Enhancing Interpretation of Uncertain Information in Navigational Commands for Service Robots Using Neuro-Fuzzy Approach

Muthugala Arachchige Viraj Jagathpriya Muthugala

(148036P)

Degree of Doctor of Philosophy

Department of Electrical Engineering

University of Moratuwa
Sri Lanka

June 2018

Enhancing Interpretation of Uncertain Information in Navigational Commands for Service Robots Using Neuro-Fuzzy Approach

Muthugala Arachchige Viraj Jagathpriya Muthugala

(148036P)

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical Engineering

Department of Electrical Engineering

University of Moratuwa
Sri Lanka

June 2018

DECLARATION

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

....

M. A. Viraj J. Muthugala Date

The above candidate has carried out research for the PhD thesis under my supervision.

....

Dr. A. G. Buddhika P. Jayasekara Date

Abstract

An intelligent service robot is a machine that is able to gather information from the environment and use its knowledge to operate safely in a meaningful and purposive manner. Intelligent service robots are currently being developed to cater to demands in emerging areas of robotic applications such as caretaking and assistance, healthcare and edutainment. These service robots are intended to be operated by nonexpert users. Hence, they should have the ability to interact with humans in a human-friendly manner. Humans prefer to use voice instructions, responses, and suggestions in their daily interactions. Such voice instructions and responses often include uncertain information such as "little" and "far" rather than precise quantitative values. The uncertain information such as "little" and "far" have no definitive meanings and depend heavily on factors such as environment, context, user and experience. Therefore, the ability of robots to understand uncertain information is a crucial factor in the implementation of human-friendly interactive features in robots.

This research has been conducted with the intention of developing effective methodologies for interpreting uncertain notions such as "little", "near" and "far" in navigational user commands in order to enhance human-robot interaction. The natural tendencies of humans have been considered for the development of the methodologies since ability of the robot in replicating the natural behavior of humans vastly enhances the rapport between the robot and the user. The methodologies have been developed using fuzzy logic and fuzzy neural networks that are capable of adapting the perception of uncertain information according to the environment, experience and user. User studies have been conducted in artificially created domestic environments to experimentally validate the performance of the proposed methods. An intelligent service robot named as Moratuwa Intelligent Robot (MIRob), which has been developed as a part of the research, has been used for the experiments.

The robot's perception of distance and direction related uncertain information in navigation commands is adapted according to the environment. According to the experimental results, a service robot can effectively cope with distance-related uncertain information when the robot's perception of distance-related uncertain information is adapted to the environment. The effectiveness can be further improved by perceiving the environment in a human-like manner. The adaptation of the directional perception in accordance to the environment remarkably improves the overall interpretation ability of uncertain notions. User feedback is used to adapt the perception toward the user while adapting to the environment and this adaptation vastly improves user satisfaction. Methods have also been proposed to interpret the uncertain information in relation to relative references and the methods are capable of replicating human-like behavior. Furthermore, the information conveyed though pointing gestures that accompany voice instructions is fused to further enhance the understanding of the user instructions. This fusion significantly reduces the errors in interpreting the uncertain information. Furthermore, it reduces the number of steps required to navigate a robot toward a goal. A vast research gap is still remaining in this particular research niche for future developments and hence possible future improvements are also synthesized.

***Keywords*-Understanding Uncertain Information; Human-Friendly Robotics; Human-Robot Interaction; Social Robotics, Service Robotics**

DEDICATION

To my beloved parents

ACKNOWLEDGMENTS

It is with great pleasure that I acknowledge the support and contribution of all the mentors, guides, and colleagues who helped me to successfully complete my PhD thesis. There are far too many to enumerate here, but I will try my best to explicitly acknowledge some of them who deserve special mention.

First, I would like to express my sincere gratitude to my supervisor Dr. Buddhika Jayasekara. I was privileged to work under the guidance of him. I thank him for his support, encouragement and patience made my success reality. Without his tireless efforts, this thesis would never have been accomplished.

Then, I would like to extend my gratitude to my progress review panel, Dr. Ruwan Gopura, Dr. Chandima Pathirana and Dr. Thilina Lalitharatne for their insightful comments, suggestions and encouragements. I specially acknowledge the efforts put by the thesis examination panel, Prof. Dileeka Dias, Prof. Chandimal Jayawardena, Dr. Achala Pallegedara, Prof. Nalin Wickramarachchi, and Prof. Rohan Munasinghe and thankful for the comments and suggestions given to refine the thesis.

I am thankful for the colleagues in Intelligent Service Robotics Group, Arjuna Srimal, Bhagya Samarakoon, Chapa Sirithunge, Ravindu Bandara, Sachi Edirisinghe, Sahan Kodikara and Sajila Wickramarathne for their support and cooperation in pursuing the research. I am grateful for them all. I am also thankful for everyone who was a part of the Robotics and Control Lab, Department of Electrical Engineering during my time specially, Maheshi Ruwanthika and Shanaka Abeysiriwardhana.

I sincerely thank Kanishka Madusanka. Thank you for all the knowledge you have shared and the time you spent for the success of this work.I am also thankful to Achintha Abayasiri, Chamika Perera, and Isuru Ruhunage for their support.

My appreciation goes to the staff and all of my friends in the Department of Electrical Engineering for their invaluable support.

I would also like to acknowledge the commitment of the volunteers, who have participated in the experiments conducted as a part of the research. Furthermore, I would like to thank the anonymous reviewers of the publication arisen from this work for their constructive comments for sharpen the pathway of the research.

I would like to extend my deepest gratitude to my family, my parents and siblings for helping me in making this endeavor a success. My friends, too numerous to mention, who simply being with me in successful realization of this thesis. Finally, I would like to thank everybody, who was important in making this endeavor a success, as well as expressing my apology that I could not mention one by one.

This work was partially supported by the University of Moratuwa Senate Research Grants SRC/CAP/14/16 and SRC/CAP/16/03.

Furthermore, some of the conference publications originated from this thesis were partially supported by travel grants awarded by IEEE Robotics and Automation Society, IEEE Computational Intelligence Society, European Society for Fuzzy Logic and Technology, and National Science Foundation of Sri Lanka.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

An intelligent service robot is a machine that is able to gather information from the environment and use its knowledge to operate safely in a meaningful and purposive manner [1]. Recent developments of intelligent service robots open up new areas of robotic applications such as healthcare [2,3], rehabilitation [4,5], caretaking [6,7], assistance [8,9], education [10,11] and entertainment [12,13]. A few examples of such service robots are shown in Fig. 1.1. Particularly, intelligent service robots are being developed to use as assistive aids for elderly or disabled people [14–17] as a solution for the widening gap between the supply and demand of human caregivers, which will create profound complications in socioeconomic behaviors of the society [18,19].

The intelligent service robots used for these kind of emerging areas of robotic applications, are anticipated to have direct interactions with human users in domestic environments where most of the users are in the non-expert category. Hence, the interaction between the service robots and the human users are preferred to be human friendly in order to provide a sophisticated service to the users [23–26]. Human friendly robots should possess human-like interaction abilities and the dream for a perfect service robot obviously depends on that. Availability of the human-human like communication abilities in human-robot interaction would enhance the overall interaction between the human user and the non-human robot partner and this would eventually increase the satisfaction of the user [23,27]. In contrast, development of human friendly interactive features in service robots is complicated for the reason that social cognitive features of

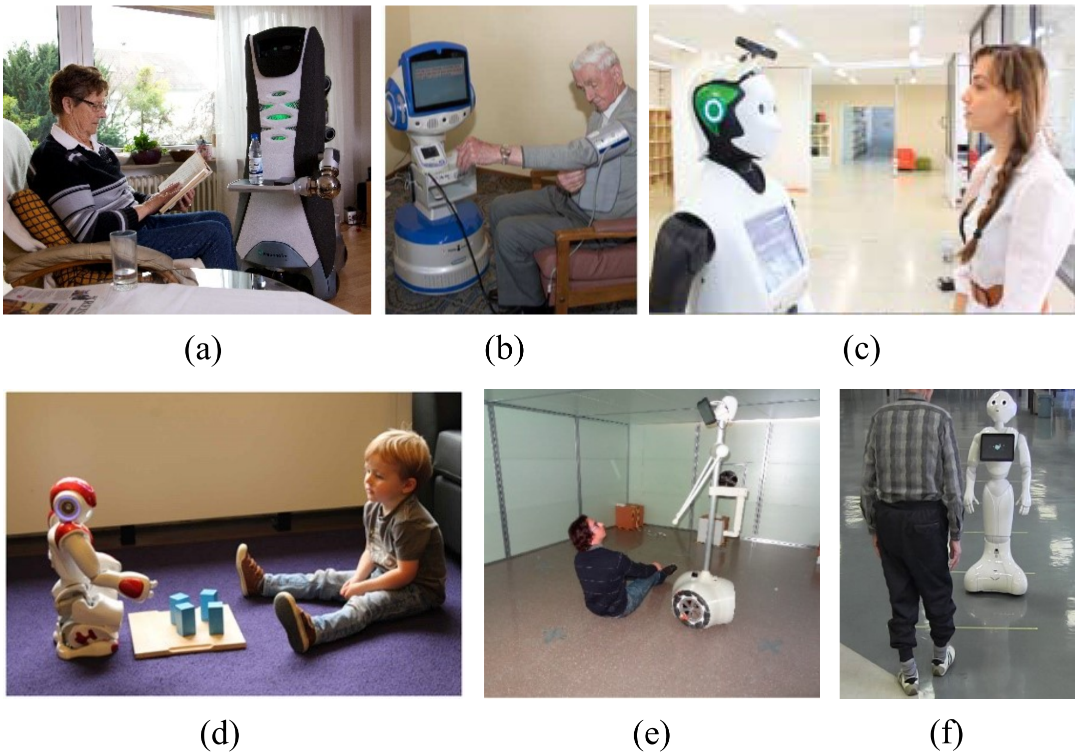


Figure 1.1: Examples of service robots used in emerging areas of robotics applications. (a) Care-O-bot 3 supporting elderly people at home.[Image courtesy: www.care-o-bot.de] (b) HealthBot [3] with a patient [Reprinted with permission ©2016 IEEE] (c) REEM robot [20] interacting with a human [Reprinted with permission ©2015 Taylor & Francis] (d) A robot tutor interact with a preschool child [21]. (e) Assistant Personal Robot (APR) [9] detecting a fallen person. (f)Humanoid service robot Pepper [22] with a older person.

the human beings should be incorporated into the robots [23].

Voice communication is one of the main communication modalities used by humans to convey instructions to the peers [28]. Therefore, human like voice communication capability of robots will enhance the overall interaction between robots and their users. This will eventually increase the rapport between the users and the robot assistants; the users can gain a more sophisticated service from the robot companions [24, 29]. Precise quantitative information is not conveyed through the voice communication and the natural voice instructions and responses often include uncertain information, lexical symbols and notions that have to be interpreted for clear comprehension. For example, a human user prefers to use the command, “move little bit forward” rather than the command, “move 40 centimeters forward” in a situation similar to the scenario shown in Fig. 1.2(a). The quantitative meaning of the term “little” has no definitive distance and the quantitative meaning depend on various factors. In here, the quantitative distance meant by the user may be in the order of 40–80 cm. However, in the situation shown in Fig. 1.2(b). The robot is commanded, “place it little bit away from the box” in order to place the cup on the table. The quantitative meaning of the term “little” in this kind of situation would be approximately in the order of 5–15 cm and the quantitative distance meant by “little” in this situation is clearly different from the earlier situation. Furthermore, the scenarios show in Fig. 1.2(c) and Fig. 1.2(d) show that the same quantitative size could be referred using completely controversial language descriptors in two different situations. Fig. 1.2(c) shows a situation where a baseball is surrounded with two golf balls and most probably in this situation the size of the baseball will be referred as “large” in a language instruction issued by a person. In scenario shown in Fig. 1.2(d), the same baseball is surrounded with a soccer ball, football and a basketball. Even though, quantitatively the size of the baseball is the same in both the situation, a person will refer the size of the baseball as “very small” in the second scenario.

However, humans have the ability to interpret a reasonable quantitative value for such uncertain terms. These kind of uncertain terms are also referred to as fuzzy linguistic information or qualitative terms. Even though the quantitative meaning of uncertain terms such as “little”, “far”, “high” and “large” depends on various factors, uncertain terms are involuntarily included in the voice instructions, suggestions and responses because of the unbridled cognitive ability of humans to understand the quantitative meaning of such terms based on the factors which affect the meaning. Therefore, the cognitive ability of a service robot to understand and appropriately respond to uncertain information in voice commands and responses is mandatory in order to provide human friendly assistance to the user.

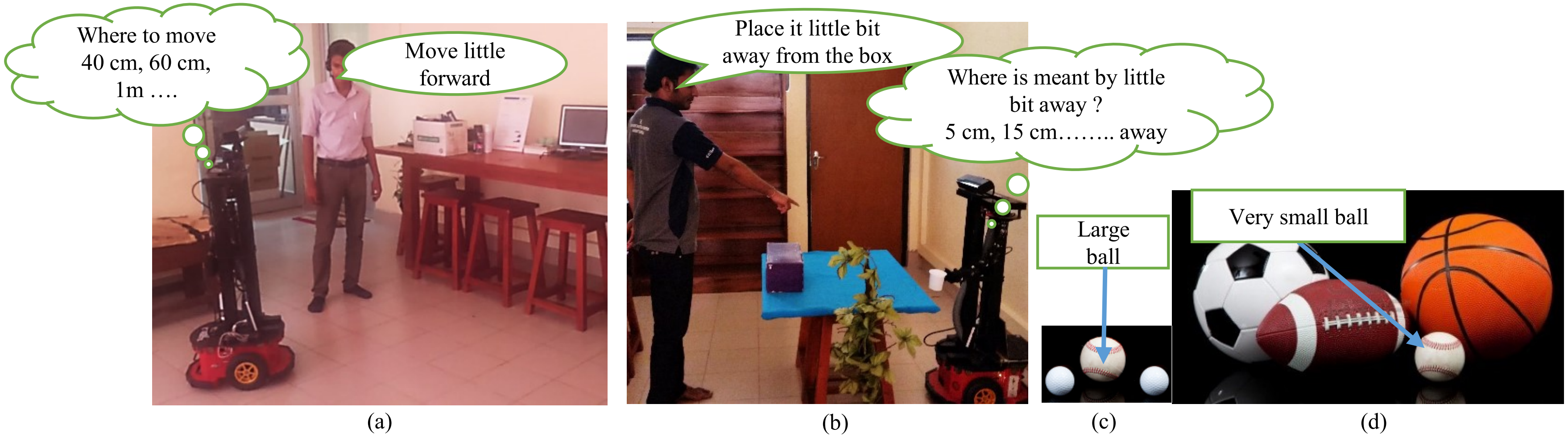


Figure 1.2: This shows example scenarios where the uncertain information is used for a purposive task and how the meanings alter in different situations. (a) shows a situation where a robot is commanded to move little bit forward by a human user. (b) shows a situation where a robot is asked to place a cup little bit away from the box on the table. (c) shows a scenario where a baseball is surrounded with two golf ball. (d) shows a situation where the same baseball is surrounded with a football, soccer ball and a basket ball.

1.1 Problem Statement

As depicted in Fig. 1.3, the robot should be capable of inferring the meaning of information conveyed from a voice instruction to perform the exact requirement of the navigation task requested by the user. This sort of navigation voice instruction may consists of uncertain lexical notions in relation to distances, directions, references, path, positions etc. Therefore, this uncertain information must be interpreted effectively and succinctly to fulfill the request of the user. The exact meaning of the uncertain information depends on various factors such as environment, experience, context and the user. Furthermore, non-verbal instructions such as gestures may be accompanied with the navigation voice instructions for enhancing the idea transferred to the peer; the meaning of the uncertain information may also depend on the non-verbal instructions associated with the voice instructions.

The overall interaction and the rapport between the user and the robot obviously depend on the ability of the robot in correctly identifying and reacting in these sort of scenarios. Therefore, this thesis investigates the methods for resolv-

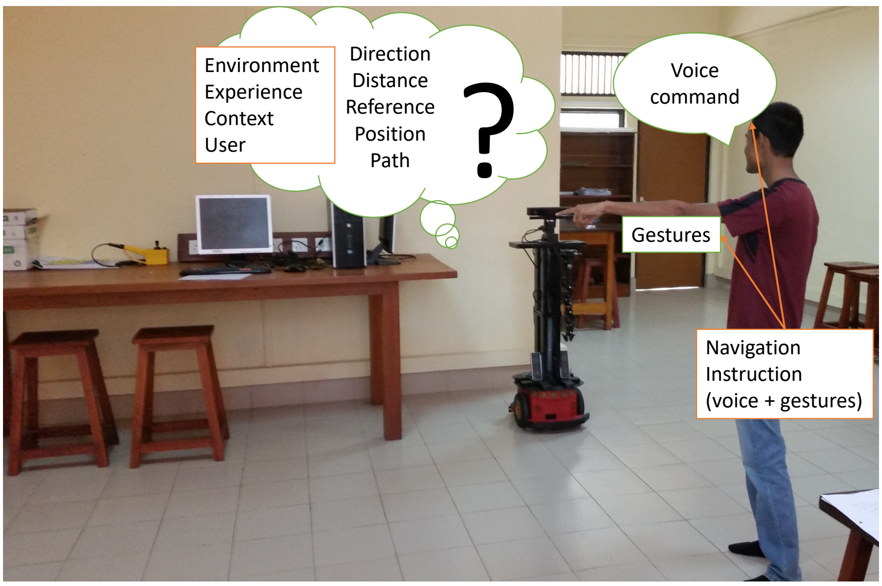


Figure 1.3: Usage of uncertain information in navigational command

ing spatial ambiguities arisen due to the inclusion of uncertain information such as “little”, “large”, “few” and “far” in navigation instructions for improving the human-robot interaction.

1.2 Thesis Contributions

This thesis contributes toward addressing the above-outlined issues by developing methods of interpreting uncertain information in relation to distances, directions and references in navigational commands.

This has been archived by resolving the issues in:

- Interpreting distance-related uncertain information by adapting the perception according to the environment
- Interpreting fuzzy directional notions by adapting the perception according to the environment
- Personalizing the perception of uncertain information while adapting according to the environment
- Interpreting uncertainties in relation to the relative references
- Adapting the perception of uncertain information according to the information conveyed non-verbally instructions

1.3 Thesis Overview

This section provides an overview of the succeeding chapters of the thesis.

Chapter 2 outlines the current state of the art in dealing with uncertain information in language instructions by the robots and systems. The existing approaches for dealing with uncertain information available in the literature have been critically analyzed in order to identify the limitations of the current state of the art and the possible future developments. The limitations of the existing system are summarized and the possible future directions are synthesized as the contribution of the review.

Chapter 3 provides an overview of an intelligent service robot named Moratuwa intelligent Robot (MIRob) that has also been developed as a part of this research. An overview of the hardware system used for the validation of the proposed concepts is explained at the beginning. Then, the functional overview of the system is briefed including all the auxiliary modules that are required for full filing the requirements of navigation instructions. Moreover, the work presented in this chapter explains the supportive modules used in the work presented in succeeding chapters. The terminologies and the names of the modules defined in this chapter will be used in subsequent chapters in this thesis. The developments of the Interaction Management Module (IMM) and the Robot Experience Model (REM), which are useful in the process of handling the interaction with the user and evaluating the uncertain information , are the main contributions presented in this chapter. Particulars on experimental validation of these auxiliary modules are also presented.

Chapter 4 contributes by proposing methods of adapting the robot's perception of distance-related uncertain information based on the environment. The chapter begins with an explanation about the rationales behind the proposed method of adapting the perception. A module called Distance Interpreter (DisI) is proposed for assigning quantitative values for the distance-related uncertain terms in navigation instructions. This module is capable of adapting the robot's perception of distance-related uncertain information according to the characteristics of the environment. Particulars on experimental validation of the proposed DisI in adapting the perception according to the environment are presented.

Chapter 5 proposes a method for enhancing the interpretation of the fuzzy notions in motional and positional navigation command by adapting the robot's directional perception based on the environmental setting. The requirement of the adaptive directional perception for the robots instead of fixed directional perception is explained at the beginning of the chapter. Then, a module called Direction Interpreter (DirI), which is deployed for the system to adapt the robots perception of directional notions according to the environment setting, is explained. At the latter stage of the chapter, the performance improvement caused to the understanding of navigational commands by the robot due to the deployment of the proposed module for adapting the robot's directional perception is discussed with the experimental results.

Chapter 6 proposes a method for enchaining the satisfaction of the user by adapting the robots perception of uncertain information based on the current environment and the corrective feedback received from the user. The advantage of learning the perception from the user feedback is discussed at the beginning of the chapter. The reimplementation of the Distance Interpreter (DisI) with fuzzy neural networks, which are capable of concurrently adapting to the environment while learning from user feedback, is explained. Particulars on the Feedback Evaluation Module (FEM) deployed to evaluate the quantitative errors of the feedback statements are discussed. At last, the improvement of the user satisfaction due to the proposed learning ability is analyzed and presented with an experimental validation.

Chapter 7 proposes a module called Relative Uncertainty Interpreter (RUI) to interpret the uncertain information in relation to the relative references. Phrases with uncertain information associated with relative references such as large table, table left of the door and table close to the door are involuntary included in navigation instructions; hence it is necessary to deploy the RUI for effective interpretation of navigation instructions. The natural tendencies in usage of uncertain information in relation relative reference are explained in the chapter with the aid of the results of a human study. The RUI has been designed by considering the identified natural tendencies of humans. The performance and behavior of the proposed RUI has been compared against a result of a human study to evaluate the performance and the outcomes are discussed and presented.

Chapter 8 investigates a way to adapt the robots perception of distance-related uncertain information by fusing the spatial information of the environmental and the influential notions conveyed non-verbally. The rationale behind the proposed adaptation method is discussed against a system that is capable on adapting the perception based only on the environment setting. The proposed method has been implemented by fusing the notions conveyed through the pointing gestures with the aid of a fuzzy inference system. The factors considered for the design of the fuzzy inference system is detailed in the chapter. Particulars on the method adopted for extracting the point referred by a pointing gesture is also explained. Finally, the experimental validation of the proposed system is presented and discussed.

Chapter 9 proposes a method to resolve the ambiguities arisen when interpreting distance-related uncertain information due to the arrangement of the environment. The intention of the user identified through the information conveyed from pointing gestures is used to resolve these sort of ambiguities in interpretation of the navigation instructions. The rationale behind the proposed concept is explained at the beginning of the chapter. A module called Motion Intention Switcher (MIS) is proposed to deploy into the system for identifying the actual intention of the user based on the accompanied non-verbal instructions. Subsequently, the MIS shifts the perceptive distance of the system between a default hypothesis and alternative hypotheses for responding the user in an effective manner. The behavior and the performance of the proposed MIS have been evaluated and the outcomes are discussed.

Chapter 10 provides the concluding remarks of the thesis together with a concise discussion on the future directions of the work.

Chapter 2

A REVIEW ON SERVICE ROBOTS DEALING WITH UNCERTAIN INFORMATION IN LANGUAGE INSTRUCTIONS

2.1 Voice and Natural Language Communication in Human-robot Interaction

With the development of voice recognition and voice synthesis engines, studies are being carried out in order to enhance the voice communication interfaces of robotic systems [30–32]. However, many of early studies in this area have primarily focused on implementing voice communication interfaces between robots and humans, and the studies are limited to ordinary control of a robotic system with limited number of user instructions such as controlling of automated wheelchairs [33–35]. Those systems are capable of understanding simple single word commands such as “go” and “stop” which are already prerecorded in the memory of the system.

For a general-purpose service robot, capability to handle only a limited number of simple instructions is not sufficient since the number of functionalities of such robotic system is much higher [27, 36, 37]. For instance, a service robot, which works in the capacity of a nurse, has to be capable of conveying empathy to a patient when communicating sensitive information related to the condition of the patient and a domestic service robot, which acts like a customer handling agent in a service desk, should be capable of adjusting the speech based on the characteristic of the customer. In addition to that, such system would not facilitate human like service. Therefore, human-like voice communication abilities in robots are preferred for service robots specially for achieving human-like human-robot communication. In this context, service robots with human like voice communication capabilities have been developed and those robots are capable of obeying natural language user instructions and responding with natural language dialogue phrases [20, 28, 38, 39].

Natural language voice instructions, responses and suggestions often include lexical symbols and notions, uncertain terms, redundant words and prepositions. Therefore, the robotic systems with human like voice communication abilities should possess the ability to understand them appropriately. Methodologies have been developed in order to compute the spatial relations referred by the prepositions such as "behind", "at" and "near" [40–43]. The methods proposed in [40,41] are capable of distinguishing the meaning of "at" and "near"; the methods proposed in [42, 43] are capable of grounding spatial relationships in human robot interactions and the method proposed in [44] can create abstract map of the working environment based on the semantic description with prepositions. Natural language voice instructions are often inaccurate or ambiguous and exact meanings of such commands depend on the context of interest. For example, the expression, "the red ball on the table near the vase" can be considered. In here, there are two alternative interpretations for the expression; the red ball is expected to be near the vase or the table near the vase. The correct interpretation among the alternatives depends on the actual arrangement of the environment. The method proposed in [45] 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; method proposed in [46] is capable of identifying a referring object in a user instructions with the aid of pointing gestures of the user, method proposed in [47] is capable of generating gestures on a robot related to object referring communications, and methods proposed in [48] is capable of fusing information from multiple modalities. Knowledge acquisition and symbol grounding through human-robot multimodal interactions have also been studied [49].

The above-mentioned methods are capable of interacting with natural language voice instructions and responses to some extent. However, the systems lack the ability of understanding uncertain information in language instructions and the methodologies for dealing the uncertain information in language instructions have not been covered in the scope of those work. Uncertain information is often included in voice instructions, responses and suggestions involuntarily in interactions as explained in chapter 1. The core contribution of this study is to investigate the methodologies used in robotic systems in order to understand the uncertain terms in language communication. Therefore, a comprehensive exploration of those systems is given in section 2.2.

2.2 Current Status: Robots Dealing with Uncertain Information in Language Instructions

As explained earlier in chapter 1, meanings of uncertain terms depend on several factors. The existing methods are capable of adapting the perception of robot about uncertain information based on different entities such as environment, experience and context. Thereby, the existing methods for understanding uncertain information are categorized primarily based on the adaptation entities in order to critically examine them.

2.2.1 Early Developments and Approaches

There have been many psychophysical studies of the perception of distance and related cognitive issues [50–52]. These studies have revealed the characteristic of the distances related cognition of human beings such as knowledge of relative location, asymmetry of cognitive distances and sources of distance knowledge. However, these concepts are limited to cognitive science and the studies have been mainly utilized to understand concept such as cognitive distances in urban environments.

Dutta [53] proposed a concept to represent spatial constraints between a set of objects given imprecise, incomplete and possibly conflicting information regarding them. Furthermore, Clementini et al. [54] developed a qualitative model for representing the positions of objects and for performing spatial reasoning as a qualitative replacement of quantitative vector algebra. However, those concepts do not directly deal with interpreting uncertain information and mostly concepts are limited to understanding of simple qualitative representations such as if object A is back of B, then B is in front of A. In adition, the concept has not been implemented on real systems and has been limited to mathematical modelling.

2.2.2 Systems with Predetermined or Fixed Interpretations

A method for communicating between robot and human using spatial language have been developed [55]. The system is capable of generating linguistic spatial description about the surrounding environment. For example, it can generate the dialogue, “The box is behind me. The object is far.” Those dialogues 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 [56]. The direction descriptors are generated categorizing the space around the robot in to 16 sub directions. The distance descriptors are generated based on the distance to the object from the robot and the distance categorization is carried out as given in Fig. 2.1. Therefore, categorization of the direction and distances are fixed with hard boundaries and the system does not possess the ability to consider the fuzziness, inherited by linguist descriptors that will eventually degrade the performance of the system.

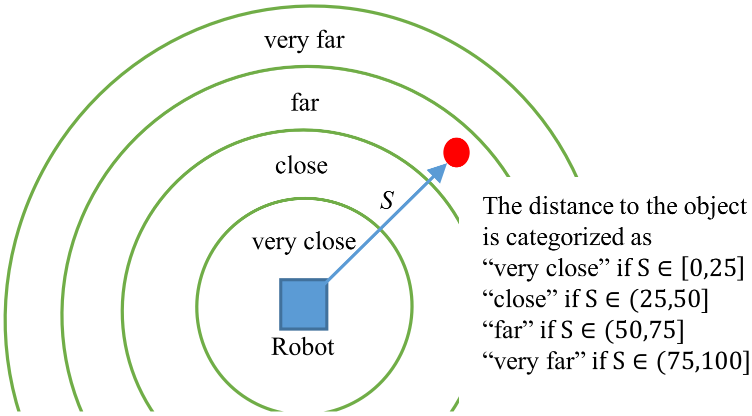


Figure 2.1: This explains the distance categorization done in the systems proposed in [55,56] in order to generate linguistic spatial descriptors about the surrounding objects.

Methodologies for controlling of a robot using information rich natural spoken user utterance have been studied with the intention of handling natural language voice instructions with fuzzy implications related to velocity of the robot while ignoring the redundant words in natural language expression [57,58]. For instance, the command "Robot, please go very fast" can be considered. In the example command the words, "Robot" and "please" are senseless within the operation domain of the robot, and only the words "go" and "very fast" are associated with the functions of the robot. The concepts are capable of ignoring the redundant words as well as interpreting fuzzy implications in natural language voice commands in order to respond appropriately. However, the concepts are not capable of identifying the context grammar and hence the systems cannot differentiate 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 of fuzzy linguistic information is defined as a linear modification factor based on the current state of the robot (i.e. the current velocity of the robot) as shown in Fig. 2.2. The desired velocity is calculated as $Desired\_Velocity = Velocity\_Factor \times Current\_Velocity$ by obtaining the corresponding velocity factor at the current speed from the graph illustrated in Fig. 2.2. The linear modification factors have been defined based on the argument that the phrases like "very fast" has a low significance when the machine

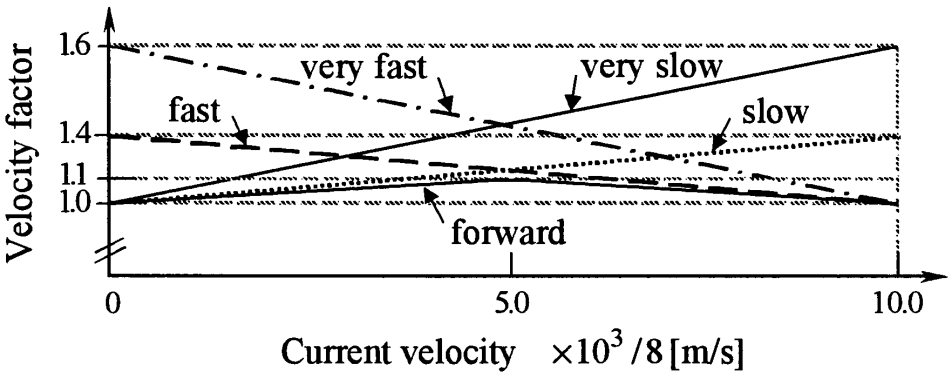


Figure 2.2: This shows the linear modification factors for obtaining the desired velocity by interpreting uncertain velocity instruction in the work proposed in [57]. Reprinted with permission ©IEEE 2004

comes close to the maximum velocity and vice versa for a phrase like "very slow". Furthermore, the linear modification factors are fixed. Hence, the output of the system is predetermined for a particular state.

A robotic aid system, which consists with a fuzzy command interpreter, has been developed for feeding the physically handicapped [59]. This system is capable of interpreting crisp values for fuzzy linguistic terms in user commands according to the current context. A fuzzy inference system is utilized and it generates a crisp output by evaluating the difference between the robot's position and the user's position as explained in Fig. 2.3. The difference between the user's position and the robot's position is calculated based on the coordinates with respect to the reference frame. The calculated position difference and the uncertain descriptor are fed into the fuzzy inference system in order to generate the crisp coordinates of the destination position. The system has been designed based on the insight that when the distance between the robot and the user is high, the command, "move closer" will move robot by a large distance towards the user to get it to a closer position and if the distance difference is small, then the movement will be a small distance because the robot is already in a closer position. This system enables the users to have much friendlier interface to instruct the robot. Even though 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, details of the fuzzy inference system in the proposed interpreter is not revealed.

2.2.3 Robotic Systems that Adapt the Perception According to the Environment

Uncertain information related to the spatial information such as sizes of objects and distances is often used in typical assistive tasks in domestic environment. The meanings of such uncertain terms obviously depend on the environmental factors. Therefore, concepts have been introduced in order to adapt the perception of robots about uncertain information based on the spatial information of the environment of interest.

A method has been introduced in order to effectively evaluate the fuzzy linguistic information in manipulation related user instructions such as "move red box little left" based on visual attention [60]. The system is capable of interpreting the quantitative distance value for the distance related fuzzy implication in a particular user command. The corresponding object in a user instruction is identified from directly mapping the lexical symbol with the object memory

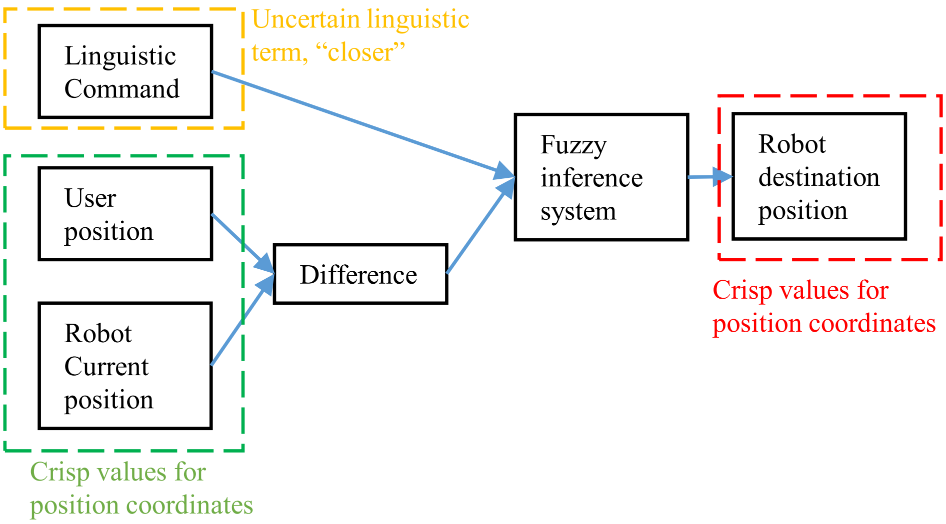


Figure 2.3: This explains the arrangement of the fuzzy command interpreter purposed in [59]. The figure is based on [59]

which considers the Hu moments [61] and RGB values as the feature set similar to the method explain in [62]. However, the object identification and movement direction related uncertain information are not handled by the system and the possible user instructions are bounded by a strict grammar model that cannot learn new patterns of user instruction. Through the visual attention, the system can perceive the working environment in order to assess the spatial arrangement of the objects in the working environment. A fuzzy inference system is used to generate the crisp distance value for the fuzzy implication by considering the average distance of the objects in the attentive vision field. In order to calculate the average distance (d_{avg}), the attentive vision field is divided into regions based on the four principle directions as shown in Fig. 2.4. Subsequently, the average distances to the surrounding object in each neighborhood are calculated. Thereafter d_{avg} is calculated by considering a higher priority for the region in the target moving direction than the other directions. Then the parameters obtained from the visual attention system are fed into the fuzzy inference system shown in Fig. 2.5. In order to evaluate the performance of the concept, variations of the evaluated distances of different fuzzy implications with the arrangements of the objects in the vision field are given as experimental results. These, results clearly indicate the ability of the system in adapting the perception according to the spatial arrangement. However, the results have not been validated against the compliance of the user. The system is only capable of interpreting fuzzy implications related to motional information and it cannot evaluate uncertain information related to positional information. As an example, it cannot evaluate the command "move blue box near to the red box" since it cannot evaluate the positional information related uncertain term "near". This is one of the major limitation of the proposed system. Furthermore, the system uses an overhead camera to perceive the environment, hence the attentive vision field is not human like.

