzy Naive Bayesian network could be used, as explained in [6]. For the fusion of multimodal interactions, it may be possible to use Bayesian networks in a manner similar to the methods explained in [44]. Such methods could be adopted as supportive aids for the interpretation of uncertain information. In addition, finite-state intention machines [1] and hierarchical structures such as the Robot Experience Model (REM) [79] are used in some systems as supportive aids for intelligent systems that interpret the uncertain information in user instructions (e.g., [79] and [80]).

#### 3) PERFORMANCE EVALUATION

Few methods of evaluating the performance of systems for interpreting uncertain information in voice commands issued to robots can be found in the available literature. An index called the user satisfactory level [76] has been used to evaluate the ability of a robot to adapt its perception toward that of the user. The user satisfactory level is calculated based on the user's acceptance of the responses of the robot to successive user instructions. Specifically, the user satisfactory level is the ratio between the number of cases accepted by the user and the total number of cases considered. In [65] and [79], user feedback was used to evaluate system performance. In [68], a human study was conducted by asking the participants to rate the sizes of objects in different scenarios using linguistic terms, and the results of that study were compared against the linguistic terms synthesized by a robotic system. Furthermore, in [67], a user study was conducted by asking users to operate a robot using a joystick. Using this approach, the findings of the user study and the output of the system could be quantitatively compared to evaluate system performance. However, these experiments and evaluation methods are highly subjective due to the subjectivity of the human participants. Therefore, the necessary human studies should be conducted in such a way that the experimental results can provide a basis for generalizability; recommendations for designing, planning and executing human studies for HRI can be found in [87]. The convergence capability of the learning function was also analyzed in the method proposed in [72], which can also be a good choice for evaluating the performance and parameter variations of such intelligent systems.

## V. CONCLUSION

This paper has presented a review of service robots coping with uncertain information in natural language voice instructions and responses. Service robots are currently being developed to cater to demands in emerging areas of robotic applications, such as health care, education, rehabilitation and assistance, and service robots with human-like interaction capabilities are preferred for such applications.

Voice communication is one of the predominant interaction modalities used to convey information between peers. Hence, service robots with human-like voice communication capabilities can provide better service. However, natural voice instructions do not convey precise quantitative information, and humans typically prefer to use uncertain terms, lexical symbols and notions rather than more precise quantitative values. Hence, the ability to interpret such uncertain information is mandatory for a human-friendly service robot.

The quantitative meanings of uncertain terms depend on several factors, such as the environment, past experience and the current context. Therefore, robotic systems should have ability to adapt their perception of uncertain information based on these entities. The existing robotic systems have been critically taxonomically investigated based on their adaptation entities.

Fuzzy logic and fuzzy neural networks are often used in the existing methodologies to interpret the uncertain information in voice instructions due to their ability to model natural human tendencies. Fuzzy inference systems and fuzzy neural networks are capable of effectively interpreting uncertain information to a great extent. However, the existing systems are nevertheless subject to limitations in their ability to interpret uncertain information in a human-like manner.

The limitations of the existing systems have been identified, and possible future improvements have been presented in this paper. In summary, the capabilities of the existing systems are far below the cognitive capabilities of human beings with regard to understanding uncertain information. Furthermore, minimal research has been done in this particular research area, although there is promising potential for future developments.

## REFERENCES


[1] R. C. Arkin, *Behavior-Based Robotics*. Cambridge, MA, USA: MIT Press, 1998.

[2] S. Koceski and N. Koceska, "Evaluation of an assistive telepresence robot for elderly healthcare," *J. Med. Syst.*, vol. 40, p. 121, May 2016. [Online]. Available: http://dx.doi.org/10.1007/s10916-016-0481-x

[3] C. Jayawardena, I.-H. Kuo, E. Broadbent, and B. A. MacDonald, "Socially assistive robot healthbot: Design, implementation, and field trials," *IEEE Syst. J.*, vol. 10, no. 3, pp. 1056–1067, Sep. 2016.