Schiffer et al. [64, 65] proposed a method that can be used by a service robot for qualitative spatial reasoning of positional information in user instruction. The

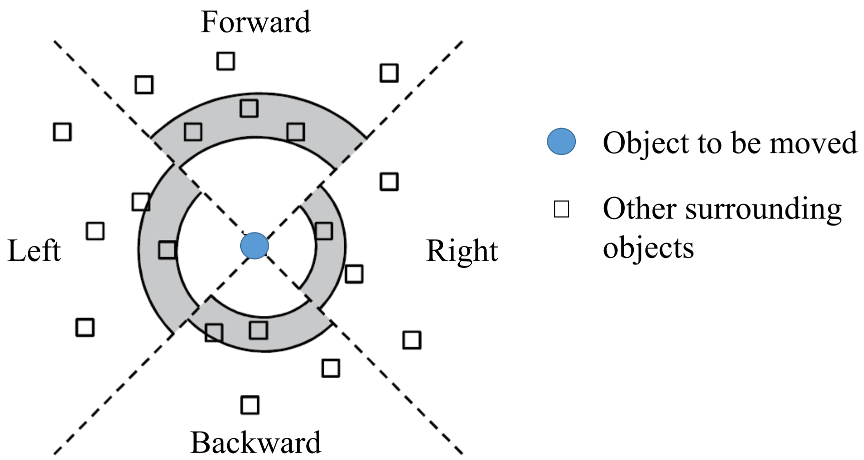


Figure 2.4: The parameter evaluation of the visual attention system proposed in [59] is explained here. The neighbourhood region of each principal direction is indicated in the shaded area. Only the objects in the neighbourhood areas are considered for evaluating the average distance by omitting other surrounding objects. The figure is based on [60].

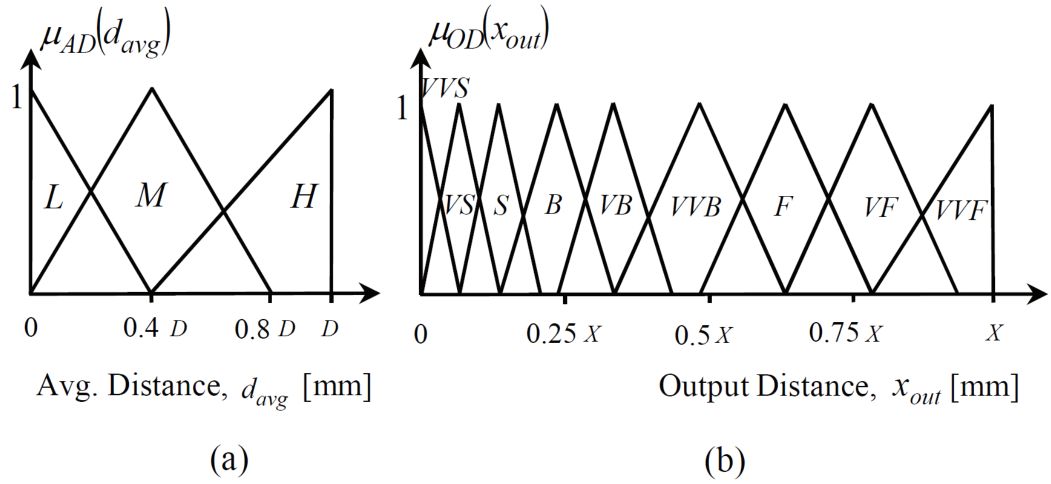


Figure 2.5: This shows the membership functions of the fuzzy inference system used in the system proposed in [60]. (a) shows the input membership function for the average distance of the surrounding objects and the fuzzy sets are adjusted according to D, which is the distance to the farthest object from all. (b) shows the output membership function of the system and the fuzzy sets in the output membership function are adjusted according to X which is the distance to the nearest object in the target moving direction. It should be noted that the fuzzy predicates that can be identified from the system (i.e. "very little", "little", "medium" and "far" according to the grammar model) are fed to the system through an input membership function with singleton fuzzy sets. The figure is extracted from [63]. Reprinted with permission ©2009 IEEE

proposed concept has been combined with a logic program language known as GOLOG [66] and a framework for reasoning about actions and changes known as situation calculus [65]. This enables the reasoning of fuzzy fluent related to the positional information in a robot operated inside a domestic environment. The basis of the reasoning method is that the fuzzy information associated to positional information in domestic environment depends on the associated frame or the point of view. The assignment of frames in an example situation is illustrated in Fig. 2.6. 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 and "far" with respect to a table in the living room has a much smaller quantitative meaning than the previous two cases. Therefore, the meanings of fuzzy terms are scaled according to the frame size, which is the size of the respective room or object such as a table. Therefore, the concept is capable of adapting the perception based on the environment. However, experimental results for variations of the interpreted quantitative values for qualitative information have not been gathered and analyzed. The adaptation entirely depends on the size of the frame and other environmental factors that influence the interpretation such as free space and object arrangements are not accounted for the adaptation. Those are the main drawbacks of their work.

2.2.4 Robotic Systems that Adapt the Perception According to Experience

Humans build up their knowledge base by acquiring knowledge through the experience. This knowledge base can be used to get an idea about the working environment, user expectations and the context. Furthermore, such knowledge acquisition enhances the capability of interpreting fuzzy linguistic information according to the current environmental context and the expectations of the user. Therefore, experience is also an important factor in adapting the perception of robot about uncertain information and systems have been developed in order to

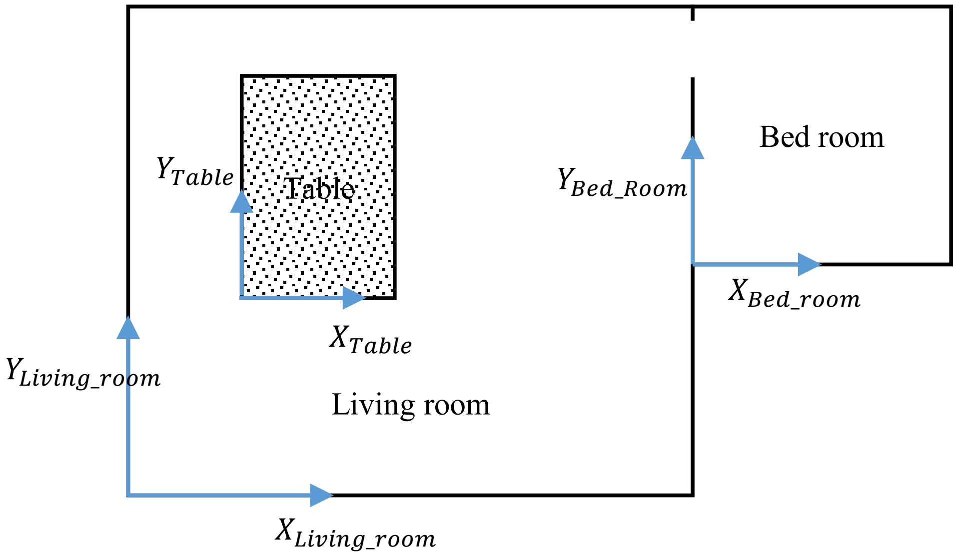


Figure 2.6: This explains the concept of frames used in [64, 65]. Each room or object has its own reference frame.

adapt the perception based on the robot's experience.

The meaning of an uncertain term depends on the immediate previous state. For example, a situation of driving a car by two persons can be considered. A person who drove the car 100 km may think driving another 10 km is a short distance while a person who drove a car 15 km may think driving another 10 km is a long distance. Based on this phenomena, the method proposed in [67, 68] assumes that the quantitative meaning of an uncertain term depends on the immediate previous movement of the robot. The proposed concept is known as fuzzy coach player system and it can be used for teaching the behaviors for robot using natural language instructions. The quantitative values for uncertain terms are interpreted by a fuzzy inference system that considers the immediate previous moment of the robot as an input. The end effector movements and single joint movements of a manipulator have been considered in [68] and [67] respectively for the implementation. The inputs and output membership functions of the fuzzy inference system used in [68] is shown in Fig. 2.7. The fuzzy sets in the membership functions are fixed and defined based on expert knowledge. Therefore, the adaptivity of the system for different conditions is hindered which is one of the

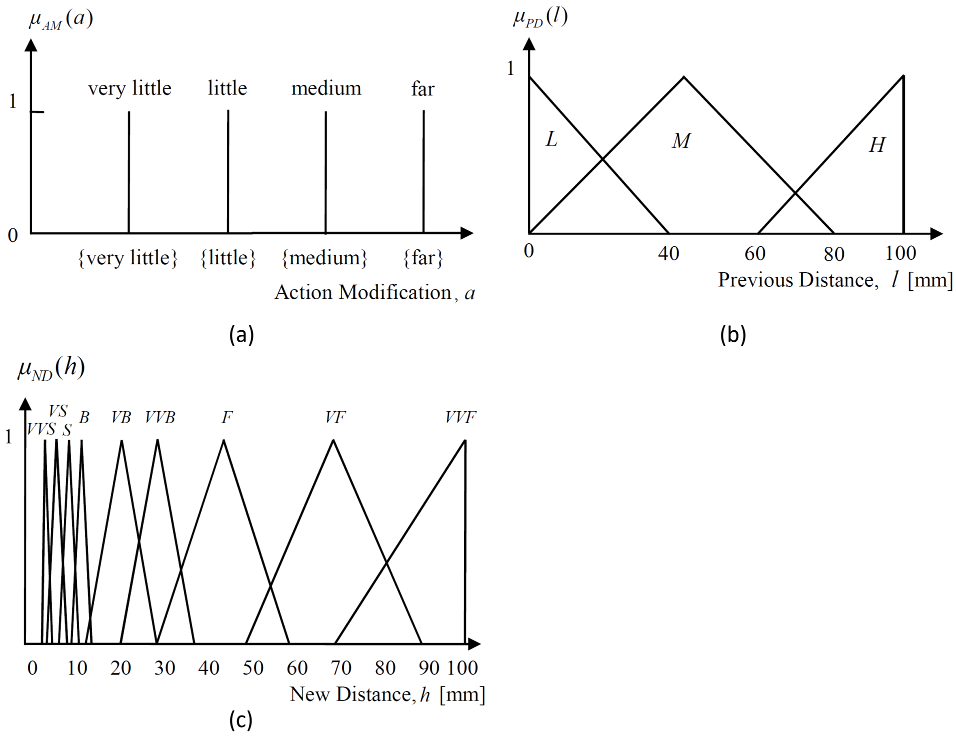


Figure 2.7: The membership functions of the fuzzy inference system used in [68] is shown here. (a) shows the input membership function for the action modifier (i.e. the uncertain term in a particular user instruction). (b) shows the input membership function for the previous movement. (c) shows the output membership function. The figure is based on [69].

major drawback of this method.

However, there are situation where merely the immediate previous state misrepresent the experience. As an example, a situation where a person has driven a car 80 km, 100 km, 70km, 90 km and 2 km for consecutive 5 times can be considered. If only the immediate pervious state is considered (i.e. only the 2km travel), his experience is not represented correctly. Therefore, a set of previous states should be accounted in order to get an enhanced assessment of the situation. Based on this, the method proposed in [70] interprets uncertain information by considering the previous set of movements of the robot. The method used the concept of internal rehearsal [71] that is an internal simulation, which simulate the ability of humans in internally perceiving and manipulating the environment, and forecasting the future [72]. The functional overview of the system is depicted in Fig. 2.8. Mainly the system consists of two sections; the fuzzy inference system and the internal rehearsal system. The functionality of the fuzzy inference system is almost similar to the fuzzy inference system used in [68] (i.e. the fuzzy inference system depicted in Fig. 2.7) even though it has been implemented as a fuzzy neural network. The fuzzy inference system is responsible for evaluating the quantitative meaning of the uncertain term in a particular instruction by considering the previous movement as similar to [68]. However, the internal rehearsal system suggests the corresponding previous movement. The internal rehearsal system consists of the Rehearsal Memory (RM), Previous Movement memory (PM) and the Rehearsal Counter. RM stores the internally simulated output value provided by the fuzzy inference system (i.e. s_r) for the suggested previous movement (PM_r). Likewise, the process will continues from $r = 1$ state to $r = N_{rh}$ state where r is the count of internal rehearsals and N_{rh} is the defined threshold limit (This indicates how many previous movements have to be accounted as the experience.) without performing any real movement. Thereafter, the simulated outputs (s_r) in the RM are integrated as given in (2.1) in order to decide the required quantified output (Y) for the movement where p_r is a constant that represents the probability of relevancy of r^{th} internal rehearsal for the final outcome. Therefore, p_r is defined in such a way that the probability of relevance decades with the time (i.e. value of p_r exponentials decades from $r = 2....N_{rh}$) as similar to human memory. Moreover, the recent previous states have a higher effect to the final output than the past states. Variations of the interpreted quantitative values for fuzzy linguist information with number of internal rehearsals have been analyzed in order to assess the performance of the proposed concept. The proposed concept has been implemented for end effector and posture controlling of a fixed manipulator. However, the system is only capable of handling a predefined set of uncertain terms and the possible user commands are bounded by a strict grammar model.

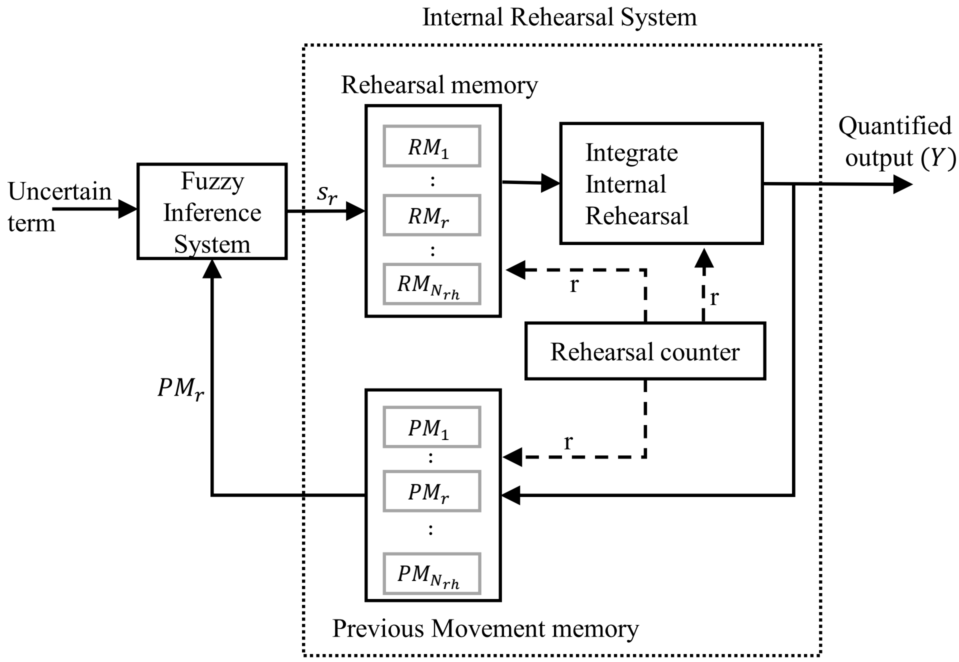


Figure 2.8: This explains the core functionality of the internal rehearsal system proposed in [70]. The figure is based on [70].

$$Y = \frac{\sum_{r=1}^{N_{rh}} s_r p_r}{\sum_{r=1}^{N_{rh}} p_r} \tag{2.1}$$

An adaptive fuzzy command acquisition network, which processes fuzzy linguistic information in spoken language commands, has been proposed [73]. The proposed concept is capable of acquiring the knowledge about fuzzy linguistic information based on the user critics. The concept has been implemented with a neural network in such a way that the system is capable of learning new user commands and on line learning. Hence, the possible user instructions are not restricted and it can acquire new knowledge while operating. However, according to its implementation, network nodes/size increases exponentially with the vocabulary size. The abilities of the system in acquiring the knowledge of fuzzy command have been verified with the experimental results. However, the ability of the system in interpreting quantitative values for the fuzzy linguistic information has not been assessed experimentally. It requires a large data set for initial training of the network and the users cannot give natural language user critics to the system. Those are the main drawbacks of the proposed concept. Furthermore, the proposed system has not been implemented on a robotic system and the concept can be applied for voice controlled robot, on-line information retrieval systems etc.

Jayasekara et al. [74] proposed a method to adapt the perception of fuzzy linguistic information based on the user feedbacks. The proposed concept has been implemented with a fuzzy neural network. The important segments and layers of the fuzzy neural network is shown in Fig. 2.9. The first layer has two types of node to acquire the inputs, the uncertain term and the previous movement. The second layer act as the fuzzification layer and the same input membership functions used in [68] and [70] are used in here also with slightly modified fuzzy sets. The third layer represents the rules by taking the algebraic product between the outputs of second layer as T-norm and the output of i^{th} node in this layer represent the firing strength of i^{th} rule (μ_i). The fourth layer links the fuzzy antecedent part to the consequent part and any node, i represents 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 the values slightly similar to the output membership function (However, a uniform distribution of fuzzy sets over the universe of discourse is considered here) of the fuzzy inference system presented in [68] and [70]. The fifth layer is the defuzzification layer and the defuzzfied output (A) is obtained from (2.2) using sum-product composition for Mamdani fuzzy systems [75] where N_R is the number of rules.

$$A = \frac{\sum_{i=1}^{i=N_R} a_i b_i \mu_i}{\sum_{i=1}^{i=N_R} b_i \mu_i} \tag{2.2}$$

The connection weights of the fifth layer of the network are adjusted based on the user feedbacks ((i.e. a_i and b_i)).This enables a more natural communication and eventually it enhances the interaction between the user and the robot. In order to evaluate quantitative values for the feedback terms, a module called

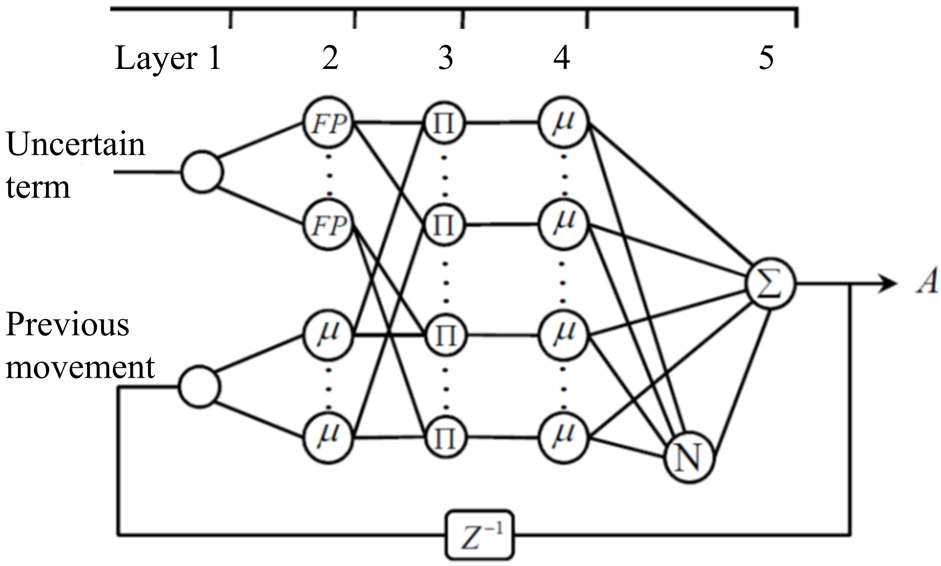


Figure 2.9: The structure of the fuzzy neural network used in [74] for interpreting the uncertain information is shown here. It should be noted that only the important segments in interpretation process are included here and hence the numbering of the layers is different from the original publication. The figure is adapted from [76].

vocal cue evaluation system has been deployed. This module has been developed with a fuzzy inference system that assumes that the quantitative meaning of feedback terms depends on the immediate previous state of the robot. The connection weights are modified though backpropagation based on the quantified error identified from the feedback of the user.

The performance of the system has been further improved by considering the willingness of the user [76] that can be used as a parameter to identify the motivation of the user to change the perception of the robot about a particular uncertain term. This parameter is evaluated by considering a series of user feedbacks. The performance improvement of the system due to the consideration of the willingness of the user for the adaptation has been analyzed experimentally by defining a performance index called user satisfactory level. The satisfactory level is the ratio between the number of feedbacks received as “good” and the total number of feedbacks. The proposed system is capable of adapting the perception of the robot about the uncertain information towards the perception of the user. The system has been implemented for controlling the end effector of a fixed robotic manipulator. The possible user commands and the feedback terms are bounded by a strict rule set. Furthermore, the system cannot evaluate the unintentional body movements of the user that can be used as feedback such as facial expression and the feedbacks have to be given explicitly in order to adapt the perception, which is an overhead duty.

All the systems mentioned in this section cannot perceive the environment through sensors, hence the systems cannot adapt the perception according to the changes in the environment. Therefore, the systems are not suitable for dynamic environment or mobile tasks since the experience is only effective for a particular environment. This is the major limitation of the systems that adapt based solely on the experience.

2.2.5 Robotic Systems that Adapt According to the Influential User Instruction

The attention can be altered based on an external stimulus such as a voice command [77] and hence attentive instructions such as “move carefully” influence the quantitative meaning of fuzzy implications in user commands. Therefore, [63] and [78] proposed a factor called attentive modification factor that can be used in order to modify the perception of fuzzy linguistic information when an attentive instruction is given. The proposed concept has been utilized with the uncertain information evaluation method proposed in [60, 74]. The users can use a set of predefined attentive instructions in order to influence the perception by varying the attentive modification factor. The attentive modification factor is a linear function (as shown in Fig. 2.10) that relies on different attentive instructions such as “move more carefully” and “move carefully” and the function is defined in a way such that it can exhibits the natural human behaviors such as fading away of the effects of an attentive instruction with the successive operation.

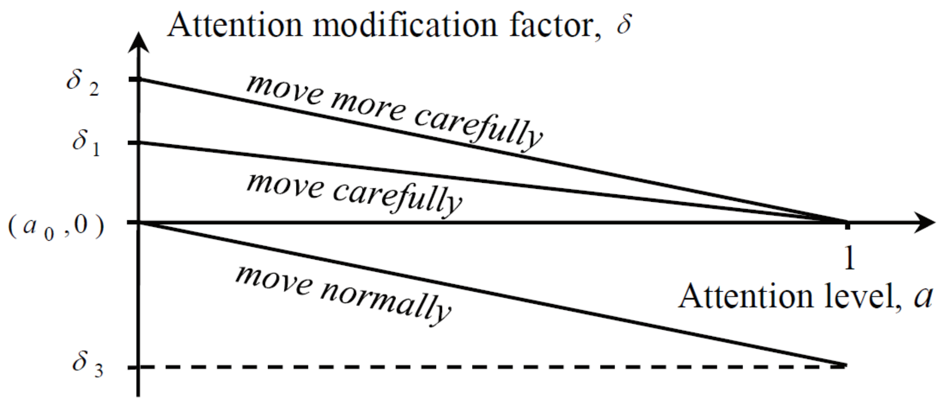


Figure 2.10: This shows the linear functions used in [63, 78] to modify the attention level. Reprinted with permission ©2009 IEEE

2.3 Limitations of the Existing Systems and the Possible Improvements

The existing systems have limitations and the performances of the existing system for understanding uncertain information are far below compared to the abilities of humans. Hence, the existing methods should be improved in order to enhance the human-robot interaction abilities. The limitations of the exiting uncertain information understanding methods have been analyzed and the possible improvements are suggested based on the following three aspects; scope, interaction and adaptation. Uncertain information links with different entities such as spatial information, time, counts etc. and such linking entities are considered as the scope. The way that the interactions between the users and the robots are taken place and the way that the perception of uncertain information is adapted are considered as the interaction and adaption respectively in the analysis. The current status of the method used for understanding uncertain information and the possible improvements are summarized taxonomically in Table 2.1.

Table 2.1: Summary of the current status of the methods used for understanding uncertain information and the possible improvements

	Current status[1]	Possible Improvements[2]
Scope	⊗ Understanding is limited to uncertain information related to • Distances in environment (e.g., [59, 60, 65, 68], [70, 74]) • Speed of movements (e.g., [57] and [73]) • Directional notions (e.g., [55]) • Object sizes,(e.g., [62] and [79]) • Joint angle (e.g., [67] and [70])	⊗ Extend the capabilities to understand uncertain information related to other aspect such as time, counts and process/task related information
Interaction	⊗ All the system are only capable of interacting using only voice communication between the robot and the user. ⊗ Very few work tried to synthesize uncertain information in vocal responses of the robot. (e.g., [55] and [79]) ⊗ All other the work focuses on addressing the issue in interpreting the uncertain information in user instructions	⊗ Extend the capabilities to evaluate the information conveyed non-verbally to adapt the perception of uncertain information ⊗ Developing methods to synthesize uncertain information in vocal responses of the robot in an adaptive manner.
Adaptation	⊗ Systems are capable of adapting the perception of uncertain information based on • Environment ∘ Frame size (e.g., [65]) ∘ Average distance between objects (e.g., [60]) • Experience ∘ Immediate previous movement (e.g., [67] and [68]) ∘ Set of previous movement (e.g., [70]) ∘ User critics (e.g., [73] and [74]) • Influential user command (e.g., [63] and [78]) ⊗ Fuzzy logic (type I) and fuzzy neural networks are often used. • Fuzzy logic (e.g., [60, 67, 79] and [59]) • Fuzzy neural networks (e.g., [70, 73] and [76]) ⊗ Performance evaluation done through user studies • User satisfactory level (e.g., [76]) • Response of the humans (e.g., [79])	⊗ Consider multiple entities to concurrently adapt the perception. e.g., • Experience and Environment • Environment and User critics • Environment and contextual knowledge of different objects and tasks ⊗ Consider more environmental factors for adapting the perception. e.g., • Available free space • Arrangement of obstacles ⊗ Perceive the environment in a human like manner for improving the perceiving effectiveness. e.g., • Use human like vision system instead of overhead cameras • Consider the human attention focusing to extract/identify the key environment parameters ⊗ Adapt the perception considering the specify knowledge of a particular context. e.g., • Common properties of an arrangement of a lunch table • Danger of hot item or flames ⊗ Investigate the possibility of using fuzzy type II systems for interpreting the uncertain information ⊗ Introduce an objective performance measurement index

[1] It should be noted that only the key publications are given as the examples.
[2] Most of the possible improvements are synthesized here. However, it should be noted that the all the possible improvements listed in here are not addressed within the scope of this thesis.

2.3.1 Scope

Uncertain terms related to many different entities are involuntarily included in voice instructions and suggestions. However, most of the existing systems are limited to handling uncertain information related to distances in environment [65, 68, 74], speed of movements [57, 73], directions of objects [55], object sizes [62, 79], and joint angles of the manipulators [67, 70]. Uncertain terms related to the other aspect such as time, counts, and processing tasks are not 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, it would be a challenging task since the factors, which effect the meaning of such uncertain information have to be identified since previous studies have not revealed those information.

2.3.2 Interaction

The present systems developed for interpreting uncertain information are only capable of interacting with humans using only voice communication. Hence the interactions are unimodal and the systems are not capable of grabbing information conveyed through interaction modalities such as hand gestures, facial expressions, and body movements. The information conveyed through these modalities other than voice can be used as supportive aid for enhancing the understanding of uncertain information included in voice instructions. Furthermore, facial expressions and sub conscious body movements can be used as substitute for voice feedback in the systems that adapt based on the user critics such as in [74]. This will eventually reduce the overhead work of the user and hence interaction will be improved.

Inclusions of uncertain terms in vocal responses of robots will enhance the human like communications abilities in the robots. However, most of the present systems are only capable of interpreting uncertain information in user instructions and only limited numbers of studies have been carried out to generate uncertain terms in vocal responses of the robot. The system proposed in [55] has fixed meanings for uncertain terms in responses and the method proposed in [79] is capable of synthesizing uncertain terms related to sizes of objects by adapting the perception based on the visual attention. Therefore, capabilities of the existing system are not sufficient in this regard and studies should be carried out in order to develop methods to generate uncertain terms in vocal responses of the robots effectively.

2.3.3 Adaptation

According to the analysis of existing literature, different methods have been used by the existing system to adapt the perception of uncertain information. These methods used different entities and different Artificial Intelligence (AI) techniques in order to adapt the perception. Therefore, limitation of the existing systems and the possible improvements in the adaptation methods are analyzed separately considering the adaptation entities and artificial intelligence techniques. Furthermore, the performance evaluation methods used in the existing approaches are also discussed.

Adaptation Entities

Existing methods are capable of adapting the perception based on different entities that affect the meaning of uncertain information. As examples methods proposed in [60, 65] are capable of adapting the perception based on the environment; methods proposed in [68, 70, 74] are capable of adapting the perception based on the experience.

The systems that adapt the perception based on the environment are capable of adapting the perception based on the spatial factors of the environment. For instance, method proposed in [65] uses room size; method proposed in [60] uses average distance between the objects. The vision feedback is used only in the system proposed in [60]. However, it does not possess the stereoscopic vision and the system uses an overhead camera that has a completely different view of the environment compared to that of a human. Therefore, the system has drawbacks in interpreting uncertain information since the environment perceiving ability is limited and not human like. In order to improve the effectiveness of the uncertain information understanding capabilities of the robot, human-like vision attentive mechanism should be incorporated into the robot. Furthermore, other environmental factors that influence the meaning of uncertain information are not utilized by the existing methods. However, studies have not been also carried out to identify the influence of the environmental factors that alter the meaning of uncertain terms. Therefore, investigations need to be carried out in order to identify the influential environment factors and those effective factors should be used in order to adapt the perception of robot on the uncertain information.

AI Techniques

Most of the systems that can understand the uncertain terms in user commands utilize fuzzy inference systems in order to interpret quantitative values for the uncertain terms (e.g., [60], [65] and [59]). In order to provide learning ability, fuzzy neural networks are utilized in systems that can adapt the perception of uncertain terms based on the user critics (e.g., [73], [76] and [74]). The fuzzy logic systems are used more often in here due to their ability in effective modelling of the knowledge of humans beings in robotic systems without the knowledge of the underlying dynamics [75, 80] and most of the behaviors related to human-robot interaction domain [81, 82].

The fuzzy inference systems that utilized in these systems are fuzzy type I systems. However, Mendel [83] showed that interval type II fuzzy sets could better represent the 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 the uncertain information of the robots. Furthermore, there is a possibility of using general fuzzy type II for improved performance since the recent development of computationally effective algorithms for implementation type II fuzzy inference systems [84]. Therefore, it would be interesting to model the systems using general type II fuzzy sets or interval type II fuzz sets in order to identify the performance gain despite of the implementation and computational complexity.

As explained in section 2.3.3, the perception of uncertain information should be adapted according to the context. In order to identify the context, fuzzy Naive Bayesian network could be used as explained in [6]. For fusing multimodal interactions, there is a possibility of using Bayesian networks similar to the methods explained in [46]. Therefore, such methods could be adopted as supportive aids for the interpretation process of uncertain information.

Performance evaluation

Few methods that can be used in order to evaluate the performance of the systems can be found from the available literature. An index called user satisfactory level [74] is used in comparing the adaptation capability of the robot towards the perception of the user. The user satisfactory level is calculated based on the agreement of the user about the responses of the robot in successive user instructions. Actually, the user satisfactory level is the ratio between the number of cases accepted by the user and the number of cases considered. In [79], a human study has been conducted by asking the participants to rate the sizes of objects using a linguistic term in different scenarios and the results of the human study have been compared against the linguistic terms synthesized by the robotic system. However, user studies that can be conducted in order to evaluate the human-robot interaction are highly subjective due to the subjectivity of human participants. Therefore, the human studies should be carried out in a way that experimental results can provide a basis for generalizability and recommendations for designing, planning and executing human studies for HRI can be found in [85]. Convergence ability of the learning function is also analyzed in method proposed in [70] in order to evaluate the performance and variations of such parameters of the intelligent systems can also be good choices.

2.4 Summary

The chapter presented a review on service robots dealing with uncertain information in language instructions and responses. Service robots are being developed in order to cater the demand 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 is one of the predominant interaction modalities used in order to convey information between peers. Hence, service robots with human like voice communication abilities could provide a better service. However, the natural voice instructions do not convey precise quantitative information and humans mostly prefer using uncertain terms, lexical symbols and notions rather than more precise quantitative values. Hence, the ability to interpret uncertain information is mandatory for a human friendly service robot.

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

Fuzzy logic and fuzzy neural networks are often used in order to interpret the uncertain information in voice instructions by most of the existing methodologies due to their ability in modelling natural tendencies of humans. The fuzzy inference systems and fuzzy neural networks are capable of effectively interpreting the uncertain information to a greater extent. However, there are limitations in the existing systems in interpreting uncertain information in human like manner.

The limitations of the existing systems have been identified and the possible future improvements are synthesized in this chapter as contributions. In summary, capabilities of the existing systems are far below compared to the cognitive abilities of human beings in understanding the uncertain information. Furthermore, it was found that minimal research had been done in this special research area and there is a wide research gap to fill out.

Chapter 3

SYSTEM OVERVIEW

3.1 Physical Overview

Moratuwa Intelligent Robot (MIRob) has been developed as a domestic service robot. The base of the MIRob consists with a Pioneer 3DX mobile robot platform developed by Adept MobileRobots[1]. It is a differential drive mobile robot with 500-tick encoders. The robot has two sonar sensor arrays one in the front and one in the back. Each sonar sensor array consists with eight sonar sensors, which have sensitivity range from 10 cm to 5 m. The base can reach maximum speed of 1.2 ms^{-1} and carry a payload of up to 17 kg. In addition to that, it has an inbuilt gyroscope for error correction in navigation. A mobile robot especially when operating in human populated environments, needs to avoid possible damages to furniture or humans because of collisions and has to be safe in this regard. As a further safety measure, it is fixed with front and rear facing bumpers to detect collisions. An aluminum structure has been placed on top of the base to increase the height of the robot to a match the height of human beings. Total height of the MIRob is 110 cm. Cyton Gamma 300 manipulator developed by Robai[2] is installed on the robot to handle objects. The manipulator has 7- DOFs and 1 DOF gripper. It can handle a maximal payload of 300 g. Full reach of the manipulator is 53.4 cm and the maximum opening of the gripper is 3.5 cm. On the very top of the robot, a Kinect version 2 motion sensor [3] is mounted with a

[1]www.mobilerobots.com
[2]www.robai.com
[3]www.wikipedia.org/wiki/Kinect

Flir-E46[4] pan-tilt unit. The pan tilt unit facilitates gaze changes of the robot. The MIRob is shown in Fig. 3.1.

The embedded motion controller of the robot automatically performs velocity control of the robot and provides robot state and control information including a position estimate of the robot in space, battery charge data, sonar range sensing data etc. An embedded computer with 2GB RAM and Intel® Core™ 2 Duo 2.26 GHz processor is integrated into the robot base for performing high level processing and controlling task such as, image processing, voice recognition and understanding, tasks planning and decision making. For communication, a WLAN adapter capable of using IEEE 802.11a/b/g is installed. In addition to that, the robot consists with stereo speakers and microphones[5] to play sound and voice recognition. The robot is powered by three hot-swappable 9Ah sealed led acid batteries and the battery power lasts for approximately 3 hours of continuous operation at full charge.

3.2 Functional Overview

Overall functionality of the system is depicted in Fig. 3.2. The system is capable of interacting with the user through voice communication and the actions of the robot. Voice commands are recognized and analyzed by the Voice Recognition and Understanding Module[6]. Voice recognition is implemented using the Speech Recognition 3.1 [7] library. Voice responses are generated by the Voice Response Generation Module, which is a text-to-speech converter implemented using the Microsoft Speech API [8]. Basic dialogue and grammar patterns, keywords, and lexical symbols are stored in the language memory.

[4]www.flir.com

[5]A wireless microphone is used to improve the accuracy of the voice recognition.

[6]In situations where the voice recognition accuracy is poor, a wizard is used to convert the voice into text.

[7]www.github.com/Uberi/speech_recognition

[8]www.en.wikipedia.org/wiki/Microsoft_Speech_API

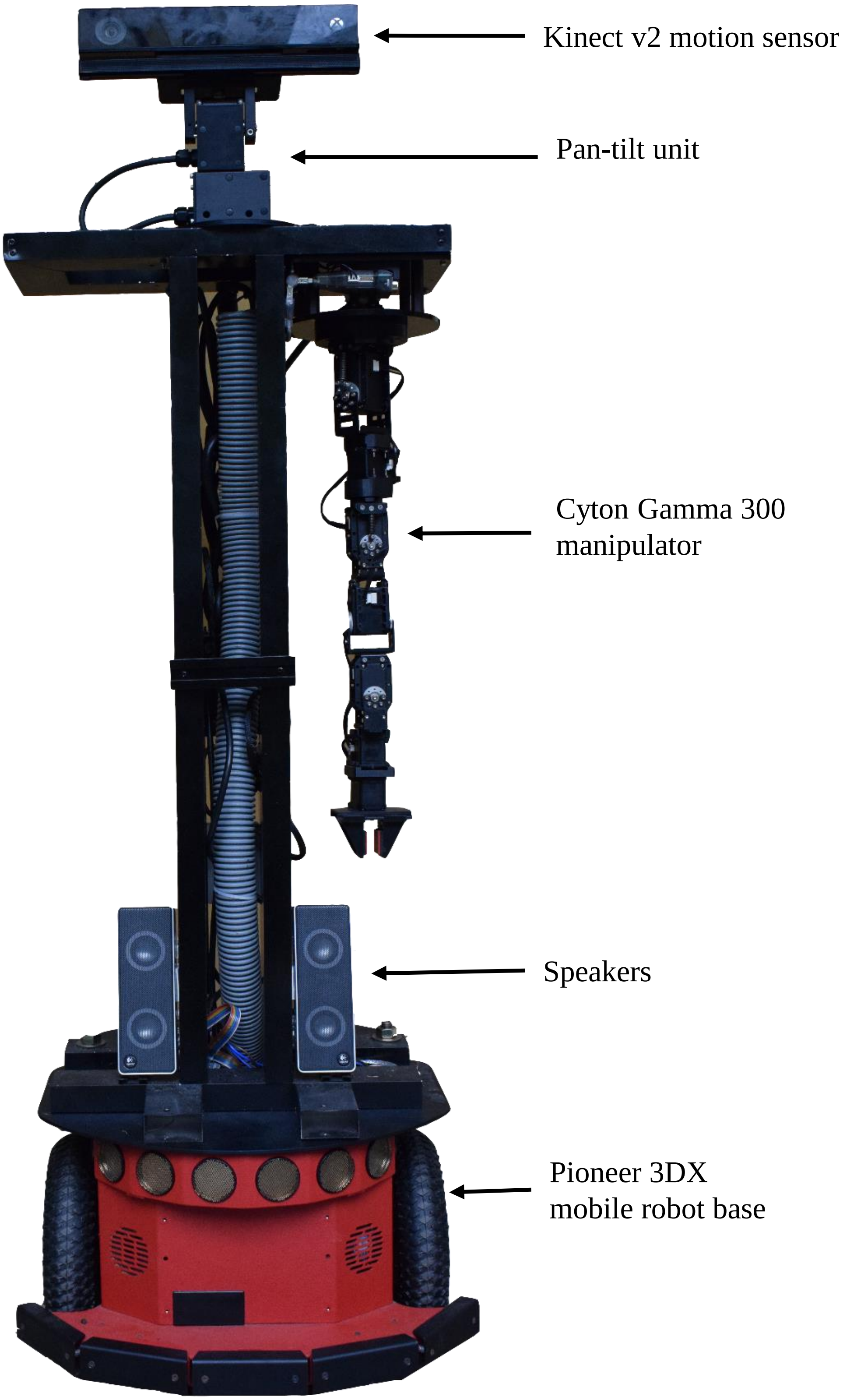


Figure 3.1: Moratuwa Intelligent Robot (MIRob).

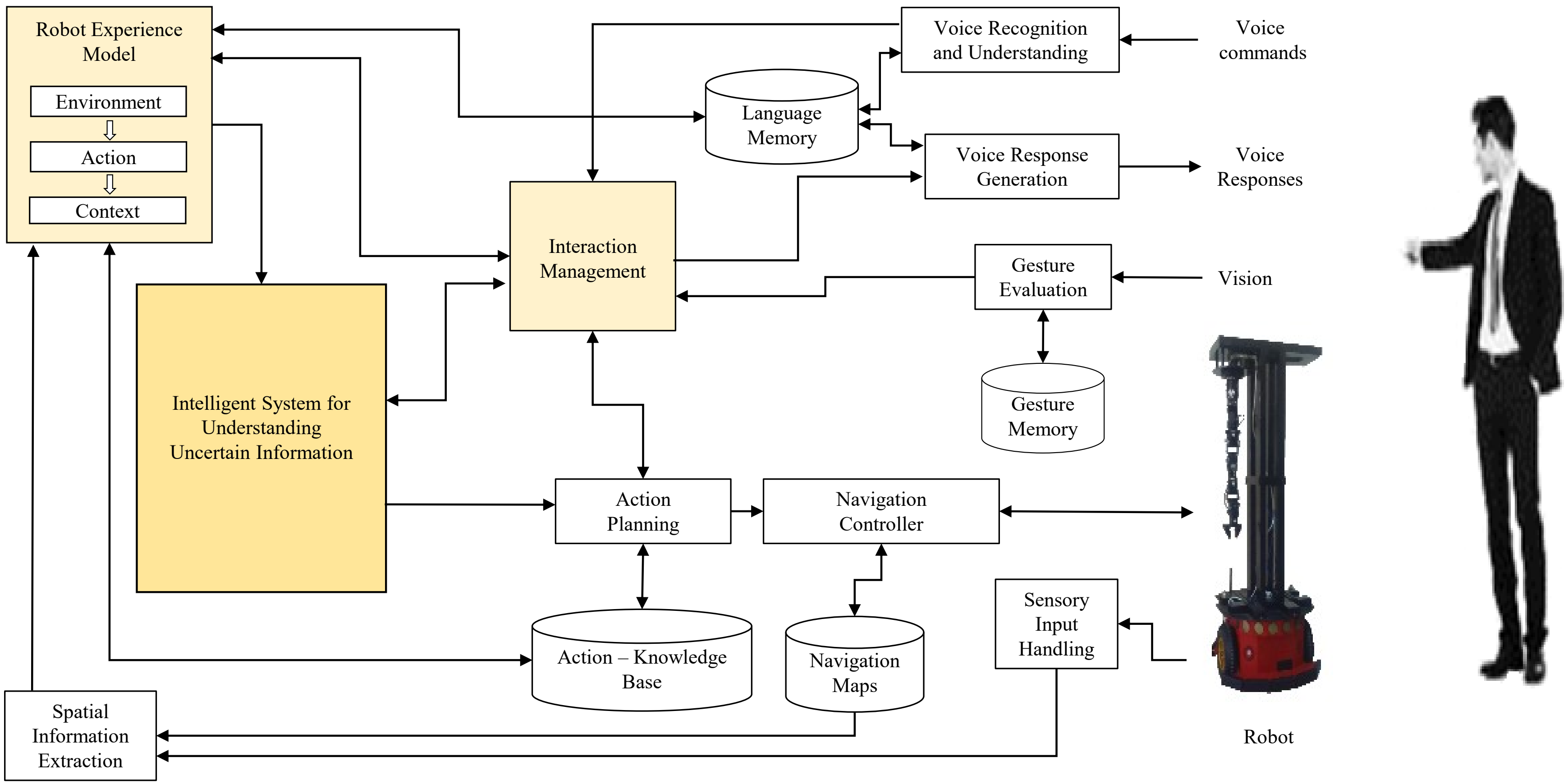


Figure 3.2: Functional overview of the system. The core contribution of this thesis is the development of the Intelligent System for Understanding Uncertain Information (IUUI). The Interaction Management Module (IMM) and the Robot Experience Model (REM) are partial contributions of this work. All other blocks (unshaded) do not possess novel contributions arisen from the work proposed in this thesis and are merely applications of existing methods.

The Gesture Evaluation Module (GEM) is deployed for identifying the non-verbal instructions accompanied with voice instructions by analyzing the skeleton of the user returned by the Kinect motion sensor attached to the robot. The interactions between the robot and the user are managed by the Interaction Management Module (IMM) in accordance with information retrieved from the Robot Experience Model (REM). In order to facilitates these behaviors, the IMM is implemented with a finite state intention module. The required set of actions for a particular interaction is determined by the IMM. Then, this required set of actions is executed by the Action Planning Module with the aid of the Action Knowledge Base and the Navigation Controller. The REM is a layered architecture that organizes the knowledge of the robot about its environment, actions, and context. In addition, the Action Knowledge Base and the Language Memory are managed by the REM.

The uncertain information in navigation instructions is interpreted by the Intelligent System for Understanding Uncertain Information (IUUI). This module consists with various submodules and specific details about those modules will be discussed in the succeeding chapters. The knowledge of the REM is used by the IUUI during the inferencing of uncertain information.

The low-level control functionalities of the robot are handled by the Navigation Controller. It is capable of navigating and path planning from an initial position to a goal position while avoiding obstacles in the environment. The required navigation maps are created using the Mapper3 application [9]. The maximum limits of motion error coefficients of the navigation controller are configured as follows for localization and path planing actions. When the robot moves linearly, the error in the distance is 0.05 mm per 1 mm distance traveled. When the robot rotates, the error in the rotational angle is 0.05 degrees per 1 degrees angle turned. When the robot moves linearly, it can also affect its orientation. This drift is 0.0025 degrees per 1 mm distance travelled. The Sensory Input Handling Module

[9] www.mobilerobots.com/Software/Mapper3.aspx

(SIHM) is used to retrieve information from the robot's built-in sensors, such as range sensors. The Spatial Information Extraction Module (SIEM) perceives spatial information about the environment by extracting information from the navigational maps and from the information retrieved by the SIHM. Then, the perceived spatial information is sent to the REM.

3.2.1 Command and robot action identification

Ability to use flexible user commands enhances the interaction between the robot and the user. Hence, the possible structures of the navigation user commands have been identified in order to develop a method to understand flexible user commands. Mainly navigation user commands can be classified into two main categories, motional commands and positional commands. A motional command is can be used to move the robot some distance from its initial position inside a room and it does not reveal information related to a position. In this kind of a command, the required distance value for the movement needs to be measured from the robot. As an example, "move far forward" can be considered. A positional command can be used to move a robot to a position, which is mentioned with respect to a reference. In this kind of a command, the required distance value needs to be measured from the reference. As an example, "move close to the table in the kitchen" can be considered.

The motion direction of a motional navigation command can be given directly with respect to the robot or with respect to a reference point in the surroundings. For simplicity of the implementation of the command identification process, it is assumed that motional commands can be classified into two types based on the manner in which the direction is given. If the direction is given directly with respect to the robot, the possible directions are assumed to be "left", "right", "forward" and "backward". Commands 1 and 2 in Table 3.1 are examples of such commands. In commands of this type, the distance that must be traveled by the robot is expressed by means of an uncertain term such as "little" or "far".

Table 3.1: Example User Commands and Corresponding Robot Actions

User Command	Command Description	Required Robot Action(s)
1. Move a little forward 2. Go far to the right	Motional	Type I
3. Move a little toward the TV 4. Move a little in the direction of the table	Motional	Type II
5. Go near to the TV 6. Move near to the table in the kitchen	Positional	If same room, Type III; Otherwise, Types IV & III
7. Go to the office	Positional	Type IV
8. Go near to the bed in the office (no bed in the office)	Erroneous or Ambiguous	Type V (voice response)
9. Too little (after action I or II) 10. Too close (after action III)	Feedback	Type VI (learning)

The robot needs to assign a quantitative value to this uncertain term and then move the corresponding quantitative distance in the given direction. Robot action type I is defined for the execution of commands of this kind. For a direction that is given with respect to a reference point, such as the location of an object in the surrounding environment, it is assumed that such commands will contain direction-related keywords such as "toward" and "direction of". Commands 3 and 4 in Table 3.1 are examples of such commands. To satisfy a command of this type, the robot first needs to identify the reference object. The environmental knowledge layer of the REM is used to identify the reference object and its location (see section 3.2.2). Subsequently, the robot needs to assign a quantitative distance to the uncertain term in the command and then move the corresponding distance. Robot action type II is defined for the execution of such tasks.

When executing a positional command, the robot first needs to identify the reference object and its location. Commands 5 and 6 in Table 3.1 are examples of commands of this kind. In this scenario, the robot needs to move to a position that is uncertain because of the uncertainty in interpreting terms such as "near" and "close". Therefore, the robot needs to assign a reasonable quantitative value to the uncertain term and then move to a position at the corresponding distance from the reference point. Robot action type III is defined for executing such tasks. However, there are situations in which the reference object is in another room and the robot needs to move from the current room to that of the reference object. Robot action type IV is defined for room-to-room navigation. Thus, the robot needs to first perform a type IV action to move to the room where the reference object is located and then perform a type III action. Positional commands also encompass room-to-room navigation commands (e.g., command 7), in which case it is assumed that there are no uncertain terms to interpret; the robot simply moves from the current room to the stated room. Robot action type IV is used for this task.

A user command may be erroneous or ambiguous depending on the arrangement of the environment or the situation. In such a case, the robot uses voice responses to ask for further information or notify the user about the situation. Robot action type V is defined for actions in which only voice interactions are involved. User responses such as "too little", "too far" and "too close" are treated as user feedback; if such feedback is received, then the robot performs a type VI action to adapt its perception (see chapter 6). Examples of user commands and the corresponding robot actions for the possible cases are given in Table 3.1.

User commands are identified by analyzing the received voice commands using the keywords, basic grammar components and lexical symbols that are available in the language memory. Subsequently, the required actions for a particular command are identified based on the knowledge of the REM. This approach allows the user to issue commands that are not bounded by a strict grammar model. However, assumptions have been made in implementing the command identification process; this usage of assumptions is considered to be valid since the main contribution of the research is the development of novel methods of interpreting uncertain information in language instructions. The grammar structures used to identify the user commands are given below in JSpeech Grammar Format (JSGF) [86][10]. Furthermore, the system is capable of mapping the synonyms with the initial tokens of the grammar model as explained in section 3.2.3.

<MotionalCommand> = <action> <distance$_{\mathrm{M}}$> <direction$_{\mathrm{M}}$>;
<direction$_{\mathrm{M}}$> = [<helping_word>] <direction$_{\mathrm{K}}$> | <helping_word><Reference>;

<PositionalCommand> = <action> <distance$_{\mathrm{P}}$> [<direction$_{\mathrm{K}}$>]<Reference>;

<actionn> = (go | move);
<distance$_{\mathrm{M}}$> = (far | medium | little);
<distance$_{\mathrm{P}}$> = (near | close);
<helping_word> =(direction | toward);
<direction$_{\mathrm{K}}$> = (forward | backward | left | right | <sub_dir>);
<sub_dir> = (front | back) (left | right);

3.2.2 Robot Experience Model (REM)

The Robot Experience Model (REM) is used to organize the robot's knowledge of its environment, actions and context. It is separated into three layers for knowledge representation, namely, the environment layer, the robot action layer and the context layer. The context layer of the REM is intended for future developments; it is currently inactive.

The knowledge of the robot about its working environment is stored in the environment layer in a hierarchical tree structure, as shown in Fig. 3.3.

[10]Redundant words such as articles that may be included in user commands are filtered out before parsing them. Therefore, the grammar structures are given without considering the possible inclusion of redundant words.

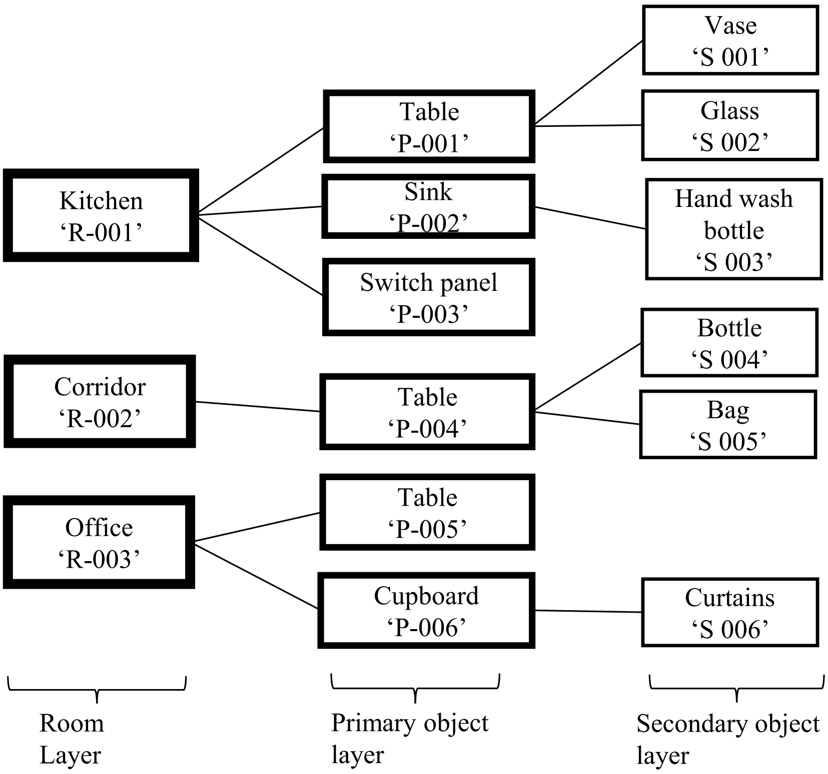


Figure 3.3: Hierarchical tree structure of the environment layer of the REM.

This enables the robot to organize its knowledge about heterogeneous domestic domains in a constructive manner such that it can be utilized for high-level decision-making. Knowledge about the rooms in the domestic environment is represented in the top sublayer. The next sublayer contains knowledge about the primary objects located inside the rooms represented in the top layer. The bottom sublayer contains knowledge about secondary objects that are often located on top of primary objects. The knowledge stored in the environment layer is used to identify the object of interest referenced in a particular user command. The characteristics of the object of interest and the room of interest can be retrieved from this layer to interpret uncertain information. In addition, this enables the IMM module to detect inaccurate user commands that do not comply with the environment and to subsequently generate responses to them. The environment layer of the REM is updated in accordance with navigational maps, sensory inputs and knowledge acquired through interactive discussions as described in section 3.2.3.

The robot action layer represents the robot's knowledge of its actions. The knowledge in this layer is used to identify the required set of actions for satisfying a particular user command based on the information in the environment layer. Five action types have been defined to fulfill the requirements of navigational commands as explained in section 3.2.1. In addition, the knowledge stored in the robot action layer is used to retrieve information on previously performed actions during execution of an interaction with a user.

For this research, the context layer of the REM is inactive and it is proposed for future developments. The language memory and the action knowledge base are also managed by the REM. The REM is updated according to the navigational maps and the sensory information received from the robot.

3.2.3 Interaction Management Module (IMM)

The proposed system is capable of acquiring knowledge through the interactive communication with the user while handling the uncertain information. In order to facilitate this the IMM has been implemented as a finite state intention module [1]. Functional overview of the IMM is shown in Fig. 3.4 as a finite state acceptor diagram.

The default intention state is set as "Waiting". In this state robot is waiting for a user instruction to perform an action. If the received user instructions is compliance with a robot action of type I or II or III or IV. Then the state is changed to "Action planning". In the "Action planning" state, the sequence of the required robot actions is decided. When the action going to be performed is a robot action type IV action then the state is changed to "Perform" state. In the "Perform" state robot perform an action of robot actions type I-IV. When the required action is an action of robot action type I or II or III then the state is changed to "Uncertainties interpretation" state. In this state robot interprets quantitative values for the uncertain terms in the user commands. After the

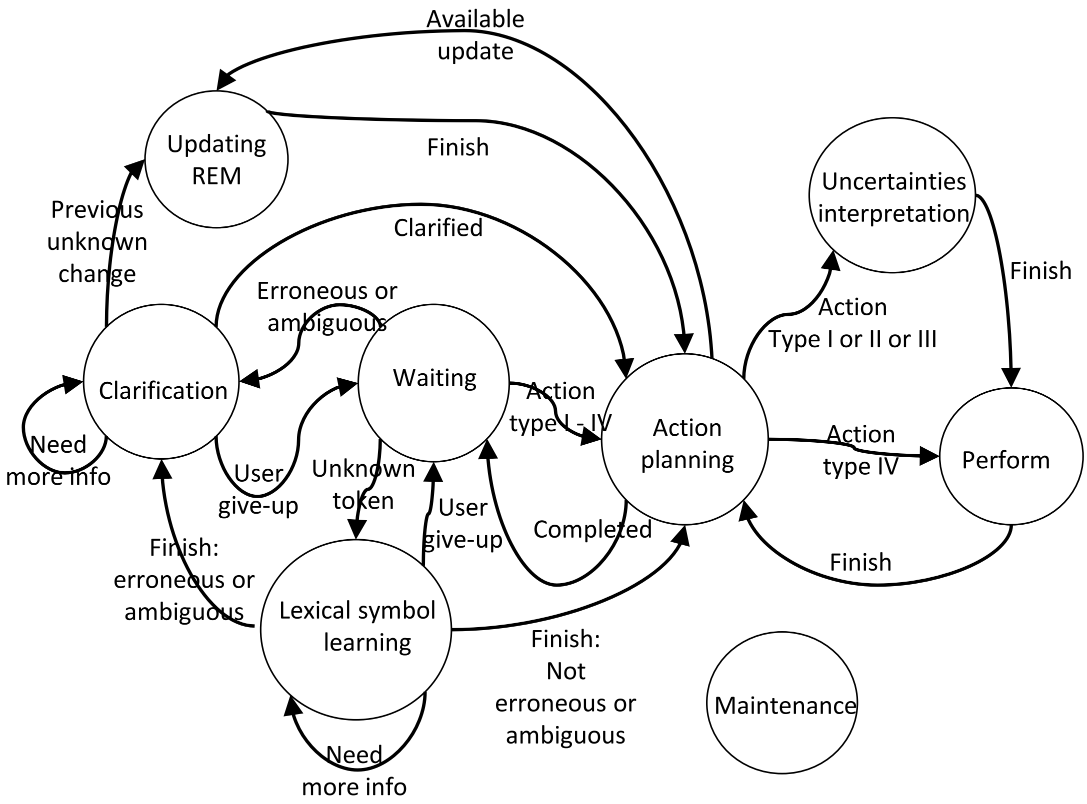


Figure 3.4: Finite state acceptor diagram of the system. It is possible to change the state to "Maintenance" from any other available states if there is a requirement of a maintenance work of the robot such as low battery condition.

interpretation process is finished then the state is changed to "Perform" state. After performing a robot action, the state is changed back to "Action planning". If an update of the REM is available then the state is changed to "Updating REM". After finishing the updating, the state is changed back to the "Action Planning" state. After finishing the required sequence of the robot actions, the state is changed back to "Waiting".

If the user command is erroneous or ambiguous according to the existing knowledge of the robot, the state is changed to "Clarification" state. In this state, the robot seeks the help of the user to clarify the ambiguity of the command or further enquire about the possible yet unknown changes in the environment. If the user intention is not to give clarification, the state is changed back to "Waiting". After successfully clarifying the ambiguity, the state is changed to "Action planing" state. If an unknown change in the environment is identified during the clarification, the state is changed to "Updating REM". After finishing the updating, state is changed to "Action planning" to plan the required sequence of actions.

If there is an unknown token in the user instruction, the state is changed to "Lexical symbol learning". In this state the robot seeks the help of the user to learn the unknown token. If the user give up the learning process then the state is changed back to "Waiting". After the unknown token is learned, the state is changed to "Action planning" if the command is compliance with an action of type I-IV. If the command is erroneous of ambiguous, state is changed to "Clarification" for further clarifications.

A set of predefined dialogue flows and patterns are used in order to acquire the knowledge while maintaining a smooth interaction with the user. The predefined dialogue flows of the "Lexical symbol learning" and "Clarification" states are given in the Fig. 3.5 and Fig. 3.6 respectively. The dialogues of the robot are given as a letter code in the diagrams and the corresponding dialogues for the letter codes are given in the Table 3.2. During the listening process, if no voice is recognized within 300 seconds or if the same listening process consecutively runs more than 5 times, the state is changed to "Waiting" with the voice response 'M' or 'C' respectively.

The state is changed to "Maintenance" if there is a requirement of a maintenance work of the robot such as recharging or hardware failure. It is possible to change the state to "Maintenance" from any other available states. During this transition robot gives speech output 'N'.

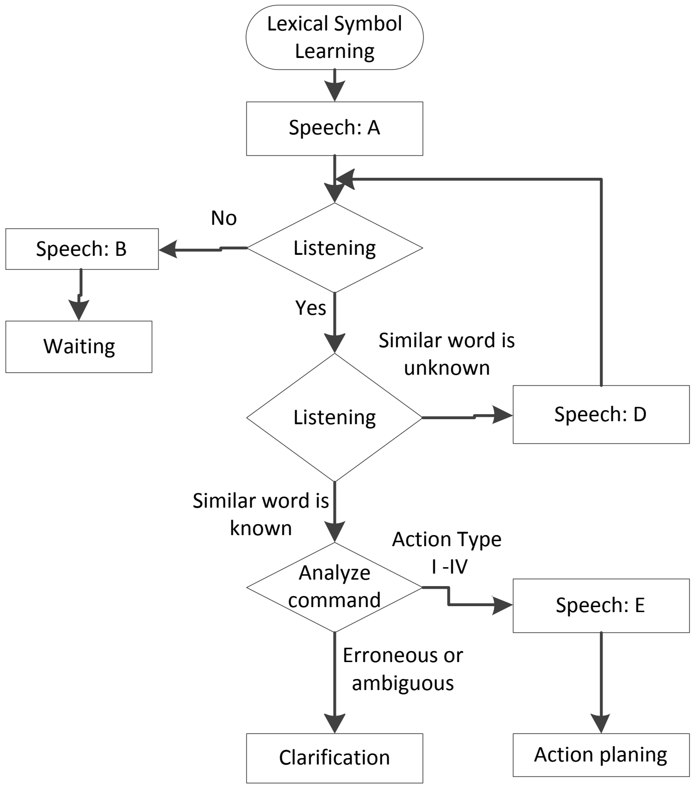


Figure 3.5: The dialogue flow of the "Lexical symbol learning" state. The voice responses of the robot is given as letter code and corresponding voice outputs are given in Table 3.2. During the listening process, if no voice is heard within 300 seconds or if the same listening process consecutively runs more than 5 times, the state is changed to "Waiting" while robot gives the voice response 'M' or 'C' respectively.

3.3 Results and Discussion

Experiments have been conducted in an artificially created domestic environment inside the laboratory facility for verification of the capabilities of the Moratuwa Intelligent Robot (MIRob). As the experiment, a series of user commands has been given to the robot and the responses of the robot have been recorded for the analysis. A selected segment of dialogues between the robot and the user during the experiment is given below.

1. ***User:*** *Move near to the tv*

2. ***Robot:*** *According to my knowledge, there is no tv here. Is there any change?*

3. ***User:*** *Yes*

4. ***Robot:*** *Where is it?*

5. ***User:*** *On the office table*

6. ***Robot:*** *OK. I will do it*

7. ***User:*** *Move near to the sink*

8. ***Robot:*** *I will do it*

9. ***User:*** *Move close to the table*

10. ***Robot:*** *I haven't heard the word close before. Can you give a similar word for close?*

11. ***User:*** *Yes. Near*

12. ***Robot:*** *I will do it*

13. ***User:*** *Move little left*

14. ***Robot:*** *I will do it*

15. ***User:*** *Move close to the tv*

16. ***Robot:*** *I will do it*

Variation of the states of the robot during the execution of this command sequence is illustrated in Fig. 3.7. The positions of the robot during execution of the command sequence is marked on the map shown in Fig. 3.8. Initially the robot was at the location marked as '0' on the map and the state of the robot was "Waiting" as illustrated in Fig. 3.7. The user issued the command "Move near to the tv". Environment layer of the robot at that time is shown in Fig. 3.3.

Table 3.2: Defined dialogues of the robot

Code	Speech output
A	I haven't heard the word $< unknown\, token >$ before. Can you give a similar word for $< unknown\, token >$?
B	What can I do ?
C	I am sorry I couldn't understand your requirement. Do you need any other help?
D	Sorry, I don't know that one also. So can you give me another similar word for $< unknown\, token >$?
E	I will do it.
F	There are more than one $< object >$ in $< location >$. Can you help me to select the correct one?
G	OK
H	Can you give me another clue?
I	According to my knowledge, there is no $< object >$ here. Is there any change ?
J	Where is it?
K	Can you provide more information ?
L	According to my knowledge, there is no $< object >$ in $< location >$. Is there any change?
M	Sorry I couldn't hear anything. Do you need any other help?
N	$< Hardware\, status >$ I require maintenance work.

According to the knowledge of the robot at that time, there was no TV in the home. Therefore, the robot considered that the user command was erroneous and the state was changed to "Clarification". Then the robot responded with dialogue 2 to clarify that from the user. Then the user responded with the dialogue 3 that there had been a change of the environment which the robot had not known yet. Then the robot asked the possible location of the TV by using dialogue 4. Then the user replied with the dialogue 5. Then the robot replied with dialogue 6 and the state was changed to "Updating REM" to update the REM according to the newly acquired knowledge. The updated REM is shown in Fig. 3.9. After finishing the updating, the state was changed to "Action planning". At this stage, the current room and the room of the reference object were different. Therefore, first the robot had to perform an action of robot action type IV for fulfill the requirements, the state was changed to "Perform". After finishing it, the state

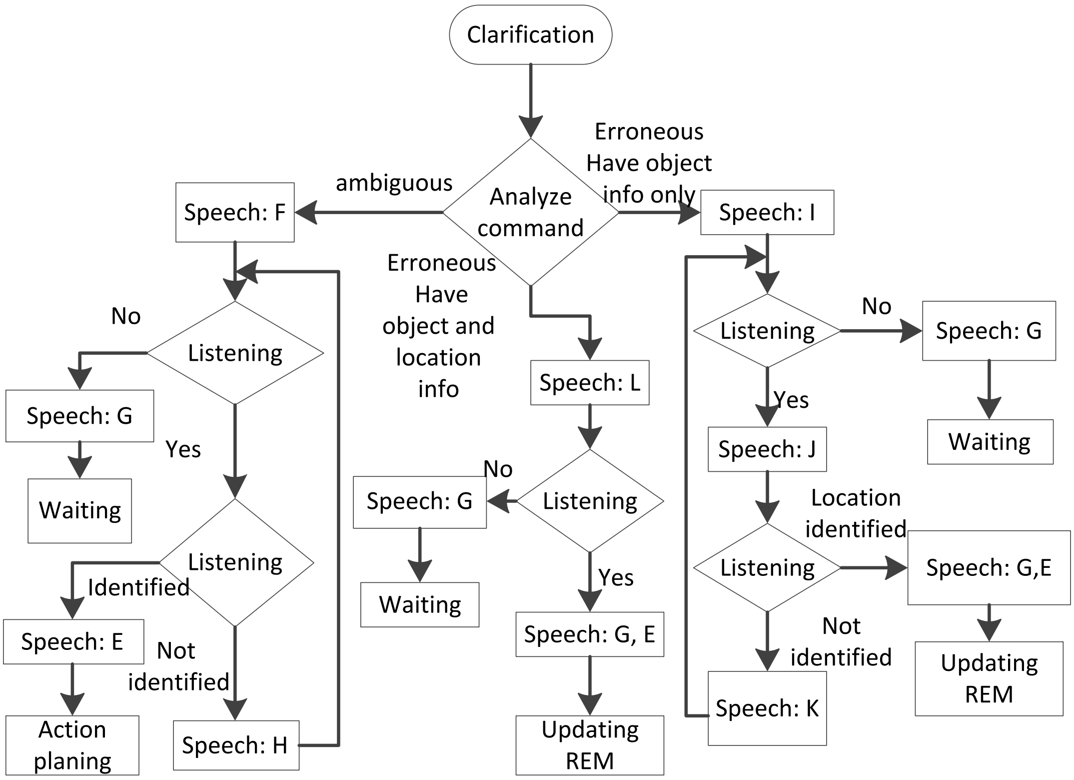


Figure 3.6: The dialogue flow of the "Clarification" state. The voice responses of the robot is given as letter code and corresponding voice outputs are given in Table 3.2. During the listening process, if no voice is heard within 300 seconds or if the same listening process consecutively runs more than 5 times, the state is changed to "Waiting" while robot gives the voice response 'M' or 'C' respectively.

was changed back to "Action planning". Then the robot had to perform type III action and then the state was changed to "Uncertainties interpretation" state to interpret a quantitative value for the uncertain term "near". After finishing the interpretation, the state was changed to "Perform". After performing the action, state was changed back to "Action planning". All the required sequence of actions had been already completed at this stage to fulfill the command issued by the user in dialogue 1 and because of that, the state was changed to "Waiting" and the robot was waiting for a new instruction from the user. At this stage robot settled at the location '1'. Captured snapshots of the robot while executing the command sequence up to this point are shown in Fig. 3.10.

In dialogue 7, the user asked the robot to move near to the sink. Then robot

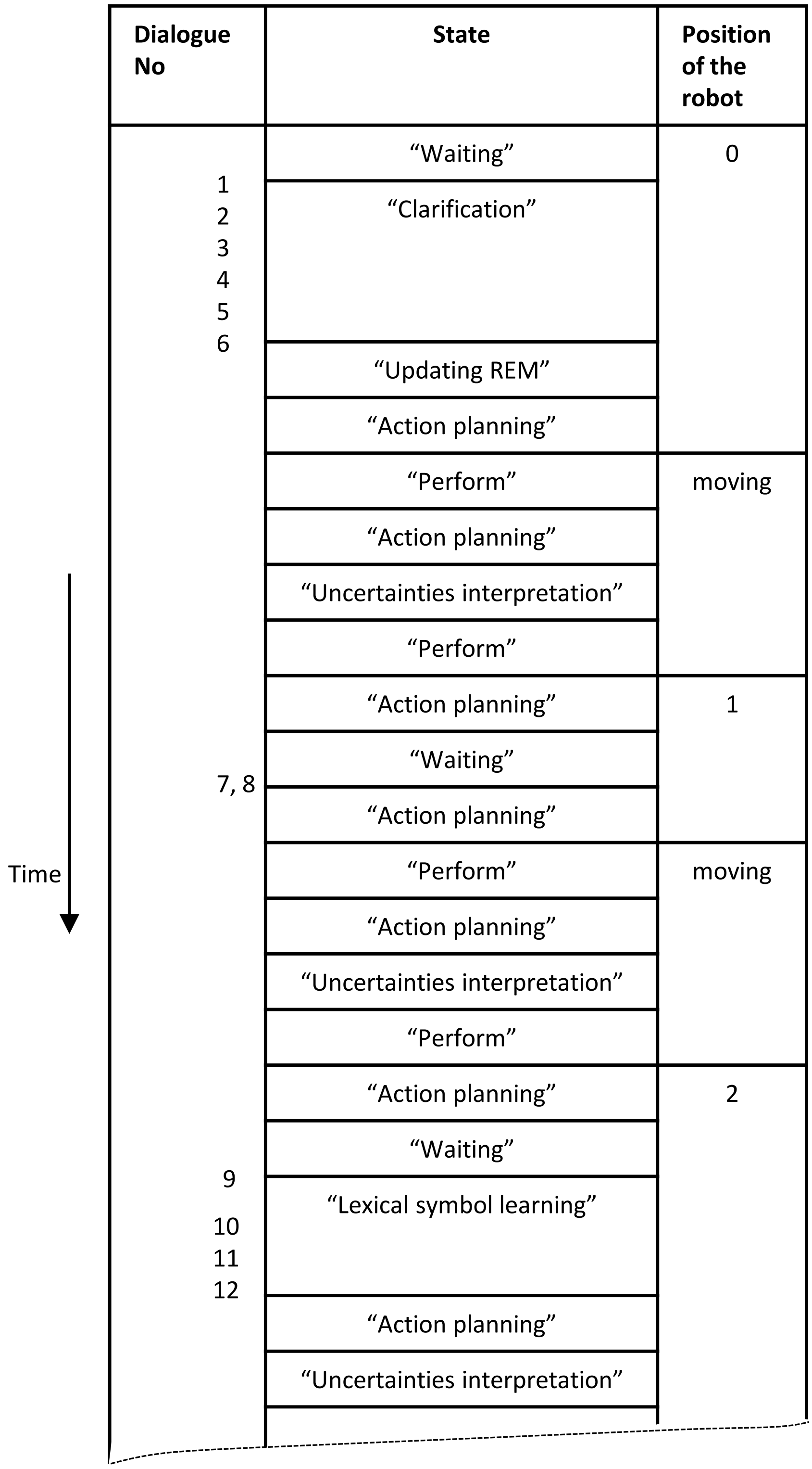


Figure 3.7: Variations of the dialogues, states of the robot and the position of the robot during the execution of the command sequence. It should be noted that the time axis is not drawn to a scale and states which involve only computation such as "Action planing" and "Updating REM" may take less than a fraction of a second.