[4] H.-M. Gross *et al.*, "ROREAS: Robot coach for walking and orientation training in clinical post-stroke rehabilitation–prototype implementation and evaluation in field trials," *Auto. Robots*, vol. 41, no. 3, pp. 679–698, 2017. [Online]. Available: http://dx.doi.org/10.1007/s10514-016-9552-6

[5] A. Tapus, C. Tapus, and M. J. Mataric, "User-robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy," *Intell. Service Robot.*, vol. 1, no. 2, pp. 169–183, 2008.

[6] R. Liu and X. Zhang, "Fuzzy context-specific intention inference for robotic caregiving," *Int. J. Adv. Robot. Syst.*, vol. 13, no. 5, pp. 1–14, 2016.

[7] K. Gerling, D. Hebesberger, C. Dondrup, T. Körtner, and M. Hanheide, "Robot deployment in long-term care," *Zeitschrift für Gerontologie und Geriatrie*, vol. 49, no. 4, pp. 288–297, 2016. [Online]. Available: http://dx.doi.org/10.1007/s00391-016-1065-6

[8] Y. Zhang, Y. Hu, P. Zhang, and W. Zhang, "Development of personal assistant system with human computer interaction," *Int. J. Human Comput. Interaction*, vol. 5, no. 3, p. 40, 2014.

[9] E. Clotet, D. Martínez, J. Moreno, M. Tresanchez, and J. Palacín, "Assistant personal robot (APR): Conception and application of a tele-operated assisted living robot," *Sensors*, vol. 16, no. 5, p. 610, 2016.

[10] N. Reich-Stiebert and F. Eyssel, "Learning with educational companion robots? Toward attitudes on education robots, predictors of attitudes, and application potentials for education robots," *Int. J. Social Robot.*, vol. 7, no. 5, pp. 875–888, 2015.

[11] A. Causo, G. T. Vo, I.-M. Chen, and S. H. Yeo, "Design of robots used as education companion and tutor," in *Robotics and Mechatronics*. Cham, Switzerland: Springer, 2016, pp. 75–84.

[12] I. Aaltonen, A. Arvola, P. Heikkilä, and H. Lammi, "Hello pepper, may I tickle you?: Children's and adults' responses to an entertainment robot at a shopping mall," in *Proc. Companion ACM/IEEE Int. Conf. Human-Robot Interaction*, 2017, pp. 53–54.

[13] M. E. Foster *et al.*, "The mummer project: Engaging human-robot interaction in real-world public spaces," in *Proc. Int. Conf. Social Robot.*, 2016, pp. 753–763.

[14] Y. Kim and W. C. Yoon, "Generating task-oriented interactions of service robots," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 8, pp. 981–994, Aug. 2014.

[15] M. Mast *et al.*, "Design of the human-robot interaction for a semi-autonomous service robot to assist elderly people," in *Ambient Assisted Living*. Cham, Switzerland: Springer, 2015, pp. 15–29.

[16] K. Swift-Spong, E. Short, E. Wade, and M. J. Matarić, "Effects of comparative feedback from a socially assistive robot on self-efficacy in post-stroke rehabilitation," in *Proc. IEEE Int. Conf. Rehabil. Robot. (ICORR)*, Aug. 2015, pp. 764–769.

[17] S. Schiffer, A. Ferrein, and G. Lakemeyer, "CAESAR: An intelligent domestic service robot," *Intell. Service Robot.*, vol. 5, no. 4, pp. 259–273, 2012.

[18] M. K. Murray, "The nursing shortage: Past, present, and future," *J. Nursing Admin.*, vol. 32, no. 2, pp. 79–84, 2002.

[19] N. Super, *Who will be There to Care?: The Growing Gap Between Caregiver Supply and Demand*. Washington, DC, USA: National Health Policy Forum, 2002.

[20] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robot. Auto. Syst.*, vol. 42, no. 3, pp. 143–166, 2003.

[21] P. Menzel and F. d'Aluisio, *Robo Sapiens: Evolution of a new Species*. Cambridge, MA, USA: MIT Press, 2001.