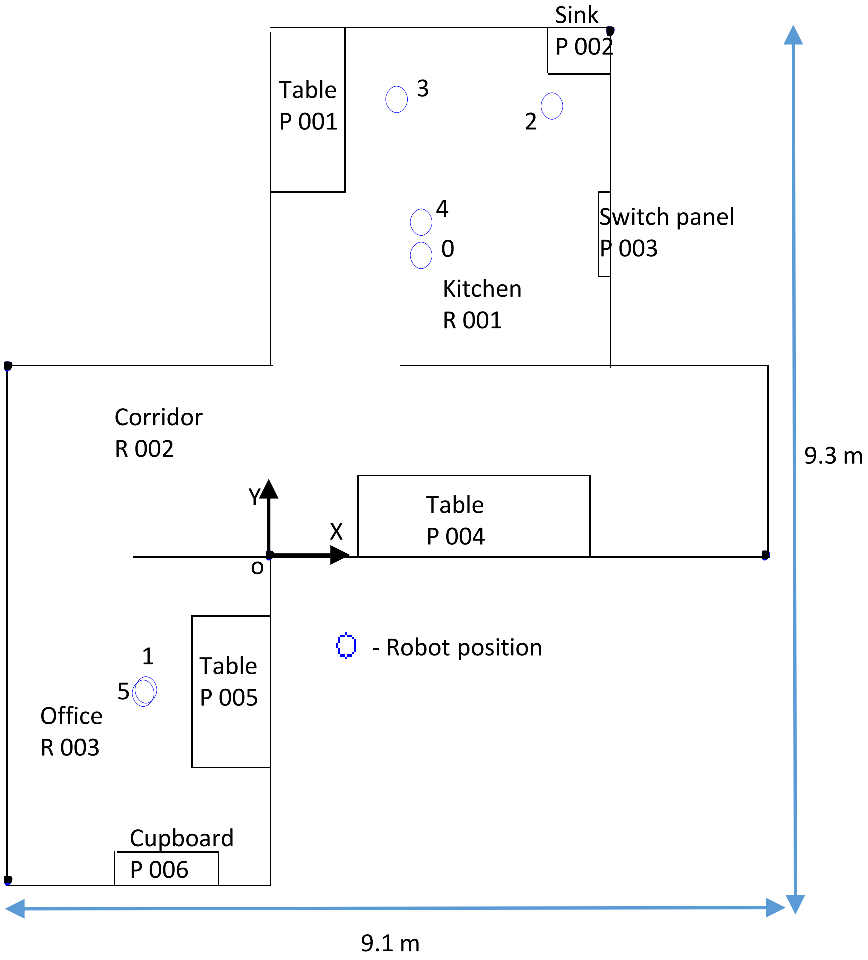


Figure 3.8: Positions of the robot during the execution of the command sequence is marked here. The map is drawn to a scale. However, the markers do not represent the actual size of the robot.

moved to the position '2'. In dialogue 9, the user asked the robot to go close to the sink. At this stage, the robot was not aware of the meaning of the token "Close". In order to learn the meaning of this lexical symbol, the state was changed to "Lexical symbol learning" and the robot asked a similar word for "close" in dialogue 10. Then the user gave a similar word in dialogue 11. Then robot learned that the meaning of the unknown token "close is similar to "near. Then it moved to position '3'. The robot moved from position '3' to '4' to obey the user instruction given in dialogue 13. In dialogue 15, the user asked the robot to move close to the TV set. At this instance, the robot had already known that there is a TV set on the office table as well as the meaning of the word "close"

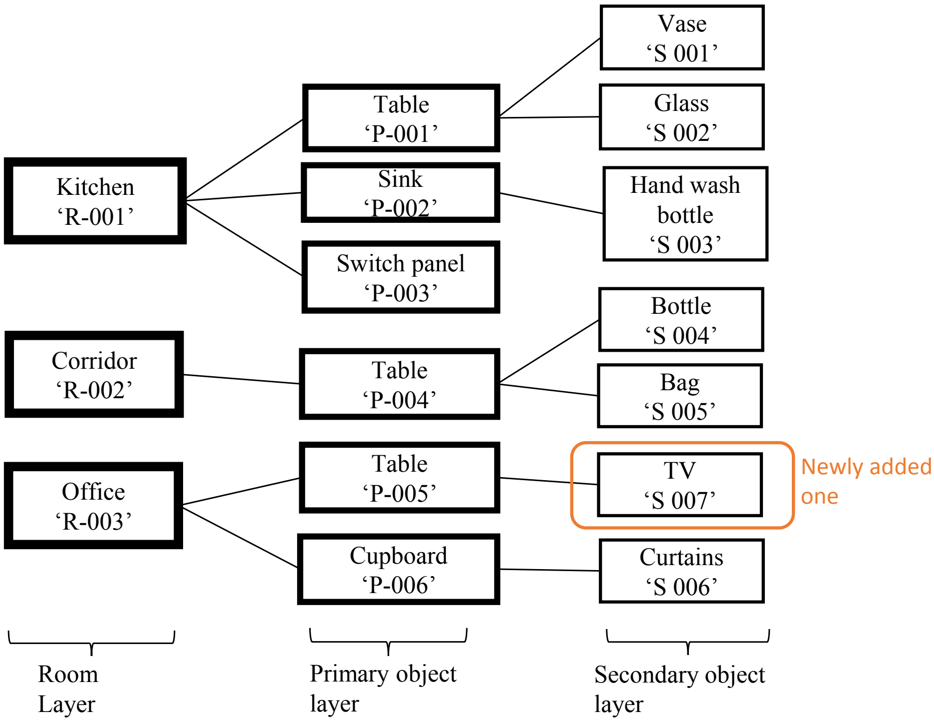


Figure 3.9: Environment layer of the REM after acquiring the knowledge about the TV set. ‘S-007’ has been added to the secondary object layer.

is similar to “near”. Therefore, it moved to the position ‘5’. This validates the ability of the robot in acquiring the knowledge about the environment and learning of unknown lexical symbols. An explanatory video of this segment of the experiment is available in the supplementary multimedia attachment 1[11].

3.4 Summary

This chapter explained the functional and physical overview of an intelligent service robot named Moratuwa intelligent Robot (MIRob) that has also been developed as a part of this research. The auxiliary modules that are required for fulfilling the requirements of navigation instructions have been discussed. Moreover, the work presented in this chapter explains the supportive modules used in

[11]Available in the attached CD and www.youtube.com/watch?v=MtIA-vc09wA

the work presented in succeeding chapters. The developments of the Interaction Management Module (IMM) and the Robot Experience Model (REM) are the main contributions presented in this chapter.

Interaction between the robot and the user is managed by the IMM that has been implemented with a finite state intention module. A set of states have been defined in order to acquire the knowledge from the user using a set of predefined dialogue patterns.

The REM has been introduced to facilitate the command understanding and the uncertain information evaluation. The user commands are not restricted by a strict grammar rule base and the system enables the users to have flexibility in issuing user commands. Hence, the flow of the interaction between the user and the robot has been improved.

Experiments have been carried out in an artificially created domestic environment in order to validate the performance of the system and the obtained results show that the system is capable of learning unknown lexical symbols and acquiring knowledge about the working environment. Furthermore, the concept is capable of updating the REM according to the acquired knowledge and the changes due to the performed actions of the robot. This has been demonstrated and validated from the experimental results.

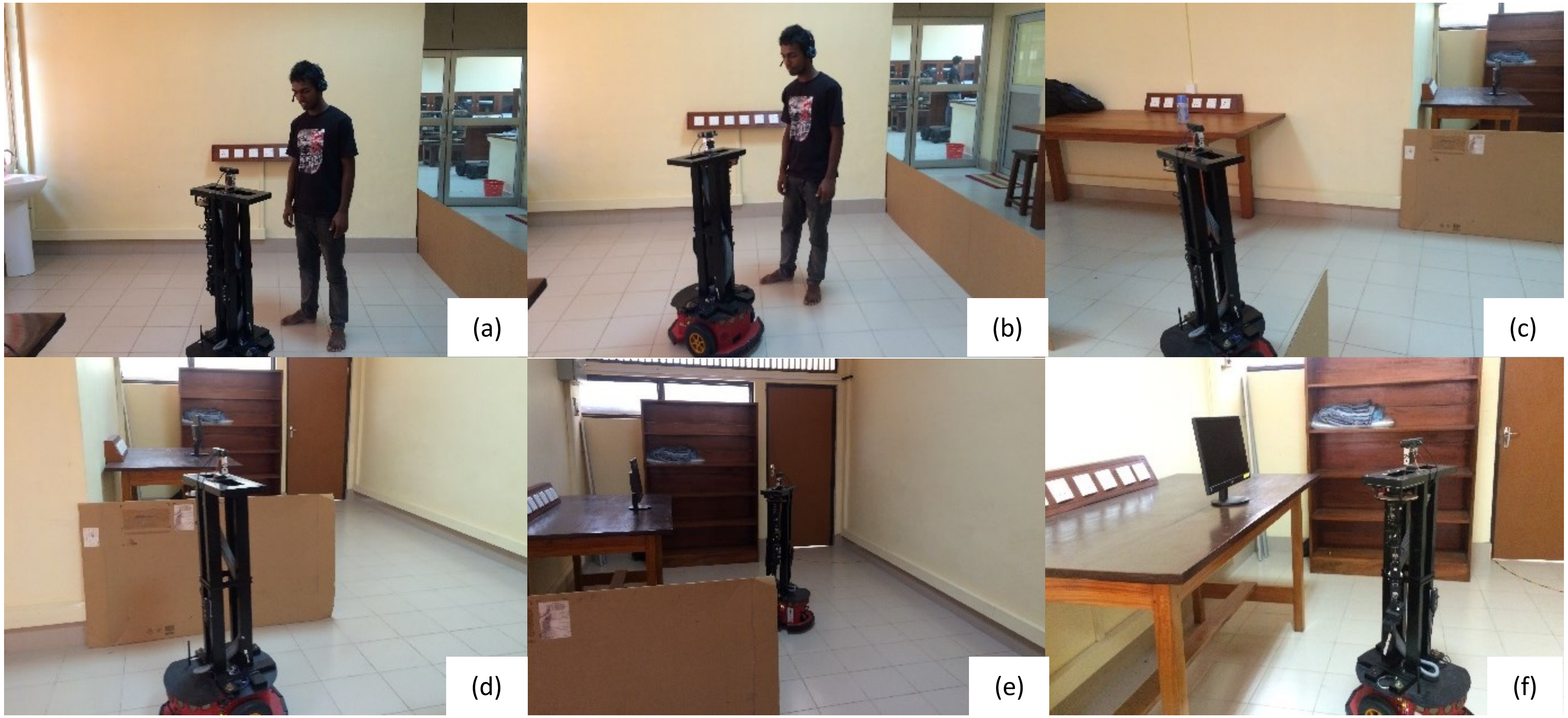


Figure 3.10: Captured snapshots of the robot until the dialogue no 7. (a) has been captured when the user and the robot was having the conversation (dialogues 1-6). In (b), MIRob started to move. (c) shows the passing of the MIRob through the door way of the kitchen. Approaching of the MIRob towards the doorway of the office room is shown in (d). (e) shows the moving of the robot inside the office room. (f) has been captured after settling the robot near the TV set.

Chapter 4

ADAPTING ROBOT'S PERCEPTION OF DISTANCE NOTIONS BASED ON ENVIRONMENT

According to the outcomes of the literature survey presented in chapter 2 , the existing systems use limited number of environmental factors for adapting the perception of uncertain information. For instance, method proposed in [65] uses room size for scaling the meaning of positional information such as "close" and "far". However, depending only on the size of the room is not enough for effective interpretation of uncertain information in navigational commands such as "move little forward" since inside the same room there may be different arrangements of the objects that affect the meaning of uncertain information such as available free space. The movement constraints caused due to 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. 4.1. In both scenarios, the person is instructed to move little forward inside the same room with different initial position for the person. In the scenario shown in Fig. 4.1(a), the person is standing in front of a wall with 50 cm gap between the wall and him. Therefore, the distance meant by the term, "little" may be approximately 15-20 cm. However, the person is standing 150 cm away from the wall in the situation shown in Fig. 4.1(b) and the distance meant by "little" in this situation may be 50-60 cm. Furthermore, the availability of free space is also influential for the mobility of people. In addition to that, the meaning of positional information given with respect to a landmark or an object depends on the saliency of the landmark [40]. For example, the

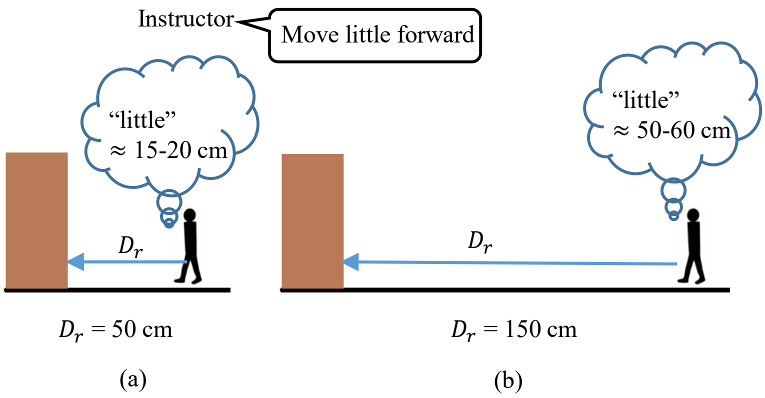


Figure 4.1: This explains the main motivation of the work proposed in this chapter. The person in both the situation is commanded to move little forward. The person is standing 50 cm and 100 cm away from the wall in scenario (a) and (b) respectively.

distance meant by "near" in the command, "go near to the pond" is larger than that of "go near to the flower pot" since the size of the land mark (i.e., size of the pond in this case) is higher than that of the flower pot. Therefore, the size of the object plays a major role in modifying the perception of distance related positional information given with respect to a reference object. Hence, the size of the reference object is used for adapting the perception of positional distance information by the system. Therefore, the Distance Interpreter (DisI) of the Uncertain Information Understanding Module (UIUM) is proposed in this chapter in order to adapt the robot's perception of uncertain information related to distances. The proposed DisI of the UIUM utilizes environmental factors; room size, available free space, size of the reference object and the movement restrictions caused due to obstacles in order to adapt the perception of uncertain information based on the spatial arrangement of the environment.

4.1 The DisI of the UIUM

The DisI of the UIUM is deployed as the Intelligent System for Understanding Uncertain Information (ISUUI) briefed in chapter 3 to interpret quantitative distance values for the uncertain terms in navigational user commands such “little”, “far”, “close” and “near”. The structure of the ISUUI with the proposed DisI of the UIUM is depicted in Fig. 4.2. The DisI of the UIUM consists with two sub module for interpretation of uncertain information in motional commands and positioning commands respectively. The suitable submodule is chosen according to the action type. For action type I and II, sub module 1 is chosen and for type III, submodule 2 is chosen. In here, the action type IV is assumed that it is just a navigation from a room to a predefined point in another room and it does not contain uncertain information for interpretation. The action type is identified based on the action layer of the Robot Experience Model (REM) as explained in section 3.2.1.

4.1.1 Module 1

Module 1 is used to interpret uncertain information in motional commands that is action type I and II actions. Commands “move far forward” and “move little towards the table” can be considered as example commands for action type I and type II respectively. In here, there is an uncertainty in interpreting quantitative

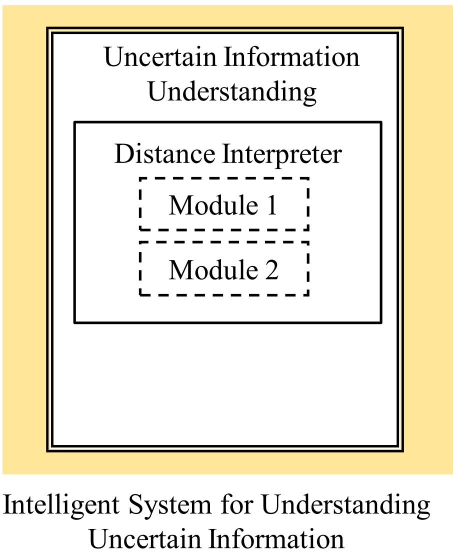


Figure 4.2: Structure of the ISUUI with the proposed DisI of the UIUM.

values for term “far” and “little” but direction of the movement is certain either a fixed direction such as forward or a direction with respect to a reference. Hence, this fuzzy inference system is used to interpret a quantitative distance value for the movement of the robot from its current position to destination position. It has two inputs; uncertain distance descriptor (i.e., $<$ distance$_{\mathrm{M}}$ $>$) of the user command and the free space of the current room. Initially as the $<$ distance$_{\mathrm{M}}$ $>$, “little”,“medium” and “far” have been defined. Other uncertain terms, which may present in user commands need to be mapped in to one of these categories as explained in 3.2.3. The output of the fuzzy inference system is the quantitative distance value estimated to the uncertain term. Centre of area method is used for the defuzzification of the output. The membership functions for the inputs and the output are shown in the Fig. 4.3.

The input membership function for the free space is adjusted based on the room size (S). The room size and the free space are retrieved from the environment layer of the REM. The output membership function is scaled according to the perceptive distance, D such that $D = D_r - d_0$ where d_0 is the safety clearance of the robot and the Dr is the distance to the nearest obstacle in the moving direction. This is illustrated in Fig. 4.4. The rule base of the fuzzy system is shown in Table 4.1.

Table 4.1: Rule Base of the Fuzzy Module 1

Input Memberships		Free space		
		S	M	L
Action modifier	Little	VS	S	M
	Medium	S	M	L
	Far	M	L	VL

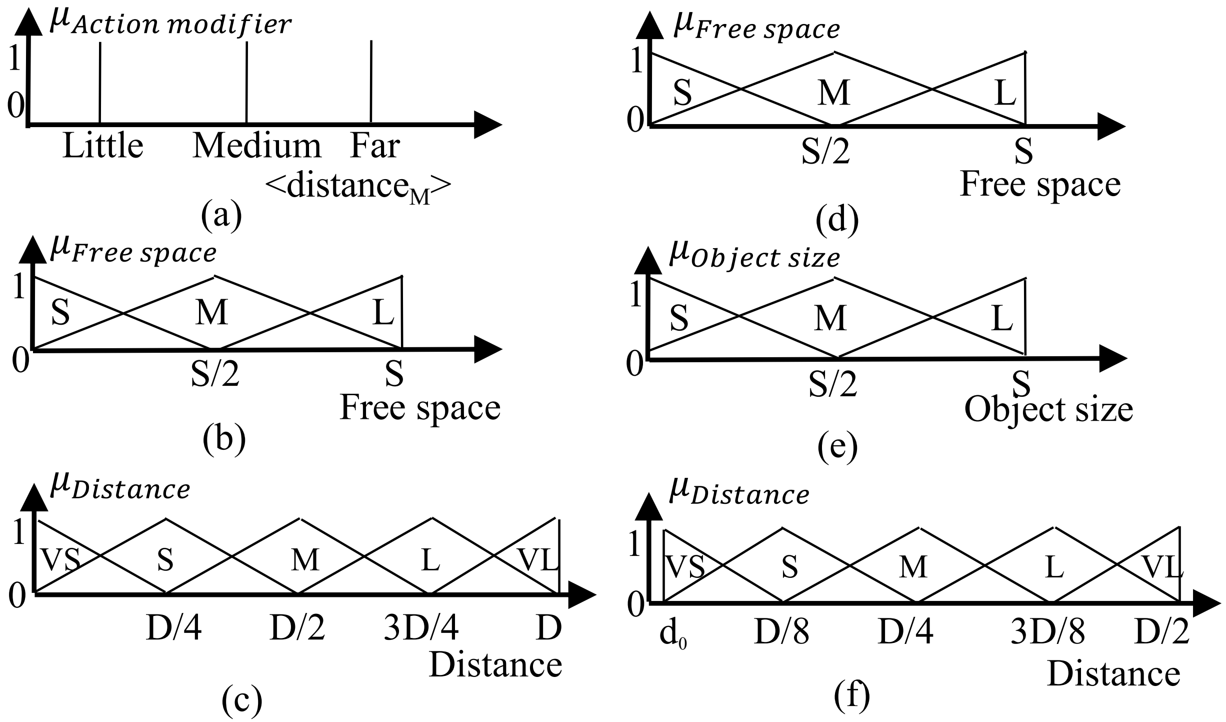


Figure 4.3: (a) and (b) represent the input membership functions of the module 1. (c) represents the output membership function of the module 1. (d) and (e) represent the input membership functions of the module 2. (f) represents the output membership function of the module 2 and only the part which is in the range $[d_0, D/2]$ is considered for the defuzzification. The fuzzy labels are defined as S: small, M: medium, L: large, VS: very small, S: small, M: medium, L: large and VL: very large.

Figure 4.4: D_r and D_{obj} are explained in this figure. The robot is to move in the direction indicated by the white arrow. The distance from the robot to the nearest obstacle or the object of interest is denoted by D_r. The distance between the object of interest and the closest nearby object in the approach direction is denoted by D_{obj}.

4.1.2 Module 2

Module 2 is used to interpret a quantitative distance value for uncertain information in a command when performing action type III. As an example command, "move near to the table in living room" can be considered. In here, the robot should move to a position where it can maintain a quantitative distance value meant by "near" from the reference object. This quantitative value for the term "near" is measured from the reference. The meaning of the uncertain terms such as "near" and "close" are assumed to be same and all the similar kind of uncertain terms are treated equally. The meaning of the positional information given with respect to a reference object depends on the saliency of the objects [40]. Therefore, it is assumed that the quantitative meaning of an uncertain term such as "close" and "near" depends on the available free space of the room and the size of the reference object. Therefore, the fuzzy inference system in the module 2 takes

Table 4.2: The Rule Base of the Fuzzy Module 2

Input Memberships		Free space		
		S	M	L
Object size	S	VS	S	M
	M	S	M	L
	L	M	L	VL

free space and the size of the reference object as the inputs. The output of the fuzzy inference system is the quantitative distance value estimated and it needs to be measured from the reference object not from the robot. The membership functions for the inputs and the output are shown in the Fig. 4.3. The rule base of the system is given in Table 4.2. Centre of area method is used for the defuzzification of the output. The input membership functions for the free space and the size of the reference object are adjusted according to the size of the room (S). The output membership function is adjusted according to either the distance between the robot and the object (D_r) or the distance between the reference object and another object, which is in the approaching direction of the robot (D_obj). This is illustrated in Fig 4.4. The smallest value is selected as the perceptive distance (i.e., D) to adjust the membership function ($D = min(D_r, D_obj)$). In order to maintain a safety clearance between the robot and the objects, the output range is considered as $[d_0,\ D/2]$ in the defuzzification stage. Required inputs for the module 2 are also retrieved from the REM.

4.2 Results and Discussion

4.2.1 Research platform and the experiment

The proposed concept has been implemented on MIRob platform. The experiments have been carried out in an artificially created domestic environment inside the laboratory. The user was asked to issue commands to navigate the robot inside the environment and the responses of the robot have been recorded.

Two sets of experiments have been carried out to analyze the usefulness of the REM for the interpretation of uncertain notions and the behavior of the DisI.

4.2.2 Usefulness of the REM for the interpretation of uncertain information

As the first experiment, a sequence of user commands was given to the robot and the responses of the robot have been analyzed to verify the usefulness of the REM in interpretation process done by the DisI. The positions of the robot during the experiment are marked on the map shown in Fig. 4.5. The issued user commands and the relevant responses of the robot are given in Table 4.3. Initially, the robot was in the position '0' that is inside the corridor ('R 002'). Then the command 1, "Move near to the table in the kitchen" was issued. This is a positioning command. The object of interest and the room of interest were changed to 'P 001' and 'R 001' respectively. Hence, the system was able to identify the correct object and the room. The room of the reference object and the current room were different and therefore robot first performed action IV for going to the required room. When the current room was changed to 'R 001', action III was performed and the robot moved to the position '1'. Then the command 2, "go near to the sink" was issued to the robot. The object of interest was changed to 'P 002" and the room of interest remained unchanged. Then action III was performed by the robot and the robot moved to position '2'. The parameters such as the room of interest, the object of interest, current room and the action type are retrieved based on the knowledge of the REM. The variation of those parameters during the execution of the full command sequence is illustrated in the Fig. 4.6. An explanatory video of this segment of the experiment can be found in the multimedia attachment 2[1].

[1]Available in the attached CD and www.youtube.com/watch?v=bkjxfRI00iI

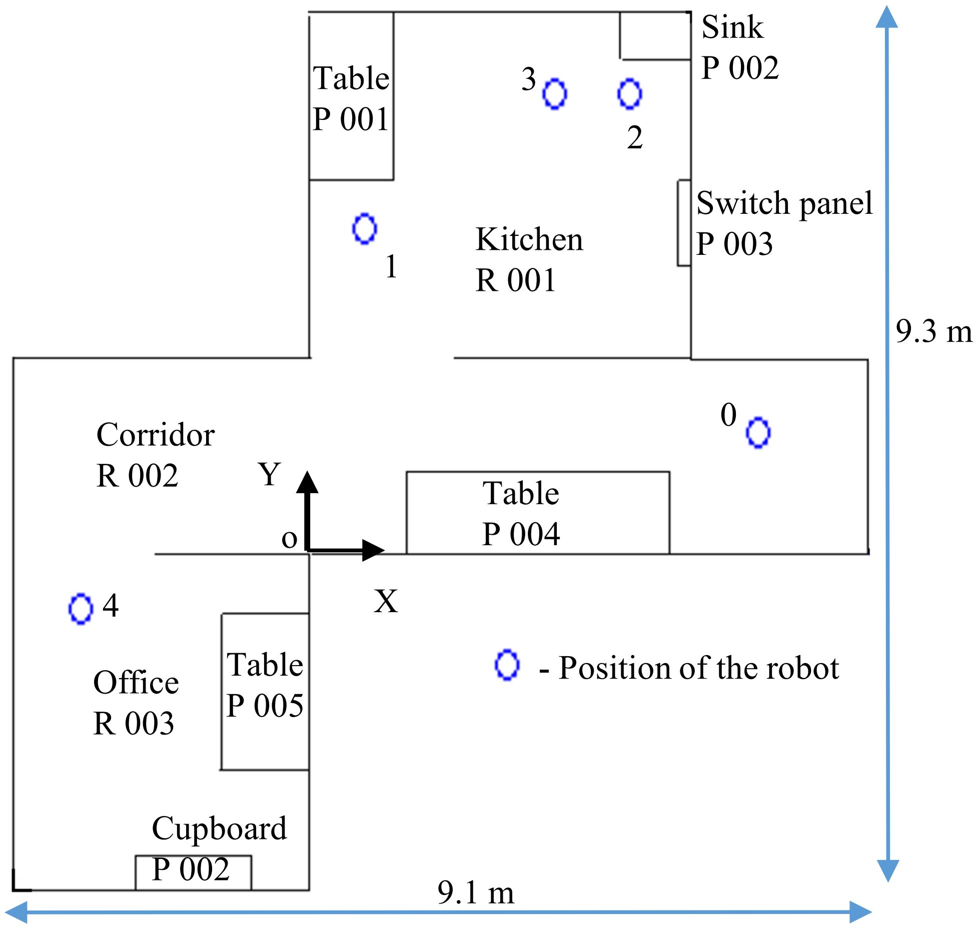


Figure 4.5: Positions of the robot recorded during the experiment I. The map is drawn to a scale and this can be used to visualize the arrangement of the environment.

Table 4.3: Issued user commands and the responses of the robot

User Command	**Response of the robot**
move near to the table in the kitchen	movement from 0 to 1
go near to the sink	movement from 1 to 2
go to the kitchen	voice response: "I am already there"
move little towards the table	movement from 2 to 3
move near to the sink in the corridor	voice response: "There is no sink in the corridor"
move to the office	movement from 3 to 4

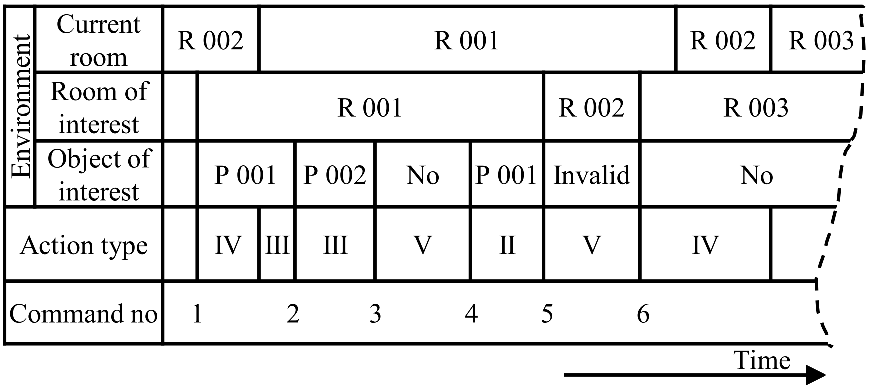


Figure 4.6: Variation of the current room, the room of interest, the object of interest and the action type with the command execution

4.2.3 The behavior of the DisI

In order to analyze the behavior of the DisI, the robot was randomly placed at different positions and the user was asked to issue commands to navigate it. The user was also asked to rate the movement of the robot with respect to his expectation. The possible set of ratings was “too small”, “too large” and “ok” for simple motional commands and “too far”, “too close” and “ok” for direct positioning command. The corresponding initial and the final positions of the robot for a given user command are marked in the map shown in Fig. 4.7 using numbered markers. The issued user commands, initial and final positions of the robot, inputs of the fuzzy systems and the interpreted distance values are given in Table 4.4. The robot positions are given in (X cm, Y cm, heading in degrees) format with respect to the marked origins of the map. It should be noted that according to the active submodule of the UIUM either the <distance$_{\mathrm{M}}$> or the object size is used as an input. Hence, both are shown in the same column. If the command is a direct positioning command, the interpreted distance value is measured from the reference and the robot is moved to the closest position where it can maintain that distance from the object. If the command is a simple motional command, then the output distance value is the distance travelled by the robot.

Table 4.4: Results of the Experiment II

	User command	Initial position (X,Y,Heading)	Uncertain term	Fuzzy controller	<distance$_{\mathrm{M}}$> or object size (m^2)	Free space (m^2)	Room size (m^2)	D (cm)	Interpreted distance D_{out}(cm)	Final position (X,Y,Heading)	User rating
1	go near to the sink	(303,155,0)	near	2	0.379	12.95	15.08	130	38	(332,497,31)	ok
2	go near to the table in the kitchen	(40,48,-86)	near	2	1.62	12.95	15.08	137	43	(52,347,111)	ok
3	go near to the switch panel in the kitchen	(40,48,-86)	near	2	0.138	12.95	15.08	332	79	(309,352,-4)	ok
4	go near to the table	(364,443,174)	near	2	1,62	12.95	15.08	242	67	(166,491,173)	ok
5	go near to the table	(221,496,176)	near	2	1.62	12.95	15.08	130	42	(141,489,177)	ok
6	go near to the office table	(-220,63,156)	near	2	1.562	9.27	11.5	175	52	(-150,-140,-33)	ok
7	go close to the cupboard	(240,138,-8)	close	2	0.652	9.27	11.5	94	36	(-126,-272,-59)	too close
8	move near to the table in the corridor	(-204,-146,-57)	near	2	0.84	16.33	18.85	310	79	(17,50,-40)	too far
9	move near to the table	(562,95,-21)	near	2	0.84	16.33	18.55	183	50	(479,51,-135)	ok
10	move little towards the table	(-262,-241,-123)	little	1	little	9.27	11.5	151	60	(-212,-215,21)	too small
11	move little forward	(-232,-73,-92)	little	1	little	9.27	11.5	255	102	(-232,-169,-89)	ok
12	go little left	(203,465,90)	little	1	little	12.95	15.08	85	36	(175,473,-179)	ok
13	move far forward	(303,290,139)	far	1	far	12.95	15.08	280	228	(143,429,142)	too large
14	move forward	(264,352,-96)	-	1	medium	16.33	18.85	355	291	(256,284,-90)	ok
15	move far left	(58,158,-88)	far	1	far	12.95	15.08	116	78	(344,167,5)	ok

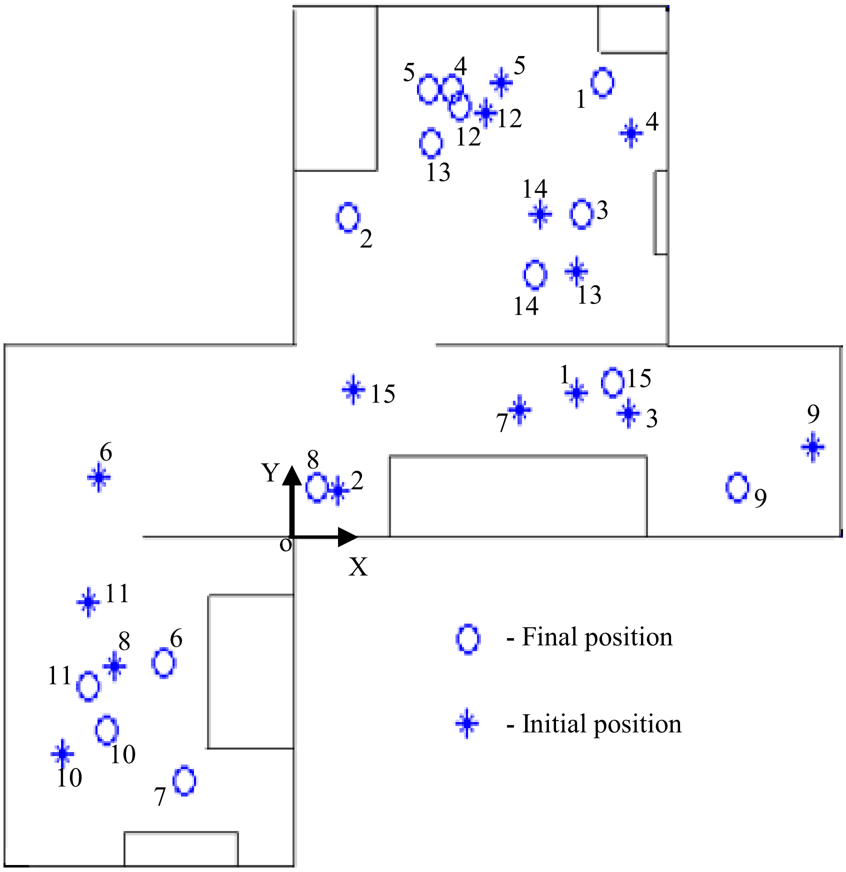


Figure 4.7: Initial and the final position of the robot during the experiment II are marked with correspondence user command number.

The following key features of the system can be pointed out by analyzing the results of the system for positioning commands.

- Interpreted quantitative value of an uncertain term varies with the reference object even though the room is unchanged. Command 1, 2 and 3 can be considered as example cases and interpreted distance values are 38 cm, 43 cm and 79 cm respectively.
- Interpretation depends on the approaching direction even though both the room and the reference are the same. Command 2 and 4 can be considered as example cases.
- Even though the approaching direction, the reference object and the room

are the same, interpretation depends on the robot initial position inside that room. Command 4 and 5 exhibit this phenomena.

- Interpretation adapts according to the environment or the room. For the same kind of reference objects, system outputs are different when the rooms are different. This is clearly showed by command 2 and 6.

For motional commands, following key features can be identified from the obtained results.

- The system adapts according to the environmental changes and the quantitative values for an uncertain term differ according to the environment where the action is being performed. Command 11 and 12 can be used to validate this.
- Interpretation depends on the arrangement of the space such as possible movement restrictions and interpreted distances are different for an uncertain term even though the room is the same. Command 10 and 11 can be considered as example cases.

Therefore, the system is capable of adapting the perception on uncertain terms based on the current environmental conditions and the previously acquired knowledge of the robot. In 11 test cases out of 15, the user has accepted the movement of the robot. However, in some cases the user expectation was different. Other factors, which were not considered in this study, may affect the interpretations. As an example, in case 7, user rated that the movement was too close to the object. In here, the reference object was a cupboard, which has quite high height. Therefore, the user was able to visualize actual size of it. However, the robot is not capable of analyzing the height of the objects to get an idea about the size of it. The limitations of the system such as navigational errors, noises in sonar sensor readings, accuracy of the navigational maps and safety limitations also effect the outcomes of the system.

4.3 Further improving the ability of perceiving the environment for adapting the perception

4.3.1 Rationale behind analyzing the arrangement of the environment in an informative manner

The three scenarios shown in Fig. 4.8 are considered in order to explain the rationale behind the requirement of analyzing the arrangement of the environment in a more informative manner. As an example, in the scenario (a) the robot is commanded, “move little left” and “move little right”. In this scenario, the arrangements of the environment in the moving directions are different. In the left side, there are two chairs and the right side is an open space. If the system analyze only the free space, room size and the distance to the nearest obstacle (distance to the nearest obstacle is indicated as “l” in Fig. 4.8(a)) in the moving direction , then in both cases the interpreted distance value for the term “little” by the system will be the same. However, the responses of the humans are different due to their ability to perceive and analyze the arrangement of the environment in the moving direction. Therefore, the system should be capable analyzing the effects

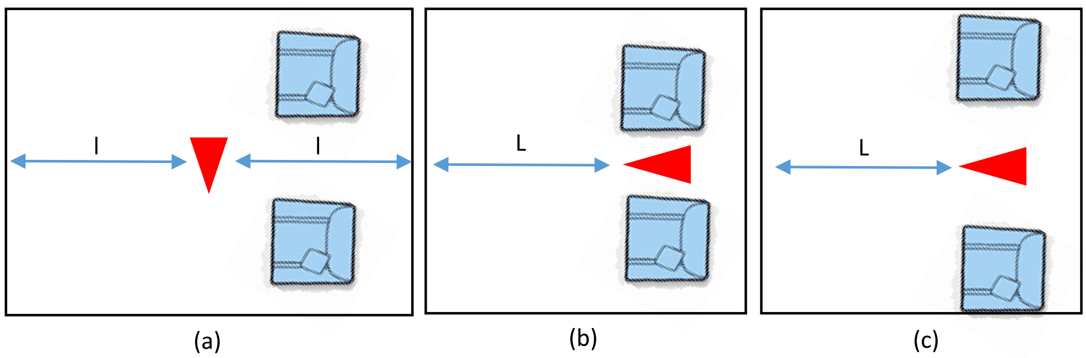


Figure 4.8: Example situations where the robot needs to analyze the arrangement of the environment in an informative manner. In situation (a), the robot is commanded, “move little left” and “move little right”. In situation (b) and (c) the robot is commanded, “move forward”.

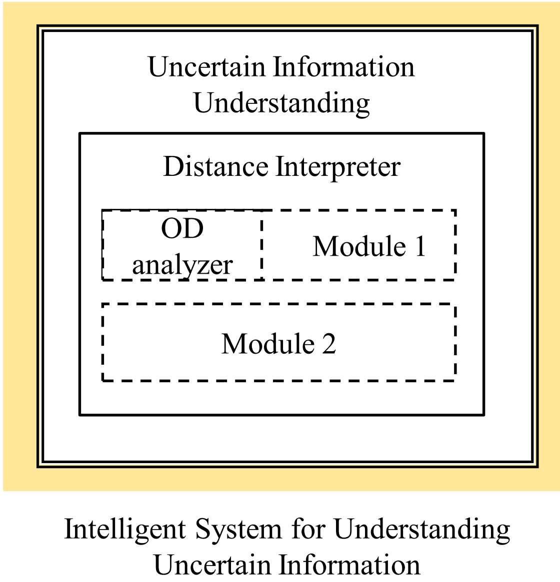


Figure 4.9: Structure of the ISUUI with the introduced OD analyzer

of nearby objects in the moving direction for effective evaluation of uncertain information.

In the scenarios shown in Fig. 4.8 (b) and (c), the robot is commanded, "move forward". In scenario (b), the robot is initially in a position where it is tightly pack with two surrounding chairs compared to the scenario (c). The humans act differently in those two cases because of their ability to adapt the perception based on the spatial arrangement. Therefore, the robot also needs to analyze the spatial arrangement of the initial location of the robot to adapt the perception.

In order to resolve the above-mentioned issues of the system, the Occupied Density (OD) analyzer is embedded to the submodule 1 of the DisI of the UIUM as shown in Fig. 4.9. The OD analyzer is capable of perceiving the arrangement of the environment in a more informative manner and subsequently modifying the robot's perception of fuzzy distance notions. This is realized by modifying the perceptive distance (i.e., D), which adapts the output membership function, by considering the arrangement of the environment and natural tendencies of humans.

4.3.2 Determination of the perceptive distance (D)

The following key features related to the mobility of the humans can be outlined by observing the natural tendencies of the human beings, .

1. Mobility decreases when moving towards an area where the occupied density is high.
2. Effects of close proximity objects are high compared to the faraway objects in the moving path.
3. Effects of the objects in different distance fields depend on the intended moving distance.
4. Mobility increases when moving to a low occupied density area from an initial position located at a high occupied density area.

Therefore, the robotic system should also have the ability to exhibit the above mentioned phenomena in order to enhance ability of evaluating uncertain information in relation to distances. The OD analyzer is used to analyze the occupied density of the surrounding environment and adapt the perception based on the distribution of the occupied density. Occupied density is the ratio between the area of the objects in a particular region and the total area of the region. The occupied density of the environment in a particular scenario is analyzed by dividing the space into regions based on the moving direction and the position of the robot. The surrounding area is divided into zones A, B, C, D and E as shown in Fig. 4.10. A field angle of 90° is considered for the region of the moving direction. Then this region is divided into 3 equal size zones in order to assign different priorities based on the proximity to the moving path. Then each of those three regions are again divided into two by the mid-point of D_r; where D_r is the distance to the nearest obstacle in the moving direction (illustrated in Fig. 4.4. Zones A, B, C and D are used to analyze the occupied density of the moving direction. Zone

E represents the occupied density around the current position of the robot. A function has been defined in order to replicate the natural tendencies of humans based on the distribution of the occupied density in each zone.

According to the fact 2 of the natural human tendencies mentioned in the above, zones A and B have low priority than the zones D and C respectively because the zones D and C are in closer proximity to the moving path than the zones A and B. It is considered that the priority is in 2:1 ratio. Therefore, the combined occupied density of the zones A and D, $OD_{\{A,D\}}$ can be obtained from (4.1) where $OD_{\{A\}}$ and $OD_{\{D\}}$ are the occupied densities of zones A and D respectively. The combined occupied density of zones B and C, $OD_{\{B,C\}}$ can be obtained from (4.2) where $OD_{\{B\}}$ and $OD_{\{C\}}$ are occupied densities of zones B and C respectively.

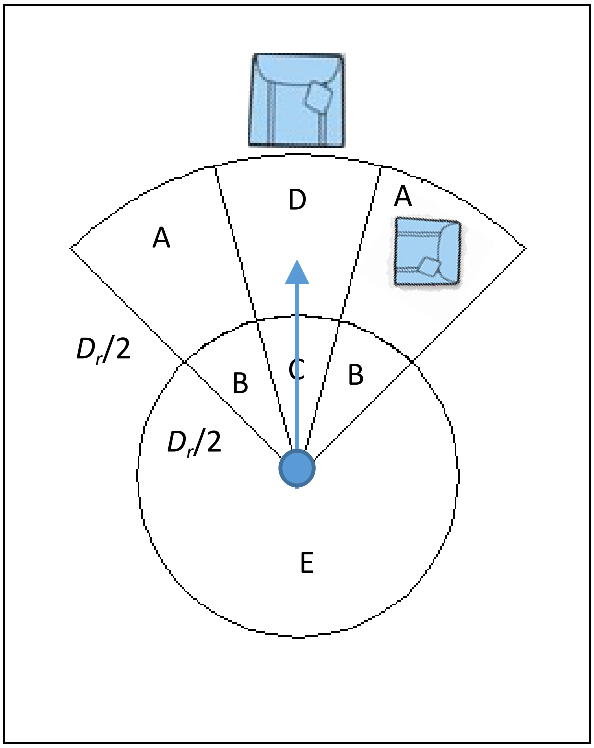


Figure 4.10: Defined zones for analyzing the occupied density is illustrated. The position of the robot is marked by the blue circle and the moving direction is marked by the arrow. D_r is the distance to the nearest obstacle in the moving direction.

$$OD_{\{A,D\}} = 0.66 \times OD_{\{D\}} + 0.33 \times OD_{\{A\}} \tag{4.1}$$

$$OD_{\{B,C\}} = 0.66 \times OD_{\{C\}} + 0.33 \times OD_{\{B\}} \tag{4.2}$$

The combined occupied density of the moving zones (i.e. zone A, B, C and D), $OD_{\{A,B,C,D\}}$ can be obtained from (4.3) where δ_{AD} and δ_{BC} are scalar constant defined based on the priority. Based on the fact 3 mentioned above, the priority between the $OD_{\{A,D\}}$ and $OD_{\{BC\}}$ is defined based on the <distance$_{\mathrm{M}}$> of the command. If the <distance$_{\mathrm{M}}$> is "far" then higher priority is given to the zone A and D over the zone B and C. If the <distance$_{\mathrm{M}}$> is "medium" then the priority is considered as equal. If the <distance$_{\mathrm{M}}$> is "little" then higher priority is given to the zones B and C. Values of the δ_{AD} and δ_{BC} are given in Table 4.5 for each action modifier.

$$OD_{\{A,B,C,D\}} = \delta_{AD} OD_{\{A,D\}} + \delta_{BC} OD_{\{B,C\}} \tag{4.3}$$

According to the fact 1, when the occupied density of the moving zone increases the mobility decreases. Therefore, when $OD_{\{A,B,C,D\}}$ increases, D should be decreased in order to decrease the travelling distance. According to the fact 4, the system should produce higher output when the occupied density of the zone E ($DO_{\{E\}}$) increases. In order to satisfy these two conditions, the perceptive distance (D) is adjusted as given in (4.4) where D_r is the distance to the nearest obstacle in the moving direction (D_r is illustrated in Fig. 4.10) and δ_{ABCD} and δ_E are scalar constant used to tune the system.

Table 4.5: Priority constant variation with the action modifier

	little	medium	far
δ_{AD}	0.33	0.5	0.66
δ_{BC}	0.66	0.5	0.33

$$D = D_r[1 - \delta_{ABCD} OD_{\{A,B,C,D\}} + \delta_E OD_{\{E\}}] \tag{4.4}$$

The obtained D is fed into the submodule 1 of the DisI to adjust the output memberships function. Then, the interpreted distance value (D_{out}) can be obtained from (4.5) where $\acute{D}$ is the defuzzified output of the system.

$$\text{Interpreted distance, } D_{out} = \begin{cases} \acute{D} & \text{if } \acute{D} < Dr \\ Dr & \text{otherwise} \end{cases} \tag{4.5}$$

4.3.3 Results and Discussion

The proposed concept has been implemented on the MIRob platform. The experiments have been carried out in an artificially created environment inside the laboratory facility. MIRob during few of these experimental scenarios are shown in Fig. 4.11. The scalar constants were experimentally chosen as $\delta_{ABCD} = 0.4$ and $\delta_E = 0.3$ by observing the performance of the system.

(a) (b)

Figure 4.11: MIRob being operated through a joystick by a user during the user study is shown in (a). (b) shows a snapshot taken during the 2^{nd} experiment.

Validating behaviors of the OD analyzer

In order to adapt the perception of uncertain terms, the perceptive distance (D) is adjusted based on the results obtained from the OD analyzer. Therefore, experiments have been carried out to validate behaviors of the OD analyzer. An example situation during the experiment is considered for explaining the behaviors of the OD analyzer. The actual arrangement of the environment in that case is shown in Fig. 4.12 (a) for better visualization of the situation. The surrounding environment is divided into zones and then the occupied density of the each zone is estimated by the OD analyzer. In this situation, D_r was 286 cm and the zones are defined based on this value. The divided zones are illustrated with the navigation map of the robot in Fig. 4.12 (b). First, the robot was commanded "move far forward" and the parameters of the OD analyzer have been recorded. Then robot was again placed at the same location and the robot was commanded "move little forward". The variations of the parameters of the OD analyzer in these two situations are given in Table 4.6.

The robot was commanded to move toward a highly occupied area. Therefore, there is a reduction of the perceptive distance (D) in both cases as expected. This validates the ability of the OD analyzer in satisfying the fact 1 of the human natural tendencies explained in the section 4.3.2. Occupied density of the far distance field is high compared to the occupied density in the near distance field ($OD_{\{A,D\}} > OD_{\{B,C\}}$). In the case 1, the robot was commanded "move far forward". Therefore, the intended moving distance is high compared to the case 2 where the command was "move little forward". Therefore, according to the fact 3 of the human tendencies explained in the section 4.3.2, there should be a higher reduction of the perceptive distance (D) in the case 1 compared to the case 2.

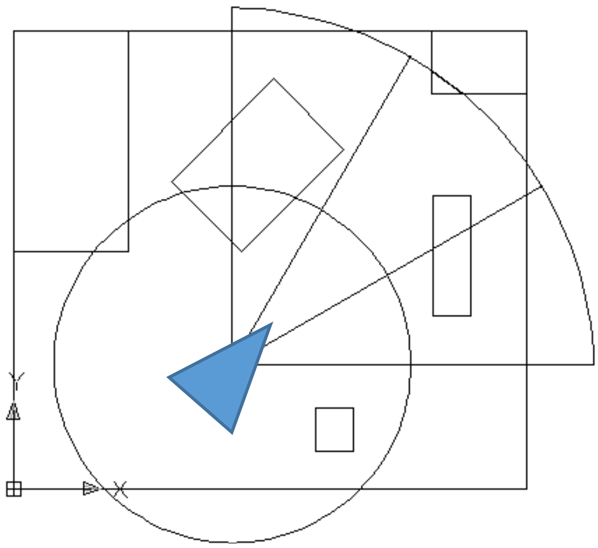
Table 4.6: Variation of the parameters of the OD analyzer

	OD_A	OD_B	OD_C	OD_D	OD_E	$OD_{A,D}$	$OD_{B,C}$	$OD_{A,B,C,D}$	D (cm)
1	0.436	0.137	0.000	0.079	0.168	0.196	0.045	0.144	275
2	0.436	0.137	0.000	0.079	0.168	0.196	0.045	0.095	283

The values of D are 275 and 283 for the case 1 and 2 respectively. Therefore, the results validate the ability of the system to perform the same phenomena.

Even though the occupied density of the zone A is 0.436, i.e. almost half of that zone is occupied by the objects. However, the reduction of the perceptive distance (D) is not very significant. This is because the ability of the OD analyzer in deprioritizing the effects due to the faraway objects from the moving path. This phenomenon is similar to the fact 2 of the human tendencies. There are few objects near the initial position of the robot as a results of that there should be a boosting of D. However, that boosting is attenuated by the effects that are caused due to the objects in other regions. According to the obtained results, the OD analyzer is capable of behaving similar to the natural tendencies of humans

(a)



(b)

Figure 4.12: (a) shows the actual arrangement of the environment for better visualization of the situation. (b) shows the divided zones in the navigation map. The position of the robot is represented by the triangle

explained in section 4.3.2.

performance of uncertain information evaluation capability

In order to evaluate the performance of the proposed system, three sets of experiments have been carried out by rearranging the spatial arrangement of the environment in different conditions. The spatial arrangements of the experimental scenarios are shown in the Fig. 4.13.

The first set of experiments has been carried out without deploying the OD analyzer (i.e in this scenario, $D = D_r$). During this set of experiments the capabilities of the robot is similar to the system explained in earlier section. The second set of experiments has been carried out with the system proposed in this section (i.e., with the OD analyzer). During these two sets of experiments, the robot was placed in the scenarios shown in Fig. 4.13 and the robot was commanded by voice instructions. The responses of the robot along with other key parameters have been recorded. The user command, initial and final position of the robot and other parameters related to the system during execution of user commands are given in Table 4.7. Occupied density variation of each zone during the second set of experiments is given in Table 4.8. The initial and final positions of the robot during the execution of each command are also marked on the maps shown in Fig. 4.13 with the corresponding number.

A user study has also been carried out in order to compare the behaviors of the system with the natural human tendencies. During the user study the robot was placed in the same locations of the environment similar to the first and the second set of experiments. Then the human participant were given a joystick and they were asked to move the robot according to the command issued to them. The same set of user commands has been issued and the movements of the robot have been recorded for the analysis purposes. The quantitative values interpreted by the users have variations. Therefore, only the minimum and the maximum interpreted values are give in the Table 4.7. These distance ranges identified from the user study are also marked on the maps shown in Fig. 4.13 for comparison and the better clarity.

The arrangement of the environment during the case 1 and 2 is shown in Fig. 4.13(a). The initial position of the robot was the same and the robot was commanded to move to opposite directions. In the case 1, the robot was commanded to move towards the area occupied with obstacles. In the case 2, the robot was commanded to move into an open area. According to the natural tendencies of humans, in these two cases, the interpreted distance values should be different and this can be verified from the results of the user study (71-77 cm for case 1 and 81-97 cm for case 2). However, the system without the OD analyzer has interpreted almost the same quantitative distance (79 cm for case 1 and 81 cm for case 2) for the two commands because of the inability of the system to analyze the spatial arrangement. The system proposed in this section (system with the OD analyzer) has interpreted 73 and 83 cm for case 1 and 2 respectively because the system is capable of analyzing the spatial arrangement in a more informative way. Therefore, the system is capable of adapting the perception based on the spatial arrangement of the moving zone. A explanatory video of this segment of the experiment (i.e. case 1 and 2) can be found in the supplementary multimedia attachment 3[2].

[2]Available in the attached CD and www.youtube.com/watch?v=-c07yec-VIw

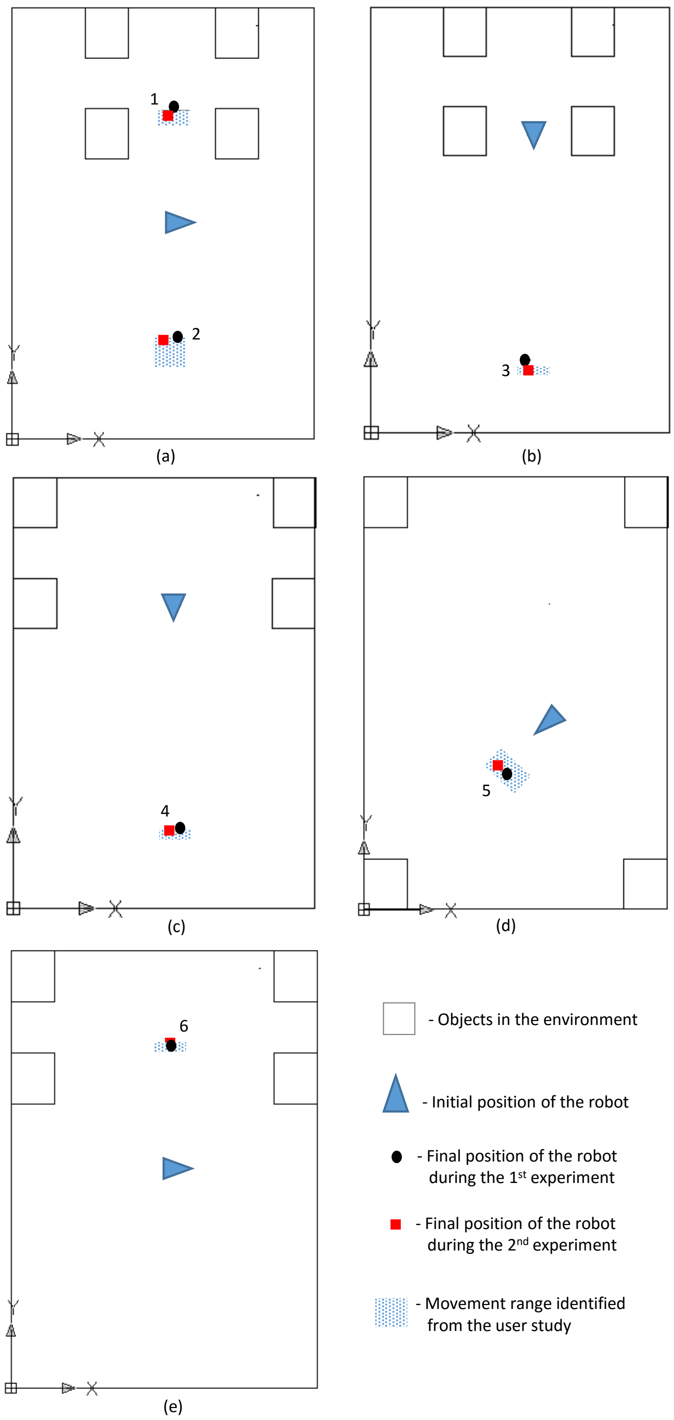


Figure 4.13: Arrangement of the environments during the experimental scenarios are shown. The initial position and the final position of the robot are also marked with corresponding case number. The gray color shaded area represents the movement range obtained from the user study. These maps are drawn to a scale. However, it should be noted that the markers are not drawn to the scale and the size of the markers do not reflects the actual size of the robot. The diagrams may not visualize the actual distance variations of the final positions due to the scaling of the map.

Table 4.7: Results of the experiments

User command	Initial position (X,Y,Heading)	Free space (m^2)	Room size (m^2)	Without OD analyzer				With OD analyzer				Distance moved by user	
				Dr (cm)	**D (cm)**	**Output (cm)**	**Final position (X,Y,Heading)**	**Dr (cm)**	**D (cm)**	**Output (cm)**	**Final position (X,Y,Heading)**	**min**	**max**
1 move left	(112,151,-2)	5.88	6.3	119	119	79	(112,230,94)	118	113	73	(109,226,92)	71	77
2 move right	(112,151,-2)	5.88	6.3	122	122	81	(115,70,-92)	121	129	83	(104,69,-94)	81	97
3 move far forward	(113,212,-89)	5.88	6.3	190	190	162	(108,50,-92)	187	195	167	(110,51,-91)	166	170
4 move far forward	(113,212,-89)	5.88	6.3	186	186	159	(117,53,-88)	185	189	162	(111,51,-91)	158	164
5 move little forward	(129,130,-143)	5.88	6.3	103	103	47	(100,93,-134)	101	99	45	(92,99,-142)	38	52
6 move left	(112,151,-2)	5.88	6.3	118	118	83	(110,234,91)	114	203	83	(109,236,92)	81	84

Table 4.8: Occupied density variation during the experiment 2

Case No	Zone A	B	C	D	E
1	0.3148	0.2155	0.0000	0.0494	0.0083
2	0.0062	0.0000	0.0000	0.0000	0.2155
3	0.2126	0.0000	0.0000	0.0000	0.1242
4	0.2126	0.0000	0.0000	0.0000	0.1242
5	0.3205	0.0000	0.0000	0.0754	0.0000
6	0.0466	0.0000	0.0000	0.0000	0.0000

In the case 3, the robot was placed in a tightly occupied position as shown in Fig. 4.13 (b) and it was commanded to move far forward. In the case 4, the robot was placed in the same position similar to the case 3. However, the arrangement of the surrounding environment was different. In this case, the location of the robot was not tightly occupied compared to the case 3. The only difference in these two scenarios was the arrangement of the environment around the initial position of the robot. Typically, humans prefer to move towards open spaces from tightly occupied areas. Therefore, the robot should move higher distance in the case 3 compared to the case 4. However, the capabilities of the system without the OD analyzer were not enough for correctly identifying the differences in the arrangement of the environment. Therefore, the interpreted distance value of the case 3 is not within the range identified from the user study. In the experiment 2 (i.e. with the OD analyzer), the interpreted distances were 167 cm and 162 cm for case 3 and 4 respectively. These values are within the range identified from the user study. Therefore, these results verify the capability of the system to adapt the perception based on the arrangement of the environment around the initial position of the robot.

In the case 5, the robot was moved towards a corner. The system with the OD analyzer is capable of analyzing the effects occurred due to the narrowness in the corner. Therefore, the moved distance in the experiment 2 (i.e. with the OD analyzer) is less than the moved distance in the experiment 1 (i.e. without the OD analyzer). In the case 6, the robot was moved from an open area to an area where there were few far away objects. However, the system neglected the effects of those objects when interpreting the uncertain information because those objects were not in close proximity to the moving path. Therefore, the interpreted distance values during the experiment 1 and 2 are the same. This validates the capability of the OD analyzer in properly recognizing the priority regions.

4.4 Summary

Methods of adapting the robot's perception of fuzzy distance information in navigation commands based on the environment are presented in this chapter. The proposed DisI is capable of adapting the perception of uncertain distance notions based on the current environmental condition, the knowledge of the robot, and the action that is being performed. According to the experimental results, the system is capable of adapting the perception with the room, reference object, approaching direction and the initial position of the robot.

The effectiveness of interpretation of motional commands has been further improved by deploying the OD analyzer that is capable of perceiving the environment in a more informative manner for adapting the robot's perception of uncertain distance notions based on the environment in such a way that the perception adaptation replicates the natural human tendencies to a greater extent.

The proposed DisI enables the user to navigate a mobile service robot in a human friendly manner and the system is capable of interpreting uncertain information in the user commands more effectively compared to the existing approaches since the perception is adapted according to the environment.

Chapter 5

ADAPTING ROBOT'S DIRECTIONAL PERCEPTION BASED ON ENVIRONMENT

5.1 Fuzziness of Directional Information

In the work presented in chapter 4 assumes that the meaning of direction-related lexical symbols are fixed entities. The example situations illustrated in Fig. 5.1 are used for the explanation of the interpretation of directions in a fixed manner. Fig. 5.1(a) shows an instance where the directions are defined with respect to the robot and this definition is used when the robot is commanded with a motional command. In this situation, that system assumes that the direction meant by "forward" is exactly similar to the current heading angle, θ. The direction meant by "left", "right" and "backward" are fixed as $\theta + 90°$, $\theta - 90°$ and $\theta + 180°$ respectively. Fig. 5.1(b) shows an instance where fixed directions are defined with respect to a reference object and this definition is used when the robot is commanded with a positional command such as "move near to the left of Obj A". However, typically the directional linguistic terms are not fixed entities like this and they are fuzzy in nature [65,87]. Therefore, the Direction Interpreter (DirI) is proposed in this chapter for embedding into the Uncertain Information Understanding Module (UIUM) to evaluate the directional notions in navigation commands considering the fuzziness associated with them. The structure of the Intelligent System for Understanding Uncertain Information (ISUUI) with the proposed DirI of the UIUM is depicted in Fig. 5.2.

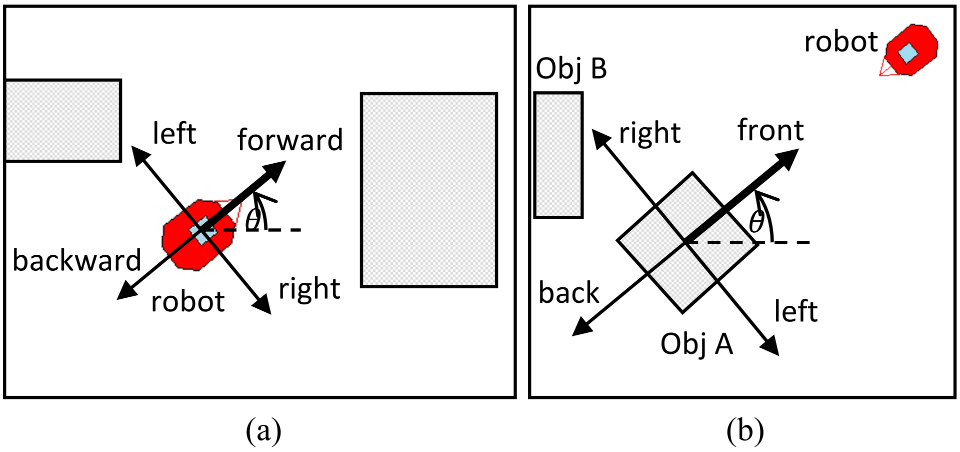


Figure 5.1: This illustrates how the directional notions are defined in the work proposed in chapter 4. The shaded color areas represent the objects in the environment. (a) represents a situation where the directions are defined with respect to the robot. In here, θ is the heading angle of the robot. (b) represents a situation where the directions are defined with respect to a reference object. In here, the direction of the front with the X-axis is annotated as θ. The orientation frame is considered based on the point of view of the robot for this kind of instance. More details related to the assignment of the absolute front of the objects can be found in chapter 7

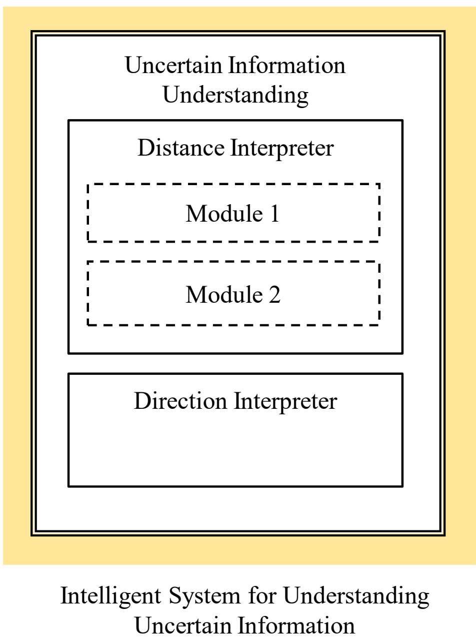


Figure 5.2: The structure of the ISUUI with the proposed DirI of the UIUM

5.2 Adapting the perception of directional notions

According to the work proposed in [87], the directional spatial descriptors are fuzzy and it proposed a method to evaluate the spatial descriptors in static situations. However, the method cannot be used in order to plan a moving direction by interpreting directional information other than evaluating natural language spatial descriptors for identifying a reference. According to [65], the linguistic directions can be modelled with fuzzy membership functions with overlapping boundaries as similar to the model shown in Fig. 5.3(b). However, the output of the system would be pre-determined if it were used for interpreting a quantified heading angle since the membership functions are predefined fixed entities. Moreover, the interpretation is not adapted according to the arrangement of the environment even though the characteristics of movements heavily depend on the arrangement of the environment (see section 4.3.2. Therefore, the proposed Direction Interpreter considers the natural tendencies related to mobility of humans in order to interpret the directional information based on the arrangement of the environment.

The directional linguistic term in a command (i.e., <direction$_{\mathrm{K}}$>) is the input of the system and the input membership function has singleton sets in order to represent the linguistic directional terms as shown in Fig. 5.3(a). The meaning of directional terms "left" and "right" has to be interchanged for motional and positional commands. Therefore, the indexes of the input sets are separately defined for motional and positional commands. The input sets and output sets are directly mapped to each other yielding to a single input single output fuzzy system. The default output membership function has been designed as shown in Fig. 5.3(b) based on the work presented in [65]. It has 8 triangular shape sets for basic Direction Sets (DS) and the sets have overlapping boundaries. If the user command is a motional command, the output is the required change of the heading angle of the robot. If the command is a positional command, the output is the angle to the destination position measured from the absolute front of the reference object. According to the natural tendencies given in section 4.3.2 related to the mobility of humans, humans prefer to move towards an area where the free space is high. Moreover, the robot should decide the direction based on the availability of free space around it instead of a predetermined direction. However, this sort of behavior cannot be achieved directly by using a fixed output membership function that is not adapted according to the surrounding environment.

Therefore, the output membership functions should be modified in such a way that the robot tends to move towards the free area. This can be achieved from weighting the base membership function with the available free space around the robot (for a motional command) or the reference object (for a positional command).

The weighted activation degree of the i^{th} Direction Set (DS_i) for angle α is defined as $\omega_{DS_i}(\alpha)$ and it can be obtained from (5.1), where $\mu_{DS_i}(\alpha)$ is the activation degree of DS_i for the angle, α. d_α is the distance to the nearest obstacle in the direction that creates α angle with the current heading of the robot (for motional command) or the absolute front of the reference object (for positional command). The ways to obtained d_α is explained in Fig. 5.4(a) and Fig. 5.4(b) for motional and positional commands respectively. $(DS_i)_L$ and $(DS_i)_U$ are the lower and upper bound of DS_i and are defined as in (5.2), where β is a scalar constant. The center of the i^{th} Directional Set, $(DS_i)_C$ is defined as in (5.3). The scalar constant, β is taken as $45°$ in order to have the default directional perception of the system similar to the directional perception of the system explained in [65].

$$\begin{aligned} \omega_{DS_i}(\alpha) = & \mu_{DS_i}(\alpha) \cdot d_\alpha \\ & \forall \alpha \epsilon [(DS_i)_L, (DS_i)_U] \end{aligned} \tag{5.1}$$

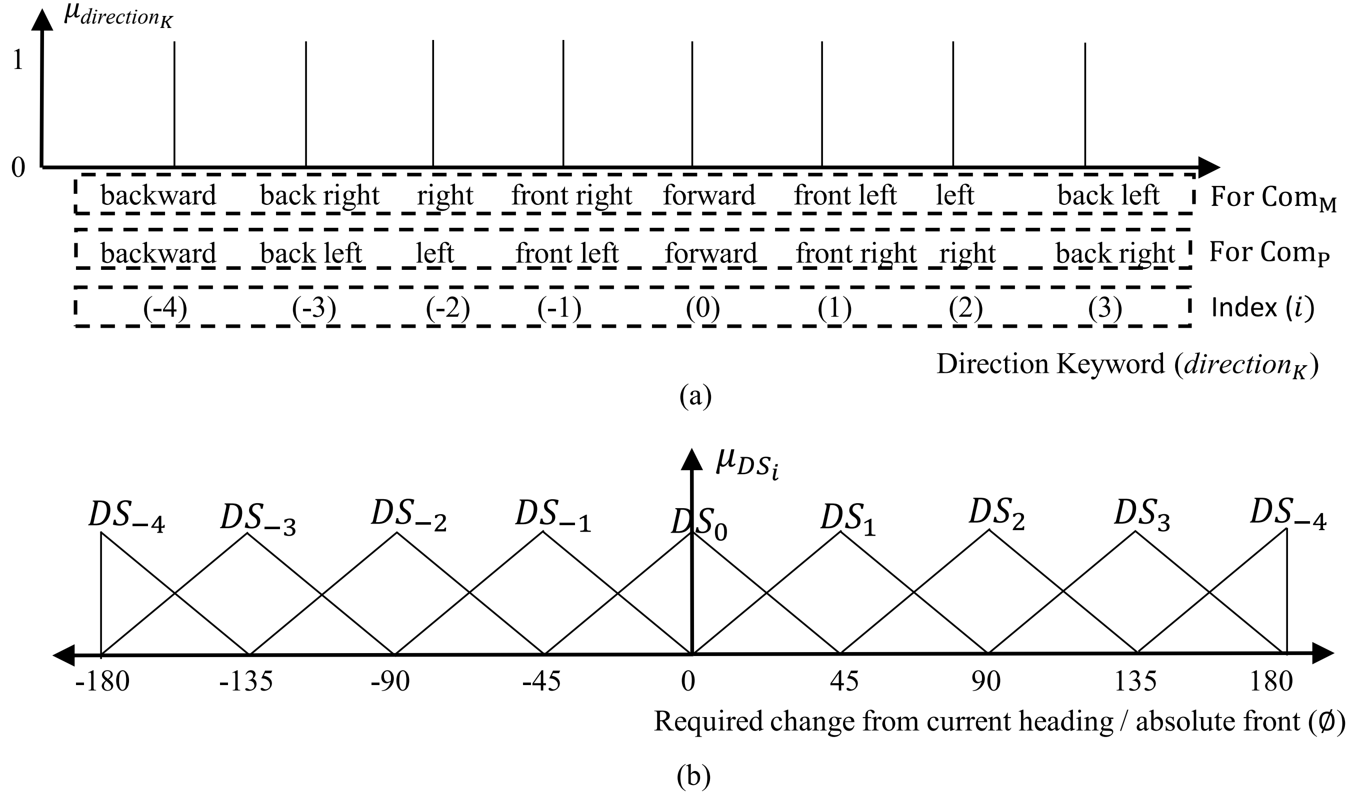


Figure 5.3: (a) shows the input membership function of the Direction Interpreter (DirI). It has singleton sets to represents the direction keywords (<direction$_{\mathrm{K}}$>). For motional commands and positional commands the indexes have to be interchanged for "left" and "right". Therefore, the keywords and index are linked differently as shown. (b) shows the output membership function. The i^{th} direction keyword in the input membership function is directly mapped to the i^{th} set of the output membership function yielding to a single input single output fuzzy system. It should be noted that the output membership function is a continues one and the ends represented by DS_{-4} are physically at the same position (-180° and 180° are physically the same).