[22] S. Frennert and B. Östlund, "Review: Seven matters of concern of social robots and older people," *Int. J. Social Robot.*, vol. 6, no. 2, pp. 299–310, 2014.

[23] S. Kleanthous *et al.*, "Analysis of elderly users preferences and expectations on service robots personality, appearance and interaction," in *Proc. Int. Conf. Human Aspects IT Aged Population*, 2016, pp. 35–44.

[24] S. Huang, T. Tanioka, R. Locsin, M. Parker, and O. Masory, "Functions of a caring robot in nursing," in *Proc. 7th Int. Conf. Natural Lang. Process. Knowl. Eng.*, 2011, pp. 425–429.

[25] N. Mavridis, "A review of verbal and non-verbal human-robot interactive communication," *Robot. Auto. Syst.*, vol. 63, pp. 22–35, Jan. 2015.

[26] F. Portet, M. Vacher, C. Golanski, C. Roux, and B. Meillon, "Design and evaluation of a smart home voice interface for the elderly: Acceptability and objection aspects," *Pers. Ubiquitous Comput.*, vol. 17, no. 1, pp. 127–144, 2013.

[27] J. A. Ansari, A. Sathyamurthy, and R. Balasubramanyam, "An open voice command interface kit," *IEEE Trans. Human-Mach. Syst.*, vol. 46, no. 3, pp. 467–473, Mar. 2016.

[28] E. Bastianelli, D. Nardi, L. C. Aiello, F. Giacomelli, and N. Manes, "Speaky for robots: The development of vocal interfaces for robotic applications," *Appl. Intell.*, vol. 44, no. 1, pp. 43–66, 2016. [Online]. Available: http://dx.doi.org/10.1007/s10489-015-0695-5

[29] K. Sugiura and K. Zettsu, "Rospeex: A cloud robotics platform for human-robot spoken dialogues," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 6155–6160.

[30] M. Mazo *et al.*, "Wheelchair for physically disabled people with voice, ultrasonic and infrared sensor control," *Auto. Robots*, vol. 2, no. 3, pp. 203–224, 1995.

[31] J. Clark and R. Roemer, "Voice controlled wheelchair," *Arch. Phys. Med. Rehabil.*, vol. 58, no. 4, pp. 169–175, 1977.

[32] K. B. Stanton *et al.*, "Psubot-a voice-controlled wheelchair for the handicapped," in *Proc. 33rd Midwest Symp. Circuits Syst.*, 1990, pp. 669–672.

[33] A. Iborra, D. A. Caceres, F. J. Ortiz, J. P. Franco, P. S. Palma, and B. Alvarez, "Design of service robots," *IEEE Robot. Autom. Mag.*, vol. 16, no. 1, pp. 24–33, Mar. 2009.

[34] M. Kim, S. Kim, S. Park, M. T. Choi, M. Kim, and H. Gomaa, "Service robot for the elderly," *IEEE Robot. Autom. Mag.*, vol. 16, no. 1, pp. 34–45, Mar. 2009.

[35] J. Fasola and M. J. Matarić, "Using spatial semantic and pragmatic fields to interpret natural language pick-and-place instructions for a mobile service robot," in *Proc. Int. Conf. Social Robot.*, 2013, pp. 501–510.

[36] E. Bastianelli, G. Castellucci, D. Croce, R. Basili, and D. Nardi, "Effective and robust natural language understanding for human-robot interaction," in *Proc. 21st Eur. Conf. Artif. Intell.*, 2014, pp. 57–62.

[37] J.-Y. Puigbo, A. Pumarola, C. Angulo, and R. Tellez, "Using a cognitive architecture for general purpose service robot control," *Connection Sci.*, vol. 27, no. 2, pp. 105–117, 2015.

[38] K.-P. Gapp, "Basic meanings of spatial computation and evaluation in 3D space," in *Proc. AAAI*, 1994, pp. 1393–1398.

[39] J. D. Kelleher, G.-J. M. Kruijff, and F. J. Costello, "Proximity in context: An empirically grounded computational model of proximity for processing topological spatial expressions," in *Proc. 21st Int. Conf. Comput. Linguistics 44th Annu. Meeting Assoc. Comput. Linguistics*, 2006, pp. 745–752.