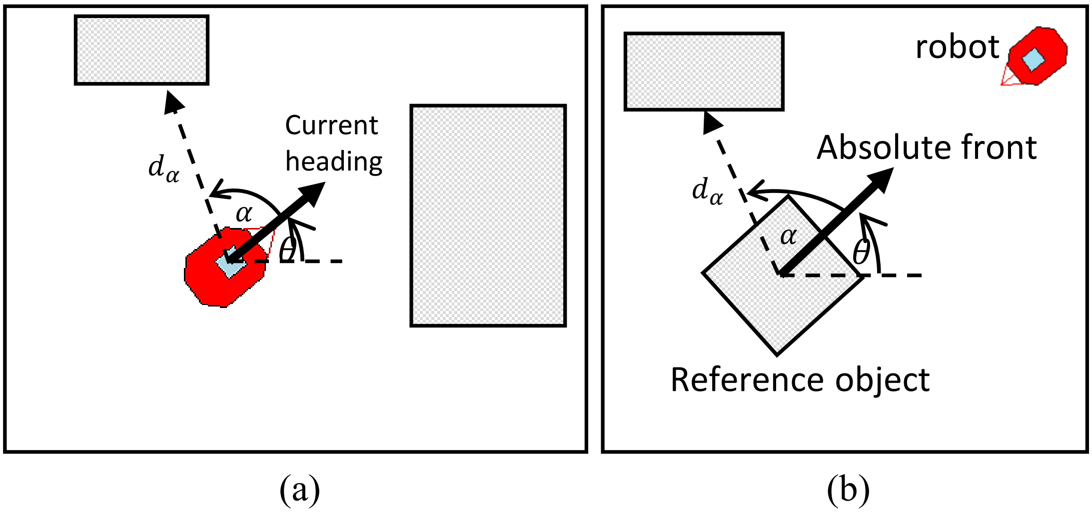


Figure 5.4: This explains the ways to obtain d_α in order to modify the output membership function of the Direction Interpreter (DirI). (a) shows an instance where d_α is obtained for a motional command. For motional commands, free space around the robot is considered for the weighting. Hence, d_α is the distance to the nearest obstacle/object from the robot in the direction that creates angle of α with the current heading of the robot. (b) shows an instance where d_α is obtained for a position command. For positional commands, free space around the reference object indicated by <Ref> is considered for the weighting. Hence, d_α is the distance to the nearest obstacle/object from the reference object in the direction that creates angle of α with the absolute front of the reference object. The absolute front of the reference object is defined based on the point of view of the robot.

$$\begin{aligned}(DS_i)_L &= (DS_i)_C - \beta \\ (DS_i)_U &= (DS_i)_C + \beta\end{aligned} \tag{5.2}$$

$$(DS_i)_C = i\beta \tag{5.3}$$

The, $\omega_{DS_i}(\alpha)$ is normalized for each i to have the variation in between [0,1]. The normalized weighted activation degree of the i^{th} Direction Set, $\hat{\omega}_{DS_i}(\alpha)$ is obtained from (5.4). The defuzzified output of the fuzzy inference system, ψ can be obtained from the Center of Area (COA) method as given in (5.5), where

$\mu_{dir_K}(i)$ is the activation degree of the i^{th} set of input membership function for the direction keyword (i.e., <direction$_{\text{K}}$>). The output of the Distance Interpreter, ϕ can be obtained as given in (5.6) and the meaning of ϕ depends on the type of the corresponding user command. If the user command is a motional command, then ϕ is the interpreted moving direction for the robot and it is achieved by changing the heading angle of the robot to ϕ. If the user command is a positional command, then ϕ is the angle to the destination position of the robot measured around the center of the reference object from the X-axis.

$$\hat{\omega}_{DS_i}(\alpha) = \frac{\omega_{DS_i}(\alpha)}{max(\omega_{DS_i}(\alpha))} \tag{5.4}$$

$$\psi = \frac{\sum_{i=-4}^{3} \sum_{\alpha=(DS_i)_L}^{(DS_i)_U} \hat{\omega}_{DS_i}(\alpha).\alpha.\mu_{dir_K}(i)}{\sum_{i=-4}^{3} \sum_{\alpha=(DS_i)_L}^{(DS_i)_U} \hat{\omega}_{DS_i}(\alpha).\mu_{dir_K}(i)} \tag{5.5}$$

$$\phi = \theta + \psi \tag{5.6}$$

In order to improve the computation efficiency and also to simplify the implementation complexity, only 6 distinct values are considered for α within the given range when weighting the default membership function with the distances as given in (5.1). Therefore, in implementation of the weighting of membership functions, α for the i^{th} DC set is defined as, $\alpha = \{\alpha_1 = (DS_i)_L, \alpha_2 = (DS_i)_C - 30°, \alpha_3 = (DS_i)_C - 10°, \alpha_4 = (DS_i)_C + 10°, \alpha_5 = (DS_i)_C + 30°, \alpha_6 = (DS_i)_U\}$ and other intermediate steps are interpolated linearly in order to generate a continuous set. The defuzzified output is generated by calculating the Center of Area (COA) of this modified membership function.

5.3 Results and Discussion

5.3.1 Experimental setup

The proposed DirI has been integrated to MIRob and experiments have been conducted in two phases for the evaluation of the behavior and performance. The first phase of the experiment has been conducted for the preliminary verification of the direction interpretation ability of the proposed DirI. According to the results of the preliminary verification process, the deployment of the DirI to the system sometimes affect the distance interpretation ability of the system. Therefore, the overall functionality of the system (considering both distance and direction interpretation abilities as a one action) has been evaluated in the next phase of the experiments. In both phase of the experiments, the behavior and performance of the proposed system (i.e. the system with the adaptable directional perception) has been evaluated against a system with a fixed directional perception (i.e. system without the proposed DirI). The evaluations have been conducted with the aid of user studies and due attention has been paid to the guidelines and recommendation given in [85] for designing and performing human studies for human robot interaction experiments since the user studies are highly subjective in nature.

5.3.2 Preliminary Verification of Direction Interpretation Ability

Three sets of experiments have been carried out in an artificially created indoor environment (Size = 11.47 m^2 and free space = 9.47 m^2) in order to verify the direction interpretation ability of the DirI. A few snapshots of MIROb taken during the experiments are shown in Fig. 5.5. The first set of experiments has been carried out by disabling the abilities of the proposed DirI (in this case, abilities of the system are almost similar to the system explained in chapter 4). The second set of experiments has been carried out with the system enabled with

(a) (b) (c)

Figure 5.5: (a) shows the MIRob during an experimental scenario. (b) and (c) show the robot being operated through a joystick by the users during the user study.

the proposed DirI. During these two sets of experiments, the robot was initially positioned in the locations marked on the map shown in Fig. 5.6 and the robot was commanded with voice instructions. The responses of the robot and the vital variables of the system that have been recorded during the experiments are given in Table 5.1 for each user instruction. The modified output membership functions during the corresponding cases are shown in Fig. 5.7 along with the default membership function for the comparison. As the third set of experiments, a user study has been carried out with participation of 10 persons whose ages are in between 24-54 years (M = 31.4, SD = 11.96) in order to compare the abilities of the two systems. During the user study, the users were given a joystick that can be used in order to operate the robot and the users were asked to move the robot according to the command issued to them. The robot was placed on the same initial positions as similar to the first two sets of experiments and the corresponding user commands were issued to them. The users were participated one by one for the study and all the participants were given one trial for each case. The movements of the robot have been recorded and the moved heading angles have been obtained. The mean value and the standard deviation of the obtained heading angle of the each case during the user study are also given in Table 5.1.

Table 5.1: Results of the experiments: Preliminary verification of the DirI

	User command	**without Direction Interpreter**			**with Direction Interpreter**							**User study ϕ°**	
		Initial position (X cm,Y cm,θ°)	ϕ°	Destination (X cm,Y cm,θ°)	d_α (cm)[1]				ψ°	ϕ°	Destination (X cm,Y cm,θ°)	Mean	Standard deviation
					d_{α_2}	d_{α_3}	d_{α_4}	d_{α_5}					
1	move medium forward	(212,204,-90)	-90	(211,180,-90)	218	184	48	57	-9	-99	(198,94,-98)	-98	5.96
2	move medium right	(50,121,-40)	-130	(35,101,-129)	28	42	83	94	-83	-123	(36,92,-123)	-120	5.82
3	move little left	(274,199,-2)	88	(282,255,88)	51	141	143	158	93	91	(272,255,90)	91.6	1.96
4	move far right	(197,283,-90)	-180	(132,283,-179)	79	69	73	92	-89	-179	(134,279,-179)	-180	2.98
5	move far forward	(163,112,0)	0	(185,112,0)	149	29	30	151	0	0	(186,112,0)	2.9	26.12
6	move far forward	(212,204,-90)	-90	(211,170,-90)	195	171	40	50	-9	-99	(195,66,-98)	-100.6	5.64

[1] There is no effect for the system from d_{α_1} and d_{α_6} since $\mu_{DC_i}(\alpha_1)$ and $\mu_{DC_i}(\alpha_6)$ are zero. Therefore, those two values are not given here.

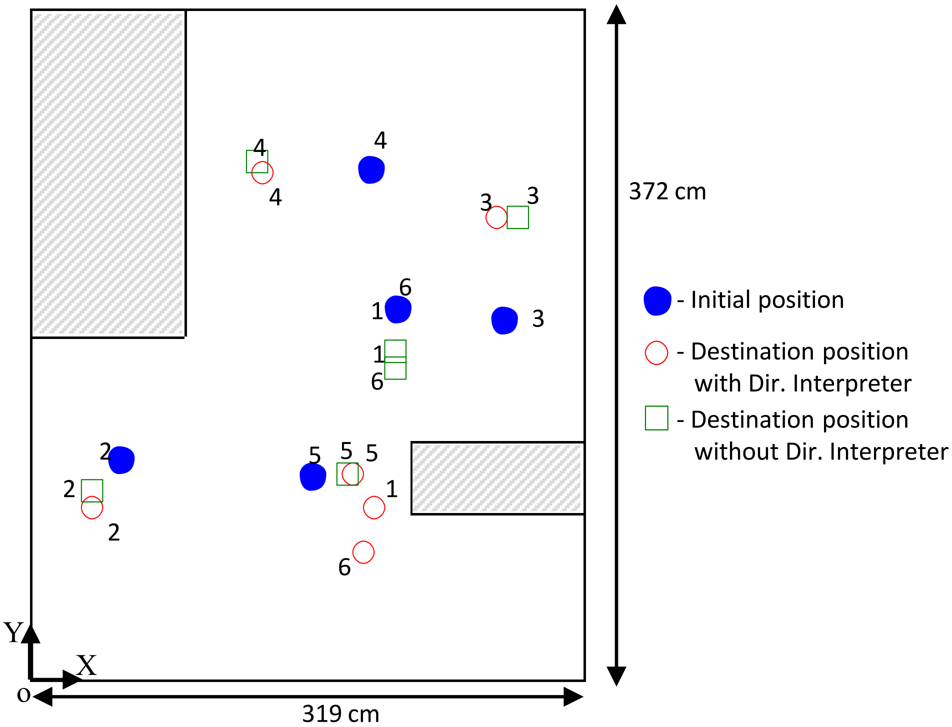


Figure 5.6: The initial and destination positions of the robot during the experiments are marked on the map with corresponding case numbers given in Table 5.1. The map is drawn to a scale. However, it should be noted that the markers do not represent the size of the robot. The obstacle/objects in the environment are marked as shaded areas.

In the case 1, the robot was commanded, "move medium forward". The robot without the DirI moved with a heading angle of -90° since the meaning of "forward" was fixed as the current heading. The distance meant by "medium" was quantified by the Distance Interpreter(DisI) as 23 cm and subsequently the robot moved to the destination location marked on the map (also the coordinates are given in Table 5.1). The system with the DirI interpreted the direction meant by the linguistic term "forward" as -99° by considering the environment based fuzziness associated with directional linguistic terms. The quantified distance in this case was 118 cm and the robot moved to the destination position. The output membership function modified by weighting the base membership function with the distances around the robot in this scenario is shown in Fig. 5.7(a). In this scenario, only the 0^{th} set is effective since the effects of other sets are null due to $\mu_{direction_K}(i) = 0$ for $i = \{-4, -3, -2, -1, 1, 2, 3\}$. The COA of this is the defuzzified output, ψ and it was $-9°$ which is different from the COA of the default membership function. Therefore, this validates the modification of the output membership function according to the environment and subsequently the modification of the robot's directional perception. The mean value of the heading angle obtained from the user study is -98°. According to t-test, the heading angle of the system without the DirI and the value obtained from the user study has a statistically significant difference ($P < 0.05$). Therefore, the system without the DirI is not capable of interpreting the directional information in human like manner. There is no statistically significant difference between the heading angle of the system with the DirI and the mean of the user study ($P = 0.61$). Therefore, the performance of the proposed DirI is acceptable in this case.

In case 2, the mean value of the heading angle obtained from user study is -120°. The moved heading angle of the system without the DirI is -130° and this is different from the value obtained from the user study with a statistically significant margin ($P < 0.05$). Therefore, the performance of the system without the DirI is not effective in this case too. The heading angle of the system with the DirI is -123° and the difference is not statistically significant ($P = 0.14$). Similarly, case 3 exhibits a similar behavior by rejecting the heading angle decided by the system without the DirI with a statistically significant margin ($P < 0.05$) while confirming a statistically non-significant difference with the mean of user study and heading decided by the DirI ($P = 0.36$). Therefore, this validates the ability of the system in replicating the natural directional perception of humans.

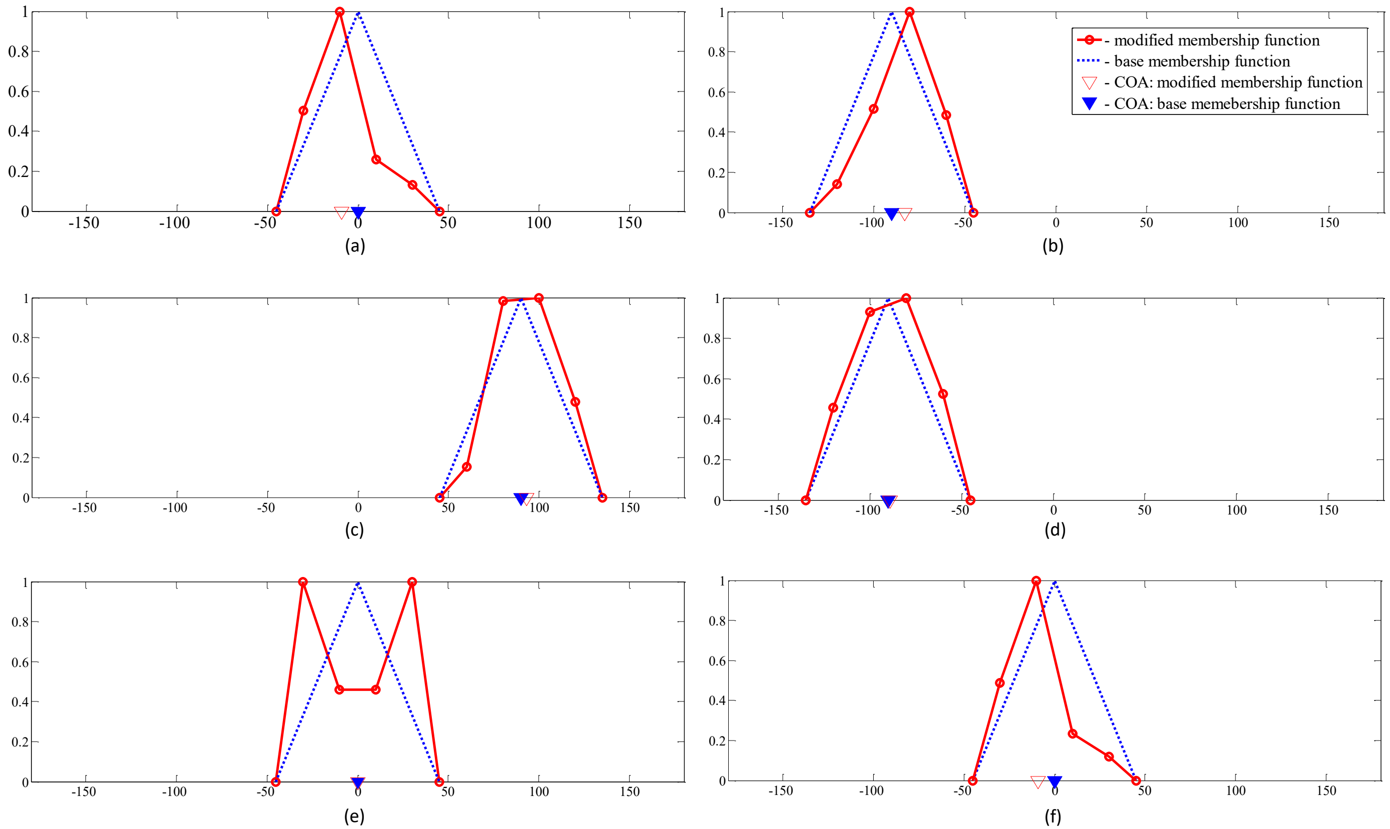


Figure 5.7: The output membership function after weighting with distances around the robot during the given experimental cases are shown here along with the base output membership function ((a) case1, (b) case 2, (c) case 3, (d) case 4, (e) case 5 and (f) case 6). Only the effective set of a particular case is given in each sub diagrams and effects of other sets are null for that case due to $\mu_{direction_K}(i) = 0$. Circles represent the points where the weighting is done (i.e. α_{1-6}). The center of areas of the both membership functions are also annotated for the each case for the comparison.

In case 4, the differences between the mean of the user study and the heading angle decided by the two systems are not statistically significant (P = 0.32 and P = 1 respectively). Therefore, in this case performance of the both systems are acceptable. In this scenario, the environment in the frontal direction is almost constant in whole region. Therefore, the directional perception modification is not significant and the responses of the system without the DirI are also acceptable for such situations. Furthermore, the modified output membership function is almost similar to the default output function (shown in Fig. 5.7(d)) and its distribution is symmetrical around the center of the default function. Therefore, the COA of the modified function is almost similar to the base case (There is only 1° difference due to the small deviations of the sensory readings). This validates that the DirI is capable of effectively evaluating this kind of scenarios also.

In case 5, the differences between the heading angles generated by the systems and the mean of user study are not statistically significant. However, the standard deviation of the heading angle decided by the users is comparatively large (SD = 26.12). Therefore, the results and the arrangement of the scenario have been further investigated in order to make a conclusion. In this situation, frontal region is almost symmetric in both right (i.e., in $-\theta$ direction) and left side (i.e., in $+\theta$ direction) directions . However, there is an obstacle just confront with the current heading angle. Therefore, the modified membership function shown in Fig. 5.7(e) is almost symmetrical around the center of the default function. Therefore, the COA is the same as the default case. Three different clusters of users can be seen from the results of the user study; a cluster of users who drove the robot toward free area by changing the heading towards the right, another set of users who drove the robot toward the free area by changing the heading toward the left, and a cluster of users who just drove the robot without changing the heading angle. Therefore, in this kind of situation, directional perception of humans also varies significantly. However, there may have been a possibility in deciding the direction by the users based on the parameters other than the environment such as information conveyed through gestures of the command issuer and context.

Therefore, effects of such parameters should be analyzed for future improvement of directional information interpretation ability of robots.

In case 6, the robot was placed exactly at the same initial position of the case 1. However, in this case the robot was commanded "move far forward". In this case, the deviation of the mean of the user study from the current heading is comparatively higher than the case 1 (case 1, ϕ = -98° and case 6, $\phi = -100.6°$). This may be due to the difference of the intended moving distance (in case 1, distance is "medium" and in case 6, distance is "far"). However, the system with the DirI moved with the same heading as in the case 1 since the system is not capable of modifying the perception according to other parameters than free space distribution of the surrounding that was almost the same as case 1 in here too. The modified output membership functions are almost the same in the both cases (shown in Fig 5.7 (a) and (f)). The ability of the proposed system is acceptable since the difference between the mean of user study and the heading decided by the system is not statistically significant. However, there is very very slight trend toward the statistically significant difference compared to the case 1 (in case 1, P = 0.61 but in here P = 0.39). The reason for the very slight trend of significant difference in case 2 (the reason for the lower P value, P = 0. 14) may be due to this issue. Therefore, it would be beneficial to investigate the effects caused by the distance descriptors toward the directional perception for improving the ability of interpreting directional notions by robots.

In most of the cases discussed in here, the direction decided by the system without the DirI (i.e., with fixed meanings for directional notions) is significantly different from the natural directional perception of humans identified from the user study. Therefore, such fixed interpretation is not effective. The proposed DirI has produced results that are not significantly different from the natural directional perception of the human users. However, the effectiveness can be further improved by considering the other factors that influence the directional perception for the adaptation of perception as explained above.

The deployment of the DirI has improved the robot's direction perception to replicate the directional perception of humans. However, the distance notions quantified by the system is significantly affected in some cases as a results of the modification of the perceptive distance (i.e., D). As an example, in the case 1, the distance moved by the robot were 23 cm and 118 cm for the system with the DirI run and without the DirI run respectively. Therefore, in order to analyze the overall performance, the overall action of the robot has been evaluated in the second phase of the experiment discussed in section 5.3.3

5.3.3 Overall Performance Comparison

The arrangement of the experimental environment is given in the map shown in Fig. 5.8. It had 3 different rooms with heterogeneous characteristics. At the start, the robot was initialized with an updated navigation map of the environment and the Robot Experience Model. Therefore, the robot was well aware of the arrangement and the characteristics of the environment during the experiments.

The user study has been conducted with participation of 12 users whose mean and standard deviation of age are 26.8 and 4.1 years respectively. The users were taken one by one to the experiment and they were advised about the structures of the user commands that can be understood by the robot. Each user has given 6 occasions to interact with the robot for each of the two systems (i.e., system with fixed directional perception and adaptable directional perception). These instances were chosen randomly deciding the initial position of the robot. The users were given the freedom to decide the user instructions. However, the users were asked to include 3 motional commands and 3 positional commands for those 6 instances and the same 6 instances were repeated to the other system. In order to minimize the subjectivity, the users were not informed about the system (either with fixed directional perception or with adaptable directional perception) that they are interacting in a particular run. After each run, the user was asked to rate the action of the robot in the scale 0-100 similar to the evaluation approach used in [3], where 100 indicates the perfect agreement and 0 indicates the null agreement. The User Rating (UR) given by the user depends on the final position of the robot. Therefore, it reflects the assessment of both direction and distance interpreted by the robot.

The results obtained from the 1st user for 6 runs using both systems are given in Table 5.2 as sample results. The parameters related to the DisI for the corresponding cases are given in Table 5.3. The initial and final positions of the robot during these experimental runs are marked on the map shown in Fig. 5.6 with corresponding indexes given in Table 5.2. The modified output membership functions of the DirI due to the weighting with the free space in these cases are shown in Fig. 5.9.

In case 1, the robot was initially placed on the location 'I_1' and the robot was commanded "move far to the left" by the user 1. This is a motional command, and <$\mathrm{dis}_{\mathrm{M}}$> and <$\mathrm{dir}_{\mathrm{K}}$> were "far" and "left" respectively. Therefore, in order to fulfill the user command the robot had to interpret the distance meant by "far" and the direction meant by "left". In the run of the system with fixed directional perception (i.e., the system with the DirI), the direction interpreted by the robot for "left" was fixed as current heading (i.e., θ) + 90° as explained in section 5.1. Therefore, the heading angle for the movement was decided by the robot as 85°. The distance meant by <$\mathrm{dis}_{\mathrm{M}}$> was quantified by the Distance Interpreter (DisI) as 28 cm based on the perceptive distance (D = 34 cm), room size (= 15.08 m^2), and the free space (= 12.95 m^2). As a result of this interpreted distance and direction, the robot moved to location 'F_1'. The action of the robot in this run has been rated by the user 1 by giving a User Rating (UR) of 24. In the system with adaptable directional perception case (i.e., the system with the purposed DirI), the robot was initially placed on the same location and issued the same user command. In this run, the direction interpreted by the system was a change of 101° from the current heading yielding to the heading of the movement to 97°.

Table 5.2: Sample results of the experiment: parameters related to the interpretation of directional notions by the DirI

	User command	**Initial position** (X cm,Y cm,θ°)	**with fixed directional perception** ϕ°	Destination[1] (X cm,Y cm,θ°)	UR	**with adaptable directional perception** d_α (cm)[2] d_{α_2}	d_{α_3}	d_{α_4}	d_{α_5}	ψ°	ϕ°	Destination[1] (X cm,Y cm,θ°)	UR
1	move far to the left	I_1(369,259,-4)	85	F_1(378,290,85)	24	10	45	187	105	101	97	F'_1(346,455,96)	84
2	move a little left	I_2(186,392,90)	180	F_2(162,392,180)	67	28	53	122	60	96	-175	F'_2(161,388,-175)	70
3	move far forward	I_3(-188,-64,-64)	-64	F_3(-150,-121.-63)	50	89	184	110	29	-6	-69	F'_3(114,-235,-67)	84
4	move near to the front of the sink	I_4(274,134,176)	-90	F_4(371,491,90)	64	140	289	227	60	-5	-94	F'_4(367,492,90)	72
5	move near to the left of the switch board	I_5(-224,130,0)	90	F_5(382,440,-90)	71	5	33	138	99	-77	103	F'_5(364,439,-90)	72
6	go near to the front of the cupboard	I_6(-169,-65,-73)	90	F_6(-122,-279,-90)	66	42	128	268	128	6	96	F'_6(-128,-279,-90)	78

[1] The destination positions are decided based on the outputs of both DisI and DirI. The parameters related to the DisI in interpreting the distance notions in the corresponding cases are given in Table 5.3.
[2] It should be noted that the effects to the interpretation from d_{α_1} and d_{α_6} are null since $\mu_{DS_i}(\alpha_1)$ and $\mu_{DS_i}(\alpha_6)$ are zero. Therefore, those two values are not displayed here.

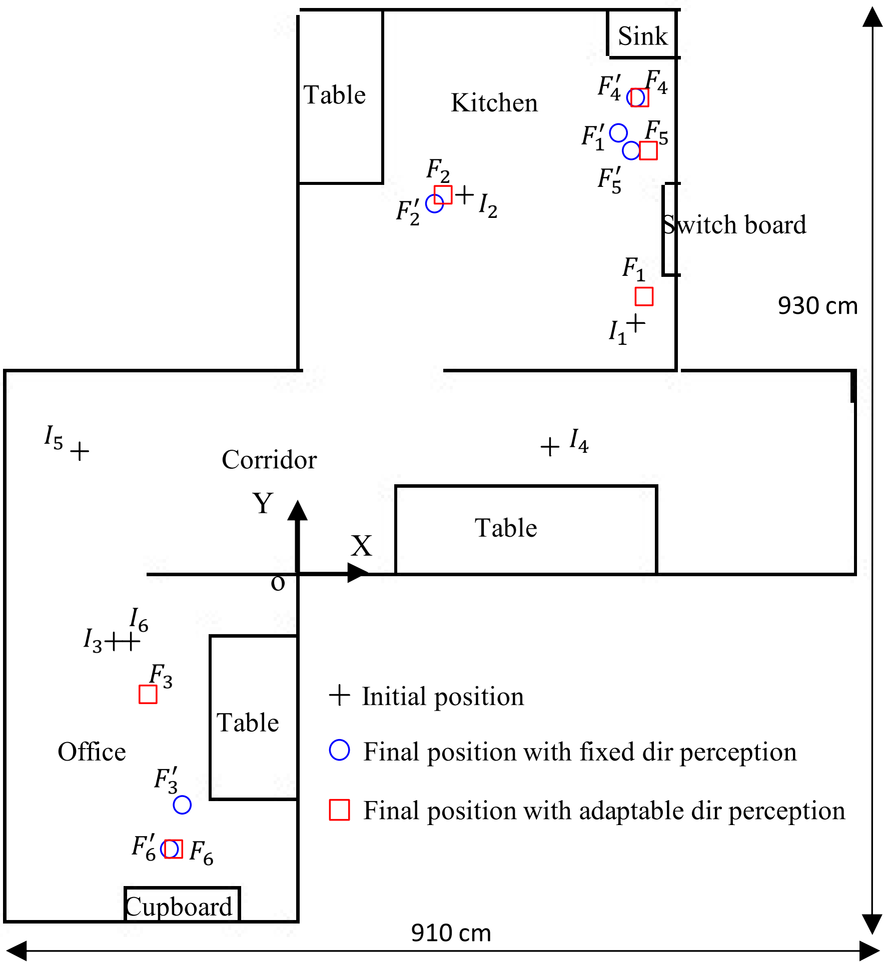


Figure 5.8: The initial and final positions of the robot during the execution of the cases given in Table 5.2 are marked on the map with corresponding letter indexes. This map is drawn to a scale in order to visualize the characteristics of the experimental environment. However, it should be noted that the markers do not represent the actual size of the robot.

Table 5.3: Parameters related to the interpretation of distance notions by the DisI

Case	Room size (m^2)	Free space (m^2)	with fixed directional perception D (cm)	with fixed directional perception D_{out} (cm)	with adaptable directional perception D (cm)	with adaptable directional perception D_{out} (cm)
1	15.08	12.95	34	28	243	198
2	15.08	12.95	60	25	65	27
3	11.50	9.27	108	86	234	186
4	15.08	12.95	100	38	100	38
5	15.08	12.95	130	37	130	37
6	11.50	9.27	94	36	94	36

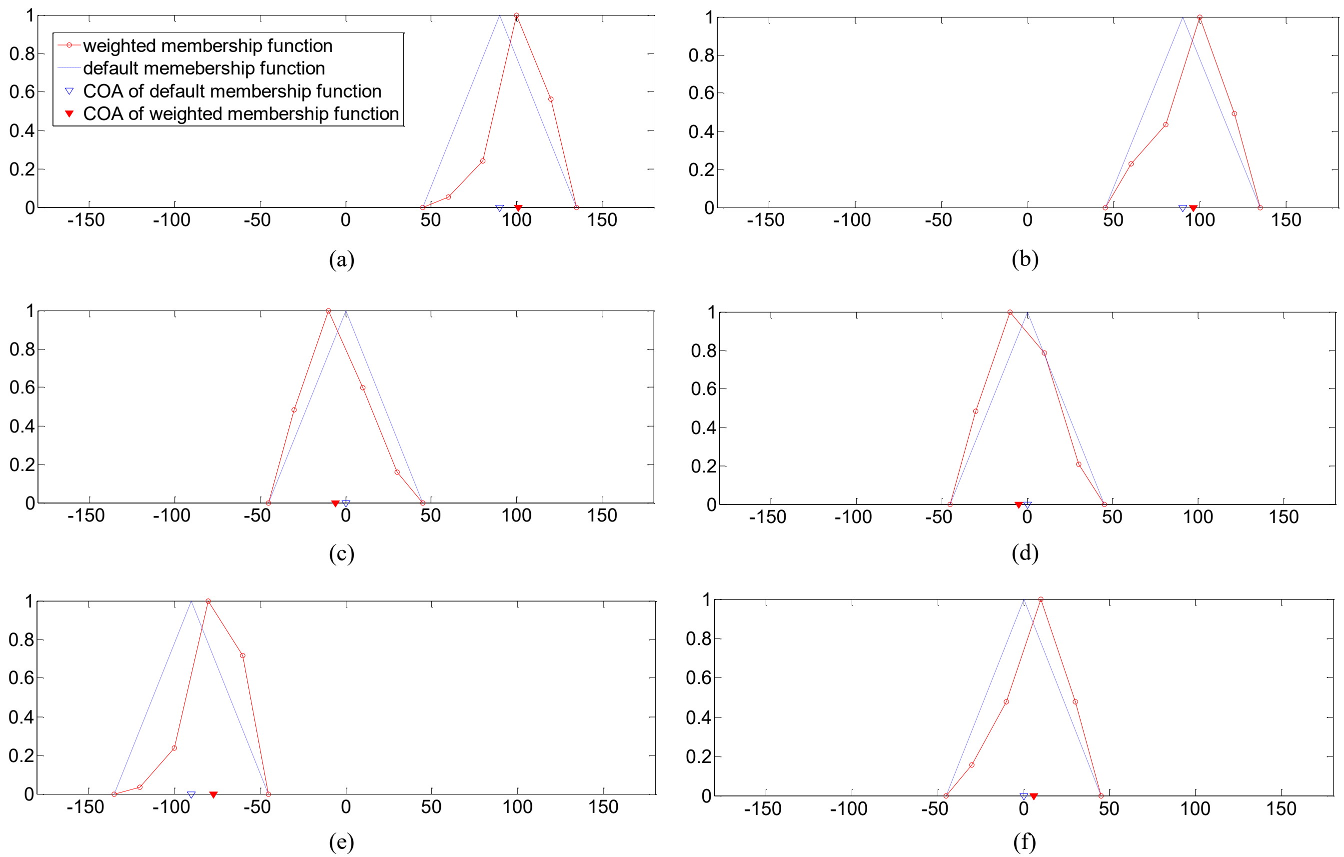


Figure 5.9: The output membership functions plotted here show the adaptation of the perception of directional notions after weighting the default perception with the available free space for the cases given in Table 5.2. (a), (b), (c), (d), (e), and (f) represent case 1, 2, 3, 4, 5, and 6 respectively. It should be noted that only the effective Direction Sets (DS) are plotted here for a particular instance and non-effective DS_i due to $\mu_{dir_K}(i) = 0$ are not shown.

The modified membership function due to the arrangement of the environment is shown in Fig. 5.9 (a). In this instance, only DS_2 was effective since the effects of other sets are null due to $\mu_{dir_K}(i) = 0$ for $i = \{-4, -3, -2, -1, 0, 1, 3\}$. As a result of adapting the directional perception according to the environment, the defuzzified output of the system was different from the COA of the default perception. This exhibits the directional perception adaptation based on the environment setting. Because of this interpreted direction, the perceptive distance (D) was significantly different from the previous case (i.e., system with fixed directional perception). Therefore, the distance interpreted by the system was 198 cm ($D = 243$ cm, free space and room size were the same as previous run) and the robot's destination position was 'F_1''. The action of the robot has been rated as 84 by the user. The increase of the UR shows an enhancement of the user agreement for the system with adaptable directional perception (i.e., the system with the DirI) than the system with fixed directional perception in this case.

In case 2, the initial position of the robot was 'I_2' and the robot was commanded with the motional command "move a little left". The system with fixed directional perception run, the heading angle decided by the system was 180° and the robot moved 25 cm resulting the destination at location 'F_2'. The UR for this run was 67. The system with adaptable directional perception run, the heading angle for the movement was -175° and the robot moved 27 cm resulting the destination at location 'F_2''. The UR for this run was 70. The URs for the two systems were not significantly different since the resulted destination positions had only a slight difference.

In case 3, the robot was commanded with the motional command, "move far forward". In the system with fixed directional case the robot moved to location 'F_3' by considering the direction meant by forward as the current heading (i.e., $-64°$) while in the system with adaptable directional perception case, the robot moved to location 'F_3'' by considering the direction meant by "forward" as heading of $-69°$. The corresponding URs were 50 and 84 respectively for the two runs.

In case 4, the initial position of the robot was 'I_4' and it was commanded with the positional command "move near to the front of the sink". In this case, the robot had to quantify the distance meant by "near" and the direction meant by "front" with respect to <Ref> (i.e., sink). The robot moved to location 'F_4' in the run where the system had fixed directional perception. In the system with adaptable direction perception run, the robot moved location 'F_4'', which is slightly deviated towards the free area with respect to the final position of the previous run. Therefore, the UR for the system with adaptable perception (UR = 72) got a slightly higher value compared to the system with fixed directional perception (UR = 64).

In case 5 and 6, the robot was commanded with positional commands. The system with adaptable directional perception runs, the robot moved to locations which are slightly deviated towards low congestions areas with respect to the moved positions in the system with fixed directional perception runs. The system with adaptable directional perception got higher UR with respect to the system with fixed directional perception. However, in case 5, the URs for the two systems were almost the same (71 and 72) even though the direction had a slightly large deviation (deviation was 13°). In system with the fixed directional perception case, the robot was not exactly settled on the direction deicide by the system since the robot cannot reach that position due to the limitation of the space for the occupancy of the robot. Therefore, the location was already deviated toward the free area due to that. This would be the reason for getting almost the same user rating for the two systems.