[40] S. Guadarrama *et al.*, "Grounding spatial relations for human-robot interaction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 1640–1647.

[41] T. Kollar, S. Tellex, D. Roy, and N. Roy, "Toward understanding natural language directions," in *Proc. 5th ACM/IEEE Int. Conf. Human-Robot Interaction (HRI)*, Mar. 2010, pp. 259–266.

[42] B. Talbot, O. Lam, R. Schulz, F. Dayoub, B. Upcroft, and G. Wyeth, "Find my office: Navigating real space from semantic descriptions," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 5782–5787.

[43] I. Zukerman, S. N. Kim, T. Kleinbauer, and M. Moshtaghi, "Employing distance-based semantics to interpret spoken referring expressions," *Comput. Speech Lang.*, vol. 34, no. 1, pp. 154–185, 2015.

[44] D. Whitney, M. Eldon, J. Oberlin, and S. Tellex, "Interpreting multimodal referring expressions in real time," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 3331–3338.

[45] H. Admoni, T. Weng, and B. Scassellati, "Modeling communicative behaviors for object references in human-robot interaction," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 3352–3359.

[46] N. Kubota and Y. Toda, "Multimodal communication for human-friendly robot partners in informationally structured space," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 1142–1151, Nov. 2012.

[47] G. Randelli, T. M. Bonanni, L. Iocchi, and D. Nardi, "Knowledge acquisition through human–robot multimodal interaction," *Intell. Service Robot.*, vol. 6, no. 1, pp. 19–31, 2013.

[48] K. Lynch, *The Image of the City*, vol. 11. Cambridge, MA, USA: MIT Press, 1960.

[49] R. Briggs, "Urban cognitive distance," in *Image and Environment: Cognitive Mapping and Spatial Behavior*. New Brunswick, NJ, USA: Aldine, 1973, pp. 361–388.

[50] D. R. Montello, "The measurement of cognitive distance: Methods and construct validity," *J. Environ. Psychol.*, vol. 11, no. 2, pp. 101–122, 1991.

[51] S. Dutta, "Approximate spatial reasoning: Integrating qualitative and quantitative constraints," *Int. J. Approx. Reasoning*, vol. 5, no. 3, pp. 307–330, 1991.

[52] E. Clementini, P. Di Felice, and D. Hernández, "Qualitative representation of positional information," *Artif. Intell.*, vol. 95, no. 2, pp. 317–356, 1997.

[53] M. Skubic *et al.*, "Spatial language for human-robot dialogs," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 34, no. 2, pp. 154–167, May 2004.

[54] M. Skubic, P. Matsakis, G. Chronis, and J. Keller, "Generating multilevel linguistic spatial descriptions from range sensor readings using the histogram of forces," *Auto. Robots*, vol. 14, no. 1, pp. 51–69, 2003.

[55] K. Pulasinghe, K. Watanabe, K. Izumi, and K. Kiguchi, "Modular fuzzy-neuro controller driven by spoken language commands," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 1, pp. 293–302, Feb. 2004.

[56] A. Chatterjee, K. Pulasinghe, K. Watanabe, and K. Izumi, "A particle-swarm-optimized fuzzy-neural network for voice-controlled robot systems," *IEEE Trans. Ind. Electron.*, vol. 52, no. 6, pp. 1478–1489, Dec. 2005.

[57] K. Kawamura, S. Bagchi, and T. Park, "An intelligent robotic aid system for human services," in *Proc. NASA Conf. Publication*, 1994, p. 413.

[58] A. G. Buddhika, P. Jayasekara, K. Watanabe, and K. Izumi, "Understanding user commands by evaluating fuzzy linguistic information based on visual attention," *Artif. Life Robot.*, vol. 14, no. 1, pp. 48–52, 2009.

[59] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inf. Theory*, vol. 8, no. 2, pp. 179–187, Feb. 1962.

[60] C. Jayawardena, K. Watanabe, and K. Izumi, "Learning of object identification by robots commanded by natural language," in *Proc. IAS*, 2006, pp. 913–920.

[61] A. B. P. Jayasekara, K. Watanabe, and K. Izumi, "Evaluating fuzzy voice commands by user guided visual attention for controlling a robot manipulator," in *Proc. ICCAS-SICE*, 2009, pp. 3136–3141.

[62] S. Schiffer, A. Ferrein, and G. Lakemeyer, "Fuzzy representations and control for domestic service robots in golog," in *Intelligent Robotics and Applications*. Berlin, Germany: Springer, 2011, pp. 241–250.

[63] S. Schiffer, A. Ferrein, and G. Lakemeyer, "Reasoning with qualitative positional information for domestic domains in the situation calculus," *J. Intell. Robot. Syst.*, vol. 66, nos. 1–2, pp. 273–300, 2012.

[64] H. J. Levesque, R. Reiter, Y. Lesperance, F. Lin, and R. B. Scherl, "GOLOG: A logic programming language for dynamic domains," *J. Logic Programm.*, vol. 31, no. 1, pp. 59–83, 1997.

[65] M. A. V. J. Muthugala and A. G. B. P. Jayasekara, "Interpreting fuzzy linguistic information in user commands by analyzing movement restrictions in the surrounding environment," in *Proc. Moratuwa Eng. Res. Conf. (MERCon)*, Apr. 2015, pp. 124–129.

[66] C. C. Lee, "Fuzzy logic in control systems: Fuzzy logic controller. I," *IEEE Trans. Syst., Man and*, vol. 20, no. 2, pp. 404–418, Mar. 1990.

[67] M. A. V. J. Muthugala and A. G. B. P. Jayasekara, "Interpretation of uncertain information in mobile service robots by analyzing surrounding spatial arrangement based on occupied density variation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 1517–1523.

[68] M. A. V. J. Muthugala and A. G. B. P. Jayasekara, "Synthesizing fuzzy linguistic vocal responses by adapting perception of robot based on visual attention," in *Proc. 7th Int. Conf. Inf. Autom. Sustainability*, Dec. 2014, pp. 1–6.

[69] C. Jayawardena, K. Watanabe, and K. Izumi, "Posture control of robot manipulators with fuzzy voice commands using a fuzzy coach–player system," *Adv. Robot.*, vol. 21, nos. 3–4, pp. 293–328, 2007.

[70] C. Jayawardena, K. Watanabe, and K. Izumi, "Controlling a robot manipulator with fuzzy voice commands using a probabilistic neural network," *Neural Comput. Appl.*, vol. 16, no. 2, pp. 155–166, 2007.

[71] K. Watanabe, C. Jayawardena, and K. Izumi, "Approximate decision making by natural language commands for robots," in *Proc. 32nd Annu. Conf. IEEE Ind. Electron. (IECON)*, Nov. 2006, pp. 4480–4485.

[72] A. G. B. P. Jayasekara, K. Watanabe, K. Kiguchi, and K. Izumi, "Interpreting fuzzy linguistic information by acquiring robot's experience based on internal rehearsal," *J. Syst. Design Dyn.*, vol. 4, no. 2, pp. 297–313, 2010.

[73] E. Erdemir, C. B. Frankel, K. Kawamura, S. M. Gordon, S. Thornton, and B. Ulutas, "Towards a cognitive robot that uses internal rehearsal to learn affordance relations," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 2016–2021.

[74] G. Hesslow, "Conscious thought as simulation of behaviour and perception," *Trends Cognit. Sci.*, vol. 6, no. 6, pp. 242–247, 2002.

[75] C.-T. Lin and M.-C. Kan, "Adaptive fuzzy command acquisition with reinforcement learning," *IEEE Trans. Fuzzy Syst.*, vol. FS-6, no. 1, pp. 102–121, Feb. 1998.