Similarly, all the 12 participants were asked to operate the robot 6 runs for each of the system (i.e., with the system with fixed directional perception and system with adaptable directional perception). This yields to 72 effective cases for each system. The mean value of the User Rating (UR) was calculated for both systems based on the individual UR for each run. The calculated mean UR scores for the two systems are given in Fig. 5.10(a) with error bars. The distributions of the UR

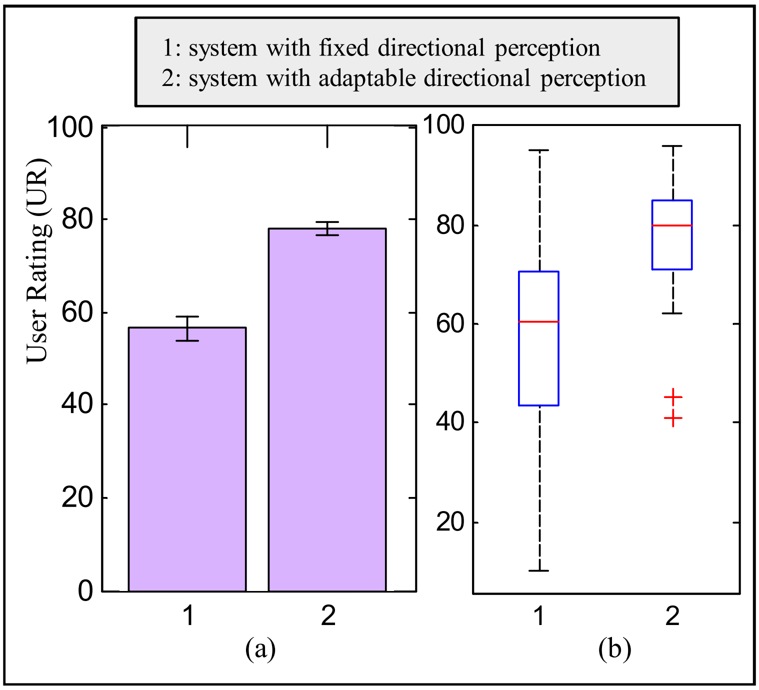


Figure 5.10: (a) shows the mean values of the user ratings for the two systems with error bars. The error bars represent the standard error. (b) shows the distribution of user ratings as a boxplot. The box plot has the usual standard notation; box: interquartile range, horizontal line: median, whiskers: minimum and maximum, and plus sign: outliers

scores are given in Fig. 5.10(b) as a box plot for better visualization of the results. The system with the proposed DirI got a mean user rating of 77.7 while the system with fixed direction interpretation (i.e., system without the DirI) got mean UR of 56.2. The difference between the means of UR is statistically significant ($P<0.05$) according to the t-test. Therefore, it can be concluded with 95% confidence that the user agreement for the system with adaptable perception is par above the user agreement for the system with fixed directional perception. Furthermore, the performance improvement caused to the understanding of fuzzy notions due to the addition of adaptable directional perception is noteworthy since Cohen's d value of greater than 0.7 can be observed from results of the user rating (since Cohen's d value greater than 0.7 is considered as a large effect [88]).

5.4 Summary

A method has been proposed in this chapter for enhancing the interpretation of the fuzzy notions in motional and positional navigation command by adapting the robot's directional perception based on the environmental setting. The major improvement of the proposed system over the existing approaches is the system is capable of interpreting the directional notions in motional and positional navigation commands by considering the fuzziness associated with natural language descriptors instead of fixed interpretations.

The directional notions in user instructions are interpreted by a fuzzy inference system that has been designed in such a way that it can replicate the natural human behavior. The perception of the directional notions is adapted by modifying the output membership function of the fuzzy inference system according to the available free space around the robot or the reference object.

Experiments have been conducted in order to evaluate the performance improvement caused to the understanding of navigational commands by the robot due to the deployment of the proposed method for adapting the robot's directional perception. The performance of the system with the adaptable directional perception (i.e., system with the prposed DirI) has been compared against a system with fixed direction perception through user studies. According to the obtained experimental results, the fuzzy navigational command understanding ability of the system with the adaptable directional perception surpasses the ability of the system with fixed directional perception with a significant margin.

CHAPTER 6

ADAPTING ROBOT'S PERCEPTION OF UNCERTAIN INFORMATION BASED ON THE ENVIRONMENT AND USER FEEDBACK

Notably, the perception of uncertain terms varies from person to person. In real-world situations, peers mutually adapt to align with each other's perceptions. Therefore, robots must also be capable of this behavior to increase user satisfaction. However, the methods discussed in chapter 4 lack a means of adapting a robot's perception toward user expectations based on corrective measures received from the user.

This chapter proposes a novel method of interpreting uncertain terms contained in navigational user commands based on the environment and prior correction measurements received from the user. The main advantage of the proposed method over the existing systems is that the proposed system is capable of concurrently adapting to the environment while learning from user feedback. In order to realize this, the Distance Interpreter (DisI) of Uncertain Information Understanding Module (UIUM) is reimplemented with fuzzy neural networks that can provide the required learning and adapting abilities. As similar to the work explained in chapter 4, the DisI of the UIUM is used to assign quantitative values to distance-related uncertain terms such as "far", "close" and "little" in user commands. In addition to that, the Feedback Evaluation Module (FEM) is integrated into the Intelligent System for Understanding Uncertain Information (ISUUI) for evaluating the error of an action reflected by a feedback statement

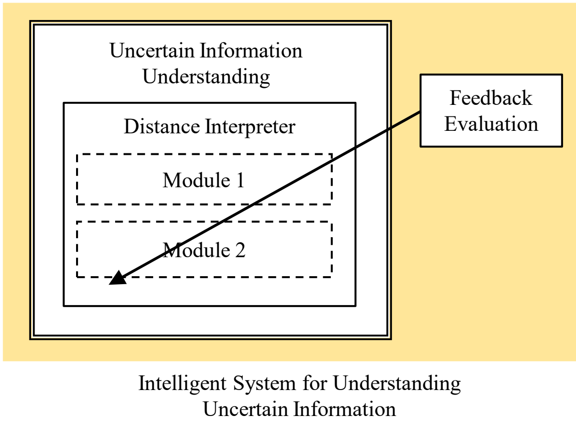


Figure 6.1: Structure of the ISUUI with the modified DisI

given by a user. The arrangement of the ISUUI with these modifications is shown in Fig. 6.1.

6.1 Reimplementation of the DisI using fuzzy neural networks

The perception of uncertain terms strongly depends on the spatial information of the surrounding environment. In addition, the perception of uncertain terms varies with the expectations of the user. Therefore, the DirI must be implemented such that it can learn from user feedback to adjust its perception to match the expectations of the user in addition to adapting to knowledge about the environment. Systems based on fuzzy logic and fuzzy neural networks are often used to understand the meaning of natural language user commands [81]. However, the existing systems cannot concurrently adapt to both the spatial information about the environment perceived from sensory information and the corrective feedback received from the user. Therefore, the DisI is reimplemented with fuzzy neural networks that can perceive the environment by means of spatial information inputs while concurrently learning from user feedback. Two independent fuzzy neural networks have been developed for the separate interpretation of uncertain terms related to motional and positional information. Submodule 1 is used to interpret distance-related uncertain terms in motional commands, i.e.,

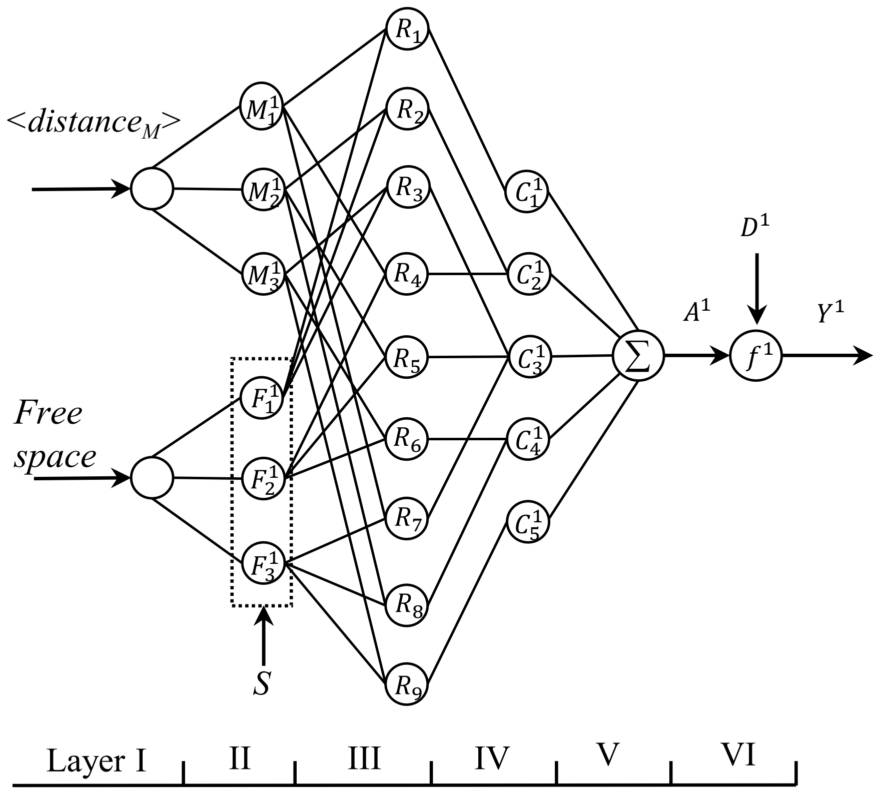


Figure 6.2: Structure of submodule 1 of the DisI. The fuzzy neural network consists of 6 layers. The <distance$_\mathrm{M}$> and the available free space are the inputs to the network. The membership functions for the free space are adjusted according to the room size (S). Therefore, the nodes that represent the free-space membership functions, which are bounded by a dotted line, take S as an input. The activation transfer function f^1 depends on the perceptive distance D^1.

when executing a robot action of type I or II. Submodule 2 is used to interpret uncertainties related to positional information in user commands; therefore, it is used when executing type III robot actions. The structures of submodules 1 and 2 are shown in Fig. 6.2 and Fig. 6.3, respectively.

Layer I of each submodule is the input layer, and it contains two types of nodes for acquiring inputs: for the 1st submodule, these nodes correspond to the <distance$_\mathrm{M}$> in the user command and the free space available in the environment, and for the 2nd submodule, they correspond to the size of the object of interest and the available free space. The neurons in this layer transmit external input signals directly to layer II, which is the fuzzification layer. The neurons in

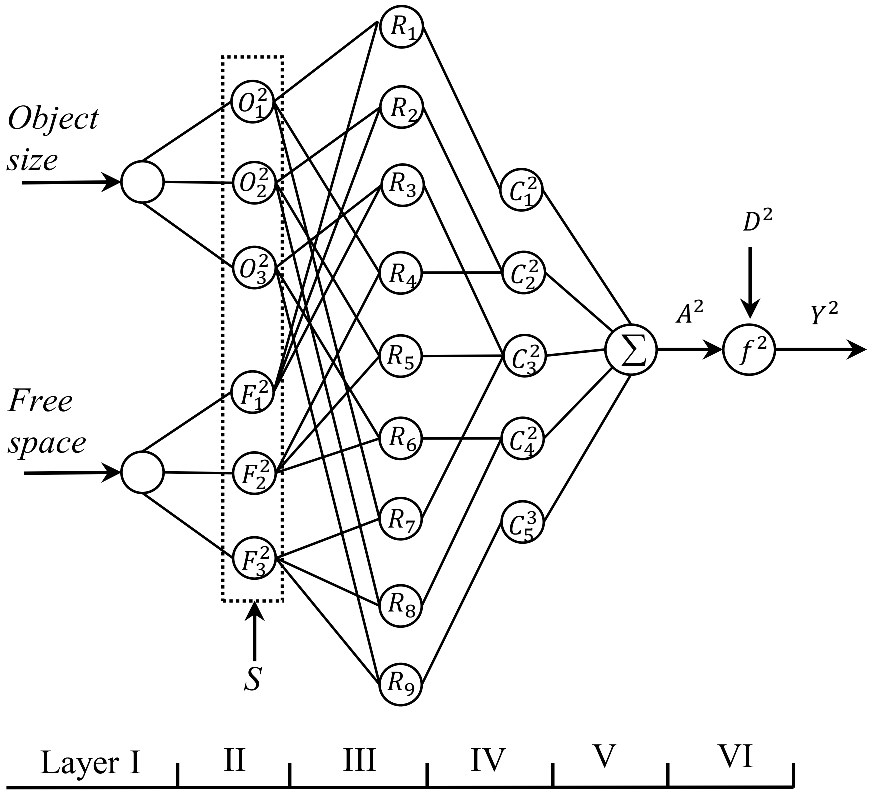


Figure 6.3: Structure of submodule 2 of the DisI. The fuzzy neural network consists of 6 layers. The size of the object of interest and the available free space are the inputs to the network. The membership functions for the size of the object of interest and the free space are both adjusted according to the room size (S). Therefore, the nodes that represent the input membership functions, which are bounded by a dotted line, take S as an input. The activation transfer function f^2 is adjusted according to the perceptive distance D^2.

this layer represent fuzzy sets used in the antecedents of fuzzy rules for the action modifier and the free space. The membership functions for the free space and size of the object of interest are adjusted according to the size of the occupied room (S). The size of the room, the available free space and the size of the object of interest can all be retrieved from the knowledge contained in the environment layer of the Robot Experience Model (REM). Layer III is the fuzzy rule layer. Each neuron in this layer corresponds to a single fuzzy rule. A fuzzy rule neuron receives inputs from the fuzzification neurons that represent the fuzzy sets in the antecedents of the corresponding rule. The algebraic product operator is used as the T-norm fuzzy operator; hence, the output of a neuron in this layer is the algebraic product of the incoming signals. Layer IV is the output membership layer. The neurons in this layer represent the fuzzy sets used in the consequents of the fuzzy rules, and an output membership neuron combines all of its inputs using the fuzzy union operator. Any node C_i^k in the k^{th} submodule, where $i = 1, ..., 5$, represents a triangular membership function with a center of $a_i^k \epsilon [(a_i^k)_L, (a_i^k)_H]$ and a width of $b_i^k \epsilon [(b_i^k)_L, (b_i^k)_H]$.

Layer V is the defuzzification layer. It takes the output fuzzy sets clipped by the respective integrated firing strengths and combines them into a single fuzzy set. The sum-product composition method can be used to simulate the center-of-area method of defuzzification for a Mamdani fuzzy system [75], and the defuzzification output is obtained from (6.1), where A^k is the output of layer V of the k^{th} submodule and μ_i^k is the integrated firing strength of the i^{th} output fuzzy set of the k^{th} submodule.

$$A^k = \frac{\sum_{i=1}^{5} a_i^k b_i^k \mu_i^k}{\sum_{i=1}^{5} b_i^k \mu_i^k} \tag{6.1}$$

Layer VI of each submodule consists of an activation transfer function that is used to scale the output. The transfer functions are given in (6.2), where Y^k is the output distance of the system, d_0 is the clearance of the robot, and the perceptive distance D^k is given in (6.3), where D_r is the distance from the robot to the object of interest or the nearest obstacle in the direction of its motion and D_{obj} is the distance between the object of interest and any other nearby object in the approach direction of the robot (as explained in chapter 4). The free space, the size of the object of interest, the room size, D_r and D_{obj} are all obtained from the environment layer of the REM based on sonar sensor readings and navigation maps.

$$Y^k = \begin{cases} (D^k - d_0)A^k & \text{if } k = 1 \\ (D^k - d_0)A^k + d_0 & \text{if } k = 2 \end{cases} \tag{6.2}$$

$$D^k = \begin{cases} D_r & \text{if } k = 1 \\ \frac{1}{2}[min(D_r, D_{obj})] & \text{if } k = 2 \end{cases} \quad (6.3)$$

The initial membership functions for submodules 1 and 2 are defined similarly to the membership functions for the system proposed in chapter 4 with slight modifications and are shown in Fig. 6.4. The initial membership functions for the output distance determine the initial connection weights of layer V, which are then adjusted based on user feedback using a backpropagation algorithm. The FEM is used to evaluate the normalized distance error ($\hat{e}$) of a particular movement by evaluating the user feedback given immediately after the robot performs an action of type I, II or III. The submodule that needs to be adjusted is chosen based on the robot action executed immediately before the feedback is received. If the previous action is an action of type I or II, then submodule 1 will be adjusted; if the previous action is a type III action, then submodule 2 will be adjusted. The robot action layer of the REM is used to identify the previous action. Then, membership parameter training (corresponding to network weight training) is performed for the i^{th} node of the k^{th} submodule with the execution of the $(t+1)^{\text{th}}$ action, as given in (6.4) and (6.5), where the $(t+1)^{\text{th}}$ action is a learning action, i.e., a type VI robot action. Here, η^k is the learning rate, and δ_a^k and δ_b^k are scalar constants that are used to maintain the variations of the parameters within the desirable ranges during the learning phase. If no feedback is given, then the weights are not adjusted.

$$a_i^k(t+1) = \begin{cases} a_i^k(t) + \eta^k \delta_a^k \hat{e} \mu_i^k & \text{if } a_i^k(t+1) \\ & \quad \epsilon[(a_i^k)_L, (a_i^k)_H] \\ a_i^k(t) & \text{otherwise} \end{cases} \quad (6.4)$$

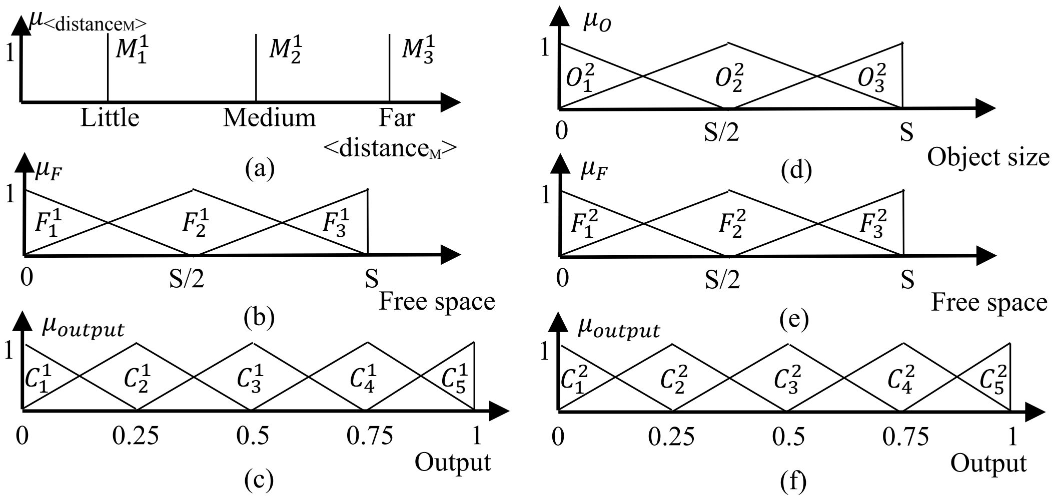


Figure 6.4: (a) represents the input membership functions for the <distance$_{\mathrm{M}}$>. It shows singleton membership functions labeled as M_1^1, M_2^1 and M_3^1 for the <distance$_{\mathrm{M}}$> “little”, “medium” and “far”, respectively. (b) represents the input membership functions for the free space. It has triangular membership functions labeled as F_1^1, F_2^1 and F_3^1, which are adjusted according to the size of the room (S). (c) represents the initial membership functions for the output of submodule 1. (d) represents the input membership functions for the size of the object of interest. It has triangular membership functions labeled as O_1^2, O_2^2 and O_3^2. (e) represents the input membership functions for the free space. It has triangular membership functions labeled as F_1^2, F_2^2 and F_3^2. These input membership functions are adjusted according to the size of the room (S). (f) represents the initial membership functions for the output of submodule 2.

$$b_i^k(t+1) = \begin{cases} b_i^k(t) + \eta^k \delta_b^k \hat{e} \mu_i^k & \text{if } b_i^k(t+1) \\ & \epsilon[(b_i^k)_L, (b_i^k)_H] \\ b_i^k(t) & \text{otherwise} \end{cases} \quad (6.5)$$

6.2 Feedback evaluation

Voice feedback includes directives from the user to modify the perception of the robot concerning uncertain terms. As an example, suppose that immediately after the robot has executed a type I action in response to a particular user command, the user issues a feedback statement of "too little". The user feedback "too little" indicates that the distance moved by the robot in response to the corresponding user command is less than the user expected. Furthermore, it conveys the intent of the user to adapt the system to generate a greater output distance on similar occasions in the future. Therefore, the robot should be able to extract the required degree of adjustment to adapt its perception to the user's expectation. However, such feedback statements do not contain precise quantitative values. Therefore, the quantitative meaning of a particular feedback statement must be evaluated to judge the required adjustment.

The FEM is implemented using a fuzzy inference system to assign a quantitative distance error (e) to a particular instance of feedback. It is assumed that the quantitative meaning of a feedback term depends on the user's observation, i.e., the immediately preceding action of the robot. Therefore, the previous output and the user feedback term are used as the inputs to the system. The output of the system is the evaluated distance error (e) corresponding to an instance of feedback on a particular robot action. The input and output membership functions of the system are shown in Fig. 6.5. The rule base of the system is shown in Table 6.1. Three singleton fuzzy sets, namely, Positive Error (PE), Negative

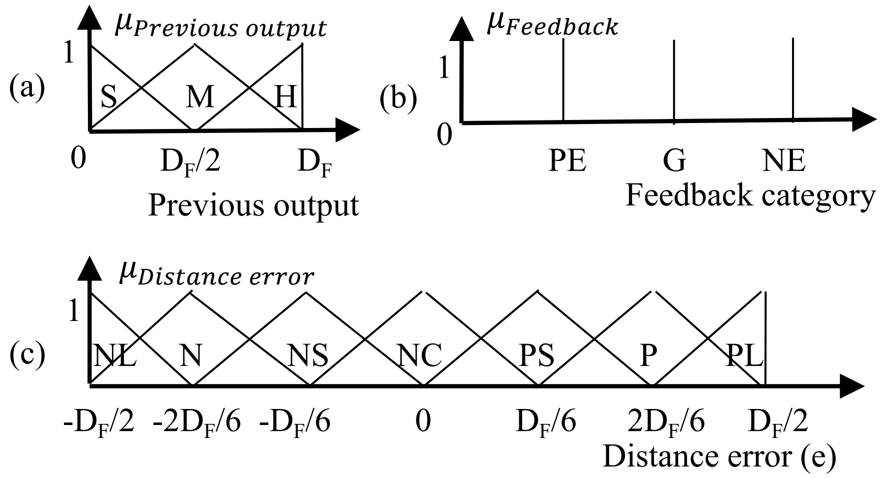


Figure 6.5: (a) represents the input membership functions for the previous output. It shows 3 triangular fuzzy sets, labeled as S: Small, M: Medium and H: High. (b) represents the input membership functions for feedback terms. It shows 3 singleton fuzzy sets, labeled as PE: Positive Error, G: Good and NE: Negative Error. (c) represents the output membership functions for the distance error. It shows 7 triangular membership functions, labeled as NL: Negative Large, N: Negative, NS: Negative Small, NC: No Change, PS: Positive Small, P: Positive and PL: Positive Large. The membership functions for the previous output and the distance error are adjusted according to D_F.

Error (NE) and Good (G), are defined as the membership functions for the feedback term. It is assumed that the user feedback will take different forms when feedback is given for different types of robot actions. For robot action types I and II, the possible feedback statements are assumed to be "too little", "too much" and "good", and for robot action type III, the possible feedback statements are assumed to be "too close", "too far" and "good". The mapping between the actual voice feedback statements and the feedback terms in the input membership functions is given in Table 6.2. The membership functions for the previous output and the distance error are adjusted according to D_F, where D_F is the maximum possible output for the particular robot action corresponding to the feedback. D_F is given in (6.6). The previous output ($Y^k(t)$) and the corresponding $D^k(t)$ are obtained from the knowledge stored in the action layer of the REM when the $(t+1)^{\text{th}}$ action is the corresponding learning action (i.e., a robot action of type VI). The normalized distance error ($\hat{e}$) can be obtained from (6.7).

Table 6.1: Rule Base of the Fuzzy Inference System for Feedback Evaluation

Input Memberships		Previous output		
		S	M	H
Feedback term	PE	PS	P	PL
	G	NC	NC	NC
	NE	NS	N	NL

Table 6.2: Mapping of user feedback terms

User feedback		Mapped feedback term
For a type I or II robot action	For a type III robot action	
Too little	Too close	PE
Good	Good	G
Too much	Too far	NE

$$D_F = \begin{cases} D^k(t) - d_0 \mid k = 1 & \text{when the } t^{\text{th}} \text{ action is} \\ & \quad \text{a type I or II robot action} \\ D^k(t) \mid k = 2 & \text{when the } t^{\text{th}} \text{ action is} \\ & \quad \text{a type III robot action} \end{cases} \tag{6.6}$$

$$\hat{e} = \frac{e}{D_F} \tag{6.7}$$

6.3 Results and Discussion

6.3.1 Experimental setup

The proposed concept has been implemented on the MIRob platform, and a user study has been conducted in an artificially created domestic environment inside the research facility to validate the performance gain of the proposed method over the existing systems. The experimental environment consisted of three rooms, namely “Kitchen”, “Corridor” and “Office”. These three rooms differed in their characteristics, such as room size, free space and object arrangement. The room names and the objects present during the experiment are annotated on the map shown in Fig. 6.6.

To evaluate the performance of the system, a parameter called the "satisfactory level" [76] is used; the definition of the "satisfactory level" (SL_{N_A}) is given in (6.8), where N_{FG} is the number of feedback instances of "good" type received following the execution of N_A previous movement-related user instructions. It should be noted that if feedback is not given for a particular action, it is assumed to be "good".

$$SL_{N_A} = \frac{N_{FG}}{N_A} \tag{6.8}$$

During the experiment, the parameters related to learning were chosen to be $\eta^k = 0.1$, $\delta_a^k = 5$ and $\delta_b^k = 3$ for $k = 1$ and 2. The definitions of the lower and upper bounds on the centers $((a_i^k)_L,\ (a_i^k)_H)$ and widths $((b_i^k)_L,\ (b_i^k)_H)$ of the output membership functions are given in (6.9), (6.10), (6.11) and (6.12), respectively.

$$(a_i^k)_L = \begin{cases} 0 & \text{if } i = 1 \\ a_{i-1}^k(0) & \text{otherwise} \end{cases} \tag{6.9}$$

$$(a_i^k)_H = \begin{cases} a_{i+1}^k(0) & \text{if } i = 1, 2, 3, 4 \\ 1.0 & \text{otherwise} \end{cases} \tag{6.10}$$

$$(b_i^k)_L = \frac{b_i^k(0)}{2} \tag{6.11}$$

$$(b_i^k)_H = \frac{3b_i^k(0)}{2} \tag{6.12}$$

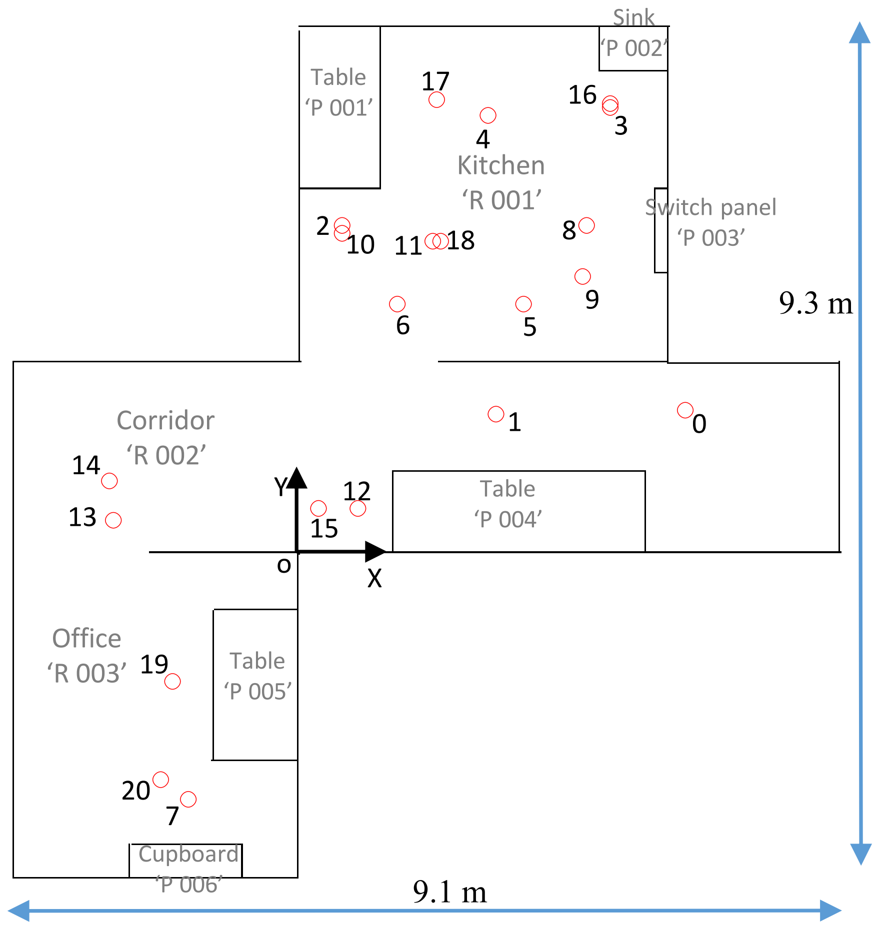


Figure 6.6: The positions of the robot after executing the user instructions listed in Table 6.3 are marked on the map with the corresponding command numbers. The map is drawn to scale. However, it should be noted that the markers are not drawn to scale and do not reflect the actual size of the robot.

6.3.2 Experiment and results

Since user studies are highly subjective in nature, the user study performed to validate the performance of the proposed method was designed and conducted with due attention to the recommendations given in [85] for designing, planning and executing human studies of human-robot interaction.

The user study was conducted with 24 participants (male = 15, female = 9) between 22 and 67 years of age (mean = 34.6, SD = 15.4). At the start of the experiment, the subjects were instructed on the possible structures of the navigation commands that could be understood by the robot. Subsequently, they were asked to freely navigate the robot such that the robot's movements would cover the entire environment. This decision was made since asking users to navigate a robot using a predefined set of commands or along a predefined path is highly restrictive for users, and consequently, the resulting behavior may not reflect actual user desires. Furthermore, this approach ensures that the intentions of the users are solely related to navigation, without any intent to perform any other task (e.g., placing/picking up an object on a table), which may influence the characteristics of the desired movement. The initial position of the robot was not fixed; instead, the initial position was selected randomly for each run. The users were also asked to follow the robot such that they could visually observe the movements of the robot and the environment. A few snapshots taken during the experiment are shown in Fig. 6.7. The users were advised to issue voice feedback about the movements of the robot (considering only the quantitative distances corresponding to the uncertain terms in the user commands) when it was necessary. To increase the voice recognition accuracy, a wireless headset with a microphone was provided to each user for issuing voice commands.

Humans have great adaptive capabilities, and during experiments, users may adapt to the behavior of robots. Therefore, to rectify the bias due to this adaptation, the participants were divided into two groups, each comprising 12 participants. In the first part of the experiment, the concept proposed in this chapter was implemented in the robot. Each user in the first group was taken individually to conduct the experiment. Subsequently, the learning ability of the system was disabled, and the abilities of the system were modified to be similar to those of the system explained in chapter 4. Then, the users in the second group were taken in for the study. Afterward, the users in the first group were taken in to conduct the study again using the system with no learning ability, since there was a considerable time gap that should have allowed the adaptation of the users toward the robot to fade. Finally, the users in the second group were taken in to conduct the study again using the system with the learning ability.

Figure 6.7: Snapshots of MIRob taken during the experiment are shown here.

The number of interactions with the robot (navigation commands only) was limited to 50 per user. This value was chosen because the satisfactory level reaches saturation before that point. The robot's movements and the parameters related to the UIUM for the first 20 commands issued by a randomly chosen user when interacting with the system with the learning ability are given in Table 6.3. The corresponding positions of the robot after executing each user command are annotated with the corresponding command numbers on the map shown in Fig. 6.6.

Table 6.3: Example results for the system with the learning ability

	User command	k	**AM or OS ($\mathbf{m^2}$)**	**Room size ($\mathbf{m^2}$)**	**Free space ($\mathbf{m^2}$)**	D^k **(cm)**	Y^k **(cm)**	**Destination position** X	**Destination position** Y	**Destination position** θ	**Feedback**	e **(cm)**	$\mathbf{SL_{10}}$
1	Move a little forward	1	little	18.85	16.33	500	204	217	146	-178	too much	-143	-
2	Move near to the table in the kitchen	2	1.62	15.08	12.95	65	46	46	355	90	-	-	-
3	Move near to the sink	2	0.379	15.08	12.95	65	44	341	486	90	too far	-22	-
4	Move a medium distance toward the table	1	medium	15.08	12.95	250	145	205	479	-179	-	-	-
5	Move far to the left	1	far	15.08	12.95	294	223	245	268	-78	-	-	-
6	Move right	1	medium	15.08	12.95	236	136	106	270	-175	too much	-70	-
7	Move near to the cupboard	2	0.652	11.5	9.27	47	36	-122	-279	-90	too close	17	-
8	Move near to the switch panel	2	0.138	15.08	12.95	165	79	315	357	0	too far	-53	-
9	Move a medium distance to the right	1	medium	15.08	12.95	163	73	310	299	-93	-	-	-
10	Move near to the table	2	1.62	15.08	12.95	97	56	46	345	90	-	-	0.5
11	Move a little right	1	little	15.08	12.95	347	95	146	337	-1	too much	-83	0.5
12	Move near to the table in the corridor	2	2.52	18.85	16.33	58	42	63	45	0	-	-	0.5
13	Move far backward	1	far	18.85	16.33	370	274	-207	32	-177	too little	124	0.5
14	Move a little right	1	little	18.85	16.33	175	42	-212	76	92	-	-	0.5
15	Move near to the table	2	2.52	18.85	16.33	160	87	18	45	0	too far	-53	0.4
16	Move near to the sink	2	0.379	15.08	12.95	65	41	341	489	90	too far	-22	0.4
17	Move near to the table	2	1.62	15.08	12.95	120	60	150	494	180	-	-	0.5
18	Move a medium distance to the left	1	medium	15.08	12.95	284	152	156	340	-89	-	-	0.6
19	Move near to the office table	2	1.56	11.5	9.27	85	48	-142	-148	0	too far	-28	0.5
20	Move far to the right	1	far	11.5	9.27	158	103	-153	-256	-97	-	-	0.5

Table 6.4: Variations in the parameters of the output membership functions with the user instructions given in Table 6.3

Command number	k (submodule)	a_1^k	a_2^k	a_3^k	a_4^k	a_5^k	b_1^k	b_2^k	b_3^k	b_4^k	b_5^k
Initial	1 and 2	0.0833	0.2500	0.5000	0.7500	0.9167	0.2500	0.5000	0.5000	0.5000	0.2500
1	1	0.0833	0.2093	0.3885	0.7500	0.9167	0.2500	0.4756	0.4331	0.5000	0.2500
3	2	0.0833	0.2017	0.3773	0.7414	0.9167	0.2500	0.4710	0.4264	0.4948	0.2500
6	1	0.0833	0.2093	0.3403	0.6277	0.9167	0.2500	0.4756	0.4042	0.4266	0.2500
7	2	0.0833	0.2702	0.4854	0.7614	0.9167	0.2500	0.5121	0.4912	0.5069	0.2500
8	2	0.0833	0.2247	0.3698	0.7585	0.9167	0.2500	0.4848	0.4219	0.5051	0.2500
11	1	0.0833	0.1722	0.3403	0.6277	0.9167	0.2500	0.4533	0.3476	0.4266	0.2500
13	1	0.0833	0.1722	0.3403	0.6763	0.9167	0.2500	0.4533	0.3476	0.4558	0.3299
15	2	0.0833	0.1801	0.3698	0.7139	0.9167	0.2500	0.4580	0.3486	0.4783	0.2500
16	2	0.0833	0.1324	0.3698	0.7054	0.9167	0.2500	0.4294	0.2759	0.4732	0.2500
19	2	0.0833	0.1324	0.2674	0.6600	0.9167	0.2500	0.3905	0.2759	0.4460	0.2500

In this run, the robot was initially placed at location '0' on the map ($X = 421$, $Y = 154$, $\theta = -178$). Then, the robot was commanded to "move a little forward" by the user. This is a motional command, and the quantitative distance value for the uncertain term "little" was interpreted to be 204 cm by submodule 1 of the UIUM. Therefore, the robot moved to location '1' by performing a type I robot action to fulfill the user command. However, the distance moved (i.e., the interpreted quantitative value for the term "little") was larger than the distance expected by the user, and therefore, the user gave the feedback "too much" to the system. Therefore, the robot performed a type VI robot action to learn from this feedback, and the FEM evaluated a quantitative error value for the feedback term (i.e., e). As a result of this user critique, the parameters of the output membership functions of submodule 1 of the UIUM were modified to the values given in the 2$^{\text{nd}}$ row (command no. 1) of Table V. Then, as the 2$^{\text{nd}}$ user command, the robot was commanded to "move near to the table in the kitchen". This is a positional command, and the quantitative distance value for the uncertain term "near" was interpreted to be 46 cm by submodule 2 of the UIUM. Therefore, the robot moved to location '2', at the corresponding distance from the table in the kitchen ('P 001'), by performing a type IV action (to move to the kitchen, 'R 001') followed by a type III action. In this case, the distance interpreted by the robot was accepted by the user, and therefore, no feedback was given to modify the robot's perception. Then, the robot was commanded to "move near to the sink". In this case, the distance assigned to the term "near" by the robot was 44 cm, and the robot moved to location '3'. The position reached by the robot was deemed to be "too far" from the sink ('P 002') according to the user's expectation. Therefore, the parameters of the output membership functions of submodule 2 of the UIUM were modified to the values given in the 3$^{\text{rd}}$ row (command no. 3) of Table V by means of a type VI robot action. Similarly, a total of 50 navigation commands were issued by the user, and the variations in the parameters of the UIUM corresponding to the commands listed in Table IV are given in Table V. The observed modification of the parameters of the output membership functions of the UIUM confirms that the system is capable of modifying its perception of uncertain information based on user feedback. The satisfactory level (SL) was calculated based on the 10 previous states, and the variation in the SL_{10} value is given for the 10^{th} user command onward in Table IV. Another similar experimental run was performed by the same user after the learning ability of the system had been disabled (i.e., the system was similar to that described in chapter 4). In this case, the parameters of the output membership functions of the UIUM were not modified in response to the feedback from the user and instead remained constant at their initial values.