[76] A. G. Jayasekara, K. Watanabe, K. Kiguchi, and K. Izumi, "Adaptation of robot's perception of fuzzy linguistic information by evaluating vocal cues for controlling a robot manipulator," *J. Artif. Life Robot.*, vol. 15, no. 1, pp. 5–9, 2010.

[77] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing; A Computational Approach to Learning and Machine Intelligence*. Upper Saddle River, NJ, USA: Prentice-Hall, 1997.

[78] A. G. B. P. Jayasekara, K. Watanabe, K. Kiguchi, and K. Izumi, "Interpretation of fuzzy voice commands for robots based on vocal cues guided by user's willingness," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2010, pp. 778–783.

[79] M. A. V. J. Muthugala and A. G. B. P. Jayasekara, "Enhancing human-robot interaction by interpreting uncertain information in navigational commands based on experience and environment," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 2915–2921.

[80] M. A. V. J. Muthugala and A. G. B. P. Jayasekara, "MIRob: An intelligent service robot that learns from interactive discussions while handling uncertain information in user instructions," in *Proc. Moratuwa Eng. Res. Conf. (MERCon)*, Apr. 2016, pp. 397–402.

[81] M. W. Eysenck and M. T. Keane, *Cognitive Psychology: A Student's Handbook*. New York, NY, USA: Taylor & Francis, 2000.

[82] K. Izumi, A. Jayasekara, K. Watanabe, and K. Kiguchi, "Attentive and corrective feedback for adapting robot's perception on fuzzy linguistic information," in *Proc. SICE Annu. Conf.*, 2010, pp. 1963–1968.

[83] I.-H. Kuo and C. Jayawardena, "Application of fuzzy techniques in human-robot interaction—A review," in *Proc. Int. Conf. Social Robot.*, 2014, pp. 249–255.

[84] E. Dadios, A. Valencia, D. Johnson, J. Biliran, and R. Garcia, *Humanoid Robot: Design and Fuzzy Logic Control Technique for its Intelligent Behaviors*. Rijeka, Croatia: InTech, 2012. [Online]. Available: https://books.google.lk/books?id=ILmHoAEACAAJ

[85] J. M. Mendel, "Computing with words: Zadeh, Turing, popper and Occam," *IEEE Comput. Intell. Mag.*, vol. 2, no. 4, pp. 10–17, Nov. 2007.

[86] J. M. Mendel, "General type-2 fuzzy logic systems made simple: A tutorial," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 5, pp. 1162–1182, Oct. 2014.

[87] C. L. Bethel and R. R. Murphy, "Review of human studies methods in HRI and recommendations," *Int. J. Social Robot.*, vol. 2, no. 4, pp. 347–359, 2010.

**M. A. VIRAJ J. MUTHUGALA** (S'15) received the B.Sc.Eng. degree (Hons.) from the University of Moratuwa, Sri Lanka, in 2014, where he is currently pursuing the Ph.D. degree with the Department of Electrical Engineering.

His main research interests include human–robot interaction, social robotics, service robotics, and intelligent systems. He is currently the Assistant Secretary of the IEEE Robotics and Automation Society, Sri Lanka Section Chapter.

**A. G. BUDDHIKA P. JAYASEKARA** (M'07–SM'17) received the B.Sc.Eng. and M.Sc. degrees in electrical engineering from the University of Moratuwa, in 2004 and 2006, respectively, and the Ph.D. degree in robotics and intelligent systems from Saga University, Japan, in 2010. He joined the Department of Electrical Engineering, University of Moratuwa, in 2010, where he is currently a Senior Lecturer.

He has been the Principle Investigator of the Intelligent Service Robotics Group, Department of Electrical Engineering, University of Moratuwa, and the Director of the Center for Advanced Robotics, University of Moratuwa, since 2017. He is currently involved in two main projects: intelligent mobile robot for domestic environments (MIRob—Moratuwa Intelligent Robot) and intelligent wheelchair for domestic environments. His main research interests are artificial intelligence techniques, human–robot interaction, human-friendly robotics, machine learning, and intelligent systems. He received the Japanese Government Monbukagakusho Scholarship to pursue his Ph.D. degree in Japan.