Similar experimental runs were conducted using the system with the learning ability (i.e., the system proposed in this chapter) and the system with no learning ability (i.e., similar to the system described in chapter 4) by all 24 participants. The variations in the SL_{10} values with the number of executed commands for both systems were calculated for all users. The variations in the mean SL_{10} values with error bars are shown in Fig. 6.8. The error bars represent the 95% confidence intervals (CIs) with respect to the mean values. The variations in SL_{10} for all users are also shown as box plots in Fig. 6.9 for better visualization of the results.

In the initial stage (after execution of the 10^{th} user command), both systems exhibit rather low mean satisfactory levels (0.5542 for the system with the learning ability and 0.5125 for the system with no learning ability), and up to the 22^{nd} user command, the difference between the two means is not statistically significant ($P \geq 0.05$) according to the t-test. Furthermore, the differences between the means for the two systems are very small, and in some situations, they overlap. The variations of the medians also exhibit similar characteristics. Therefore, it can be concluded that there was no initial prejudice in the users' evaluations of the two systems due to their adaptation toward the system in earlier runs.

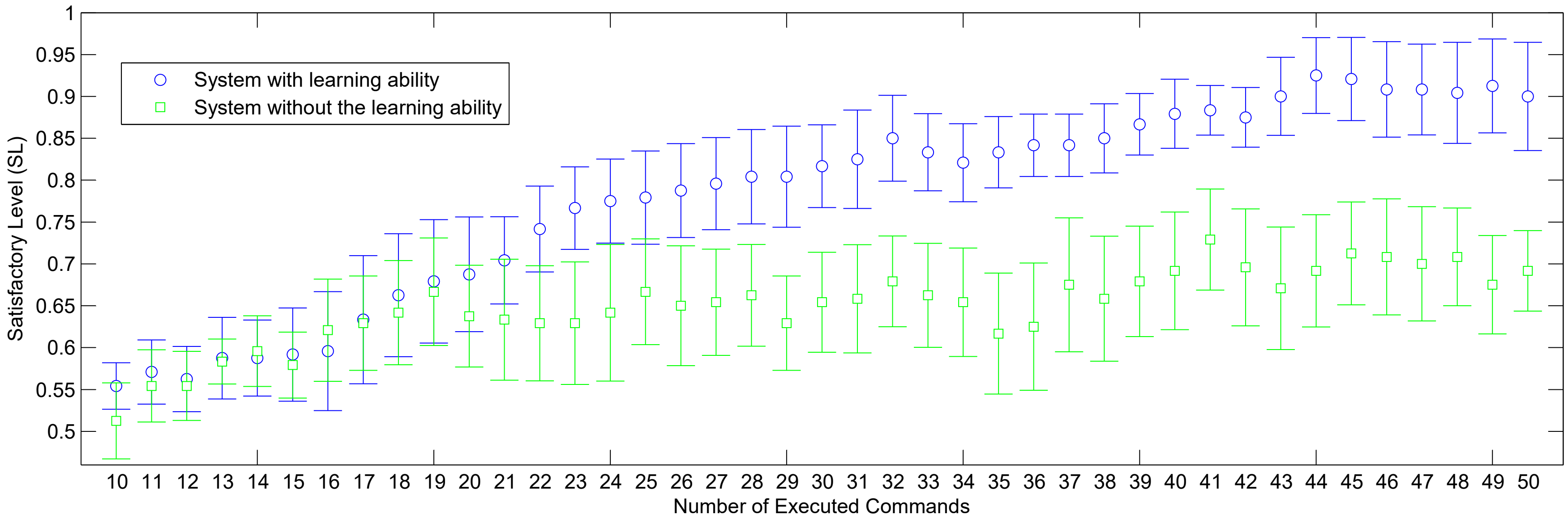


Figure 6.8: This plot shows the variations in the SL_{10} values of the system with the learning ability and the system with no learning ability. The markers represent the mean values, and the error bars represent the 95% confidence intervals (CIs) for the means based on a t-distribution.

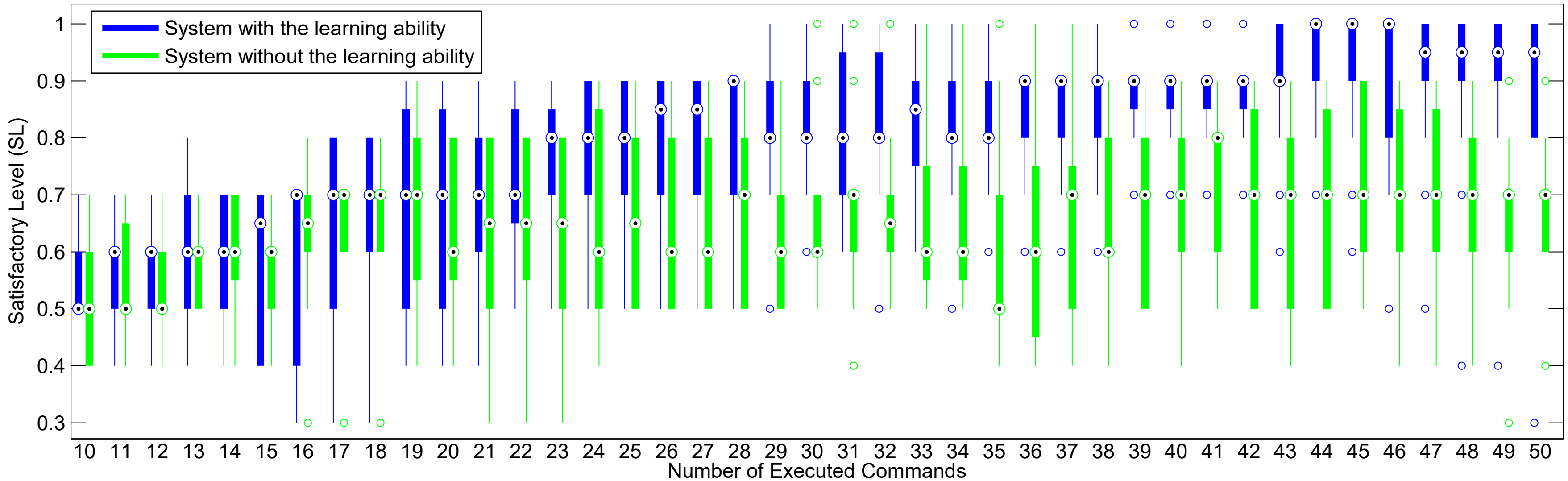


Figure 6.9: This figure shows box plots of the variations of the SL_{10} values with the number of executed commands for all users. The results for the system with the learning ability are shown in blue, and the results for the system with no learning ability are shown in green. The black dots in white circles represent the medians, and the boxes represent the interquartile ranges. The whiskers represent the maximum and minimum values of each distribution. However, the maximum length of the whiskers is limited to 2.7σ; any outliers are marked with circles in the color of the corresponding data set.

The satisfactory level of the system proposed in this chapter (i.e., the system with the learning ability) increased gradually over time, and finally, the mean of SL_{10} settled at approximately 0.9 (after the 46th command, $SL_{10} = 0.9083$; after the 47th, $SL_{10} = 0.9083$; after the 48th, $SL_{10} = 0.9042$; after the 49th, $SL_{10} = 0.9125$; and after the 50th, $SL_{10} = 0.9000$). Therefore, the learning ability of the system facilitates the adaptation of its perception of uncertain information to match the perception of the user. Hence, in this experiment, user satisfaction increased with successive interactions. Meanwhile, the satisfactory level of the system with no learning ability was also increased at the end of the experiment compared with the initial stage (initially, after the 10th user command, $SL_{10} = 0.5125$; after the 50th command, $SL_{10} = 0.6917$). This occurred because humans have a great cognitive ability to adapt their perceptions in accordance with the actions of their peers. Hence, the users adapted to the perception of the robot during the experimental runs. Therefore, user satisfaction increased with successive interactions even though the system did not adapt to their perceptions. However, the SL of the system with no learning ability was lower than that of the system with the learning ability; from the 22nd command onward, the differences between the means are statistically significant at the 95% confidence level ($P < 0.05$). When the power values of the statistical analysis are considered, from the 22nd command onward, the power values are also greater than 0.8 (according to Cohen's four-to-one weighting of the beta-to-alpha risk criterion [88], power values greater than or equal to 0.8 can be considered as good). Therefore, it can be concluded that the experimental results correctly indicate rejection of the null hypothesis (i.e., H_0: the mean SL values of the two systems are the same) when the alternative hypothesis (i.e., H_1: the mean SL value of the system with the learning ability is greater than that of the system with no learning ability) is true (from the 22nd command onward). Furthermore, from the 26th command onward, Cohen's d values of greater than 0.8 can be observed. This implies that there is a large effect (values above 0.8 are considered to be large [88]). Therefore, it can be concluded that there is a definite, noticeable effect on the SL due to the addition of the learning ability. Based on these statistical observations regarding user satisfaction, it can be concluded that the experimental results confirm that the performance enhancement of the system with the learning ability (i.e., the system proposed in this chapter) over the system with no learning ability is significant and reliable. Ultimately, with regard to user satisfaction, the system with the learning ability (i.e., the system proposed in this chapter) surpasses the system with no learning ability (i.e., similar to the system explained in chapter 4).

The medians of the SL scores also exhibit a phenomenon similar to that of the means of SL, as seen from the box plots shown in Fig. 6.9. According to these box plots, there are both positive and negative outliers for both systems. The existence of outliers for a user study of this kind is natural, since there may be users whose expectations and perceptions are significantly different from those of others. Except for a single user, the individual variations in the SL outliers are similar to the variations of the majority of the data, although the absolute SL scores are above or below the others. Furthermore, the variations for the older users were separately analyzed and were found to exhibit characteristics similar to those of the overall results. Therefore, even though not all participants were older or challenged users, this aspect of the study population showed no significant effect on the evaluation of the performance of the system. Furthermore, assistive robots can be used indirectly for assisting elderly/disabled people by using them as support agents for human caregivers in care facilities such as nursing homes [89], and hence, not all users of such systems will be older people.

The proposed method enables a robot to learn from user feedback while concurrently adapting its perception of uncertain information according to the spatial information of its current working environmental. Moreover, the proposed system is capable of modifying the parameters of the output membership functions of the system proposed in chapter 4. Therefore, the key characteristics of the scheme for perception adaptation based on environmental factors that is presented in chapter 4 are clearly well preserved in the method proposed in this chapter. By contrast, the methods proposed in [76,90] are capable of perception learning based only on user feedback, and after the learning process is complete, the meanings of uncertain terms are fixed. Systems that assign such fixed meanings to uncertain information are suitable for use only in a fixed working environment and cannot be used in dynamic working environments. The experiment performed in this study was conducted in an environment that was static with respect to the global frame. However, this environment was dynamic with respect to the robot's frame since the environmental parameters perceived by the robot varied with its current position in the globally fixed environment. Therefore, the environment perceived by the system was not a static one. Moreover, the working environment was also dynamic due to the changing position of the robot during navigation. Therefore, the proposed system is capable of adapting the perception of a robot toward that of its user based on user feedback while concurrently adapting its perception in accordance with its current environment. This is the major improvement of the proposed method over existing methods.

6.4 Summary

This chapter proposed a method of effectively interpreting uncertain information related to navigation commands by adapting a robot's perception of uncertain information based on both the environment and user critiques. The main improvement of the proposed concept over existing systems is that the proposed system is capable of concurrently adapting its perception based on the spatial information of the environment while learning from user feedback.

The DisI has been reimplemented using fuzzy neural networks. These fuzzy neural networks enable the system to learn from user feedback while simultaneously adapting its perception based on sensory information related to the environment. The FEM has been deployed to evaluate the quantitative meanings of feedback terms.

Experiments were conducted to validate the performance enhancement achieved by the system due to its learning ability. An index called the user "satisfactory level" was used for the performance analysis. The experimental results confirm the performance improvement of the proposed method over the existing methods. The performance of the system with the learning ability surpasses that of a system with no learning ability. Therefore, the proposed concept is capable of enhancing user satisfaction by adapting a robot's perception of uncertain information based on the environment and user feedback.

CHAPTER 7

INTERPRETING UNCERTAINTIES RELATED TO RELATIVE REFERENCES

7.1 Relative References

As mentioned in chapter 3, the motional and positional navigational commands may consist with a reference (i.e., <Reference>). This reference is useful in identifying the direction of the movement (for motional command) or the position to be moved (for positional command only). Therefore, the robot must be capable of identifying the reference to fulfill a navigation task requested by the user. However, this reference may be expressed as a phrase by using a combination of other references. For instance, "table close to the vase" can be considered as an example phrase. In here, the reference of interest is a table, which is in close proximity to a vase (a second reference). Often uncertain information is included in these sort of referential phrases and the uncertain information is expressed with relative to any other aspect such as features of other objects in the environment, and position and orientation of the user or the performer. Therefore, the robot must be capable of effectively interpreting the uncertain information associated with these sort of relative references. This chapter proposes the Relative Uncertainty Interpreter (RUI) to embedded into the Uncertain Information Understanding Module (UIUM) for realizing the requirement. The arrangement of the Intelligent System for Understanding Uncertain Information (ISUUI) with the proposed RUI is shown in Fig. 7.1

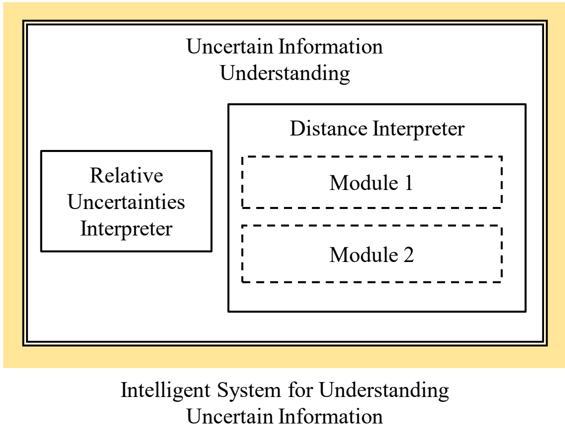


Figure 7.1: The ISUUI with the proposed RUI

7.1.1 Actions Related to Identification of Relative References

Voice commands that include relative references are often used in issuing navigation instructions to the companions. In order to analyze the commands, the navigational commands with relative references are categorized based on the number of referencing entities included in the command. As an example, the command "move near to the large table" can be considered. In here, "large table" is the reference and it is defined as Ref_1. This type of referential phrases is defined as Reference Phrase Type I (RP Type I). The robot needs to identify the reference by interpreting the uncertainties related to relative references if there are uncertainties. The robot Decisive Action type I (DA-I) is defined for the Ref_1 identification action. In some cases, there may be more than one references. As an example "move near to the table close to the large cupboard". In this command, the robot needs to move near to the table that is close to the large cupboard. "Table" is considered as the Ref_1 and "large cupboard" is considered as the second reference (Ref_2). In order to interpret this kind of commands, first the Ref_2 should be identified and then the Ref_1 should be identified based on the Ref_2. The Decisive Action type II (DA-II) is defined for the process of identification of the second reference (Ref_2). Therefore, in this kind of commands, the robot first needs to perform DA-II in order to identify the Ref_2 and then DA-I action in

Table 7.1: Example referential phrases and required sequence of decisive actions

	Referential Phrase	Available references	Type	Required decisive action[1]
1	the large table	Ref_1 = "large table"	RP-I	DA-I
2	the table close to large cupboard	Ref_2 = "large cupboard" Ref_1 = "table close to Ref_2"	RP-II	DA-II DA-I
3	the chair left of the table close to the sink	Ref_3 = "sink" Ref_2 = "table close to Ref_3" Ref_1 = "chair left of Ref_2"	RP-III	DA-III DA-II DA-I
4	Ref_1 ..... Ref_n	Ref_n . . . Ref_1	RP-n	DA-n . . . DA-I

[1] These decisive actions are named for the sake of clarity of explanation. In reality, these are merely processing tasks of the robot.

order to identify Ref_1. Similarly, if there is n number of references in a particular command then it is categorized as Reference Phrase Type n (RP type n). If there is n number of reference in a command then the robot needs to perform DA-n to DA-1 successively in order to finally identify the Ref_1 (DA-n for Ref_n then DA-$(n-1)$ for Ref_{n-1} ...... DA-II for Ref_2 finally DA-I for Ref_1 identification). Example referential phrases and the corresponding sequence of required decisive actions are summarized in Table 7.1

After identification of the Ref_1 in a command if there is a Ref_1, the robot needs to execute robot actions (as explained in chapter 3) to move to the specified position described by the uncertain information related to positional and motional user commands.

7.2 Usage of Uncertain Information in Relation to Relative References

In order to identify the natural tendencies of human related to usage of relative references, a study has been carried out with participation of 55 healthy people (Age: M = 34.38 and SD = 16.1 years). The participants were asked to write down 10 to 15 referential phrases that can be used in order to refer objects for navigating a person/robot inside the given environment. Furthermore, they were asked to explicitly mention the referred object along with the phrases. For this, the objects in the environment were labeled with distinct numbers. For the study 10 different domestic environment were used and 5-6 volunteers were taken for a single environment. Altogether, 620 valid relative referential phrases have been gathered. The gathered data have been analyzed in order to identify the natural tendencies of humans associated with relative references[1].

The first analysis has been conducted to investigate the inclusion of the uncertain information in referential phrases. The percentage availability and non-availability of the uncertain information in the phrases are shown in Fig. 7.2. Approximately 46% of the gathered phrases contains the uncertain information yielding to necessity of having a mechanism to interpret the uncertain information in referential phrases for enhancing the navigation command understanding of service robots. The number of references available in a phrase has also been analyzed and the outcomes are presented in the plot shown in Fig. 7.3. More than 90% of the phrases contain either a single reference or two references and there are only few cases where three references are available in a phrase. Moreover, RP Type I and RP Type II are the noteworthy types.

The composition of the uncertain information has been analyzed based on the entities related to the uncertain information such as "distance", "size" and "direction". The percentage composition of the uncertain information in relation to different entities is given in the plot shown in Fig. 7.4. Most of the uncertain information related to relative references can be classified into two categories; object saliency related uncertainties and position related uncertainties. This classification is illustrated in Fig. 7.5 with examples. There are uncertain information related to the features of objects and such information is categorized as uncertainties related to saliency of objects. Object size (e.g., "large table"), shape (e.g., "round table") and color (e.g., "dark chair") related uncertain terms could be considered as examples. Position related uncertain information mainly

[1] It should be noted that all the analyses have been conducted manually based on human knowledge.

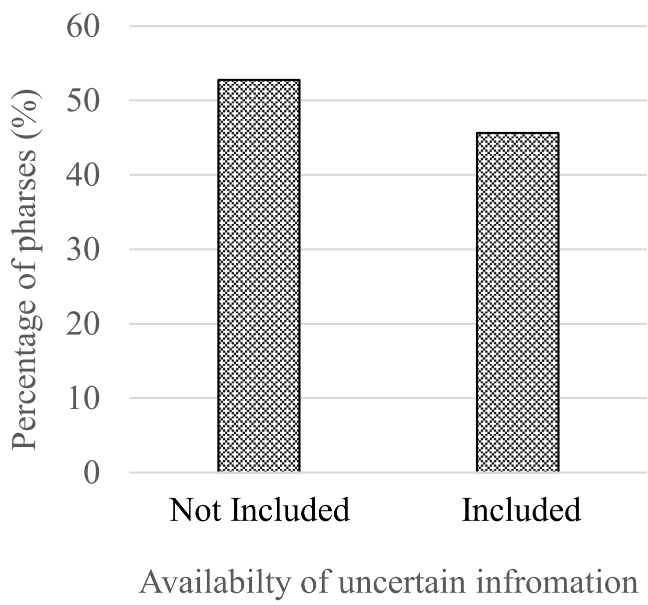


Figure 7.2: Percentage frequency of phrases vs. availability of uncertain information

includes distance (e.g., “table near the door”), direction (e.g., “left of the table”) and location (e.g., “corner of the kitchen”) related uncertainties. Furthermore, there are few situations where the uncertain information is related to the age of the reference objects (e.g., “old cupboard”) and those are categorized into the

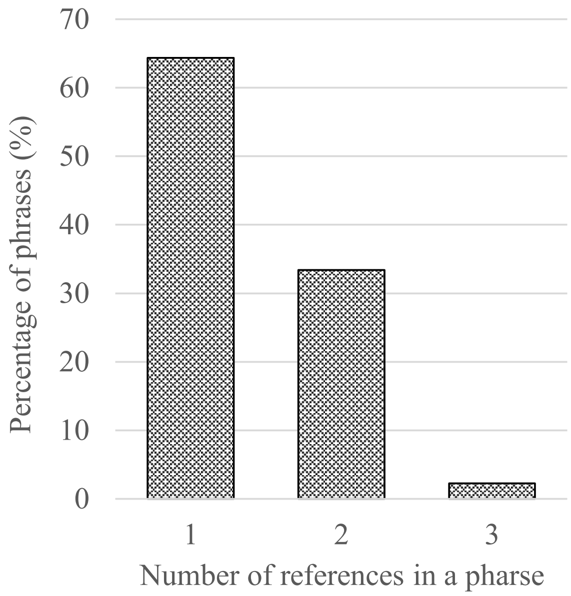


Figure 7.3: Percentage frequency of phrases vs. number of reference in a phrase

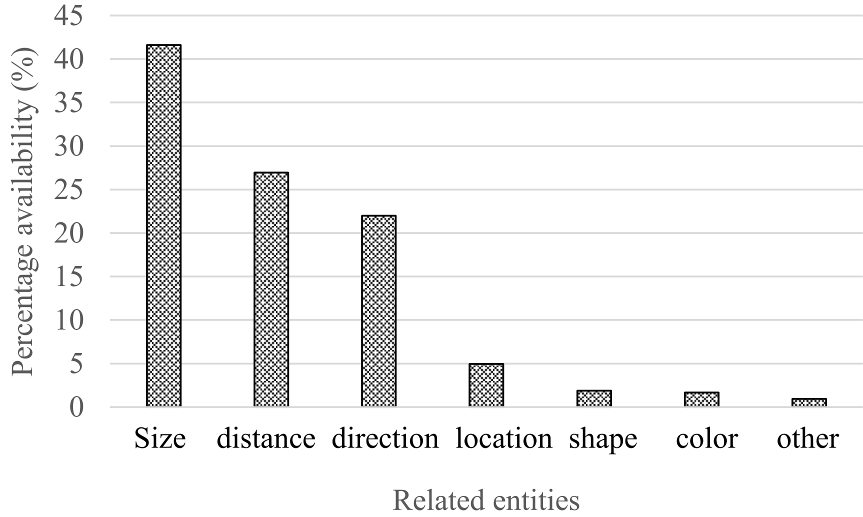


Figure 7.4: Percentage frequency of uncertain terms vs. entity of uncertain information

other category in the plot.

According to the outcomes of this survey, size, distance and direction are the mostly associated relative referencing entities in navigation (approximately 90%). Therefore, only the relative referencing uncertain information related to relative size, distance and direction is considered for this paper. Relative uncertainties may associate with any of the references in a command. Therefore, the relative uncertainties associated with them need to be interpreted by the robot in order to identify them. i^{th} and $(i+1)^{\text{th}}$ references in a command are considered for the explanation of the identification of the i^{th} reference when there is n number

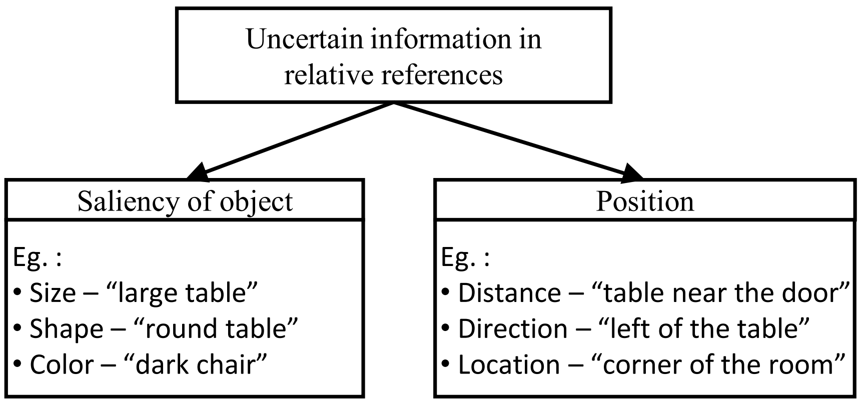


Figure 7.5: Classification of relative references

of references. Based on the outcomes of the study, it has been assumed that the linkage between the i^{th} and $(i+1)^{\text{th}}$ references (i.e Ref_i and Ref_{i+1}) can be expressed in following format given in JSpeech Grammar Format [86].

$$<Ref_i>=[<R_{size}>]<\overline{Ref_i}>[(<R_{dis}>|<R_{dir}>)<Ref_{i+1}>]$$

It should be noted that the redundant words in the natural language referential phrases are omitted in the given structure and only the key information is represented. In here, R_{size}, R_{dis} and R_{dir} represent language descriptors related to relative size, distance and direction respectively. The lexical symbol that represent the object type of the Ref_i is given by $\overline{Ref_i}$. Relative size descriptors may be included in i^{th} reference and the descriptor is included before the object type lexical symbol of the reference (i.e. $\overline{Ref_i}$). If there is a Ref_{i+1}, then the Ref_i is linked with the Ref_{i+1} through a relative distance (R_{dis}) or a relative direction (R_{dir}) descriptor that is included after $\overline{Ref_i}$. Therefore, Ref_{i+1} should be evaluated first and then the Ref_i should be evaluated based on the selected Ref_{i+1}. Moreover, the evaluation should be carried out from the n^{th} reference to the 1^{st} reference ($i = n, n-1, ..., 2, 1$ order) as explained earlier.

7.3 Interpretation of Relative References

Three separate algorithms have been implemented inside the RUI based on the identified natural tendencies of humans in interpreting such uncertain information. The required algorithm is selected based on the entity of the uncertainty. Relevant entity of an uncertainty is identified based on the knowledge of the language memory.

7.3.1 Relative sizes

Size of an object is a relative entity based on the saliency of the considered object type. As an example, "large table" means that the object is a table and the size of it is relatively larger than the other tables in the environment. The object size related uncertain information can be associated with any i^{th} reference. Therefore, the robot needs to interpret object size related uncertain information in order to identify Ref_i if relative size related uncertain information is associated with it. The algorithm 1 has been developed in order to interpret the object size related uncertain information by considering the following natural human tendencies.

- Object size related uncertain information such as "large" and "small" is expressed relative to the similar type of other objects in the environment. Most of the times there are few objects of the same type. e.g., two tables in kitchen, three chairs in office.
- Predominantly two categories are used in order to distinguish the object of interest from others based on the size. "large" and "small" are the two most commonly used categories. In order to categorize into more, humans prefer to use other relative references related entities such as "position" and "directions" instead of size in such cases.
- "Large object" means the largest object among the similar types of objects in a given room or the entire house. "Small object" means the vice versa of this.
- When room related information is not included in the command, the selection is done based on the objects in the current room if more than one candidates are available in the current room. If there are no or only one candidate then the entire house is considered for the selection. e.g., A situation where the current room is kitchen and there are two tables and a

rack in it, is considered. In this situation, "large table" means one of the table inside the kitchen. However, "large rack" may not be the rack in the kitchen and this is evaluated considering the racks in the entire house.

In order to evaluate the algorithm, the required information of the objects such as sizes of objects and relationships between objects and rooms is retrieved from the environment layer of the Robot Experience Model(REM). The exceptions of the algorithm is handled by the Interaction Management Module (IMM) by changing the state of the robot to clarification state that can be used to get clarifications from the user.

Algorithm 1 Ref_i selection based on relative size

```
INPUT: Uncertain_category, object_type
OUTPUT: Selected_Ref_i
if room_info_available_in_command then
    if no_of_object(object_type) in the room ≥ 1 then
        Search only in the room
        return SELECTION(objects_in_given_object_type)
    else
        return Clarification_state
    end if
else
    if no_of_object(object_type)in Room_C > 1 then
        Search only in Room_C
        return SELECTION(objects_in_given_object_type)
    else
        Search entire house
        return SELECTION(objects_in_given_object_type)
    end if
end if

function SELECTION(possible_candidate)
    if Uncertain_category = "large"  then
        return largest_object_among_possible_candidate
    else if Uncertain_category = "small"  then
        return smallest_object_among_possible_Candidate
    else
        return Clarification_state
    end if
end function
```

7.3.2 Relative Distances

The uncertain information related to the relative distances is associated with Ref_i only when there is a post reference to it (i.e. if Ref_{i+1} exists). Hence, this information is useful in identifying the Ref_i by evaluating the relative distances between candidates for the Ref_i and the Ref_{i+1}. However, every time the Ref_{i+1} may not be an object and it may be any other entity such as performer. Algorithm 2 is used in order to identify the Ref_i by evaluating the uncertain information associated with the relative distance of that from the Ref_{i+1}. The algorithm has been developed in such a way that it can replicate the following natural human tendencies.

- The distance information is expressed relative to the distances between the possible candidates for Ref_i and the Ref_{i+1}.
- Possible candidates for the Ref_i are only in the room of Ref_{i+1}.
- Most of the times only two categories are used to distinguish the object of interest from other possible candidates. If there is a requirement of categorizing more than two, the humans tend to use others aspects.
- "Close" and "far" are the predominantly used uncertain categories and "close" means the closest one and far means the furthest one from the Ref_{i+1}. Synonyms are also used other than these two tokens. e.g., "near" for "close".

The information required for evaluation of the algorithm such as distance between the objects are retrieved from the environment layer of the REM. Language memory is used to map the possible uncertain terms associated with the uncertain term categories given in the algorithm.

Algorithm 2 Ref_i selection based on relative distance with Ref_{i+1}

INPUT: Uncertain_term_category, object_type, Ref_{i+1}
OUTPUT: Selected_Ref_i

```
Search only in room of Ref_{i+1}
if no_of_object(object_type) ≥ 1 then
    if Uncertain_term_category = "close" then
        return Ref_i_candidate_shortest_distance_to_Ref_{i+1}
    else if Uncertain_term_category = "far" then
        return Ref_i_candidate_longest_distance_to_Ref_{i+1}
    else
        return Clarification_state
    end if
else
    return Clarification_state
end if
```

7.3.3 Relative Directions

In order to distinguish the object of interest from the possible candidates for Ref_i, the relative direction with the Ref_{i+1} can be used. Therefore, navigational commands often include relative direction related uncertain information. The following natural human tendencies have been considered in order to develop the method to select Ref_i by interpreting the uncertain information related to relative directions.

- There are basic four directions and these directions depend on object characteristic and the point of view. Those directions have overlapping boundaries.
- Objects that have a useful side most of the times have fixed direction with them. As example objects; TV, Cupboard and Fan have a fixed front side based on their usage. Other directions; back, left and right are also fixed based on the front. This has the highest priority among other parameters that decide the orientation.
- The directions depend on the point of view if the Ref_{i+1} is not an object with fixed directions. The direct pointing direction is the front side and

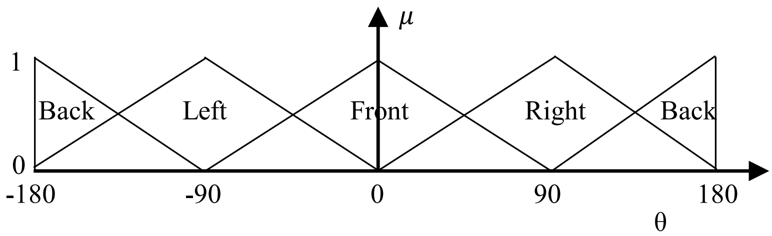


Figure 7.6: Membership function for the evaluation of relative directions. The absolute front is defined at the origin. The input, θ is in degrees and anticlockwise measurements are considered as positive.

other directions are given based on that. However, there are situations where the point of views of the action performer and the commander are different. According to [91], most of the times the point of view of the performer is dominant and performer tends to consider the point of view of himself in such scenarios. However, there is small probability of clarifying it from the commander or considering the point of view of the commander by the performer.

- If there is no Ref_{i+1} in the command, most of the cases the Ref_{i+1} will be the performer when relative directional information is available.

The fuzzy membership function shown in Fig. 7.6 is used to evaluate the likelihood of possible candidates for Ref_i to be the correct Ref_i. The object that has the highest activation degree in the considered direction is chosen as the Ref_i among the possible candidates. The angle between the absolute front and the vector $\vec{Ref_{i+1}Ref_i}$, θ is given as the input of the membership function. The absolute front is decided based on the context of objects or the point of view. If the Ref_{i+1} is an object with fixed directions then the absolute front is defined along the exact front of that object. Context layer of the REM is used in order to retrieve the knowledge of such objects. In other cases, the absolute front is defined along the point of view of the robot since the point of view of the performer is dominant over the point of view of the commander.

7.4 Results and Discussion

7.4.1 Experimental Setup

The proposed concept has been implemented on MIRob platform. At the start, the robot was initialized with a updated REM and a navigation map of the environment. Therefore, the robot is well aware of the arrangement and the characteristics of the environment during the experiments. The experiments have been conducted out in a simulated domestic environment inside the research facility in order to validate the applicability of the proposed concept. A human study has also been carried out to find out how the human participants execute the same set of tasks. The snapshots taken during the experiments are shown in the Fig. 7.7. Finally the experimental results of the developed robotic system and the human study have been compared in order to validate the performance of the proposed system.

(a) (b)

Figure 7.7: (a) shows MIRob is being instructed by a user during the experiment. (b) shows a snapshot taken during the user study. In here, 1 is issuing instructions to 2 during the command no 11 of the user study. Details are in Table 7.3.

7.4.2 Overall Behavior of the Proposed Concept

In order to analyze and verify the overall functionality of the proposed system, a sequence of voice instructions has been issued to the robot and the responses of the robot have been recorded as the experiment 1. An explanatory video of this segment of the experiment can be found in the supplementary multimedia attachment 4[2]. The issued user instructions and the corresponding responses of the robot are given in Table 7.2. The positions of the robot during the execution of the commands in the experiment 1 are annotated on the map shown in Fig. 7.8.

Instruction 1: Initially the robot was at position 'A' and the instruction 1, "move near to the small table in the kitchen" was issued to the robot. This command contains only one reference (i.e. only the Ref_1) and hence it is a RP Type-I phrase. In order to fulfill this command, first the robot had to identify the Ref_1 and then move near to that. The room related information is available in the command. Therefore, possible candidates for the Ref_1 should be inside the kitchen ('R 001'). There are two tables in the kitchen ('P 001' and 'P 003'), which were the candidates for the Ref_1, and the robot had to choose the correct table by interpreting the relative uncertain information associated with the Ref_1. In this case, the Ref_1 was associated with uncertain information related to the relative

Table 7.2: Issued commands and the responses of the robot

	Instruction	**Response of the robot**
1	"move near to the small table in the kitchen"	movement from A to B
2	"move near to the sink"	movement from B to C
3	"move far towards the table left of the sink"	movement from C to D
4	"move near to the cupboard close to the sink"	Voice response: "there is no cupboard close to the sink"
5	"move near to the table in the corridor"	voice response: "There are two tables in there. Please clarify"
6	"move near to the table close to the vase"	movement from D to E

[2]Available in the attached CD and www.youtube.com/watch?v=GIie6Zohelk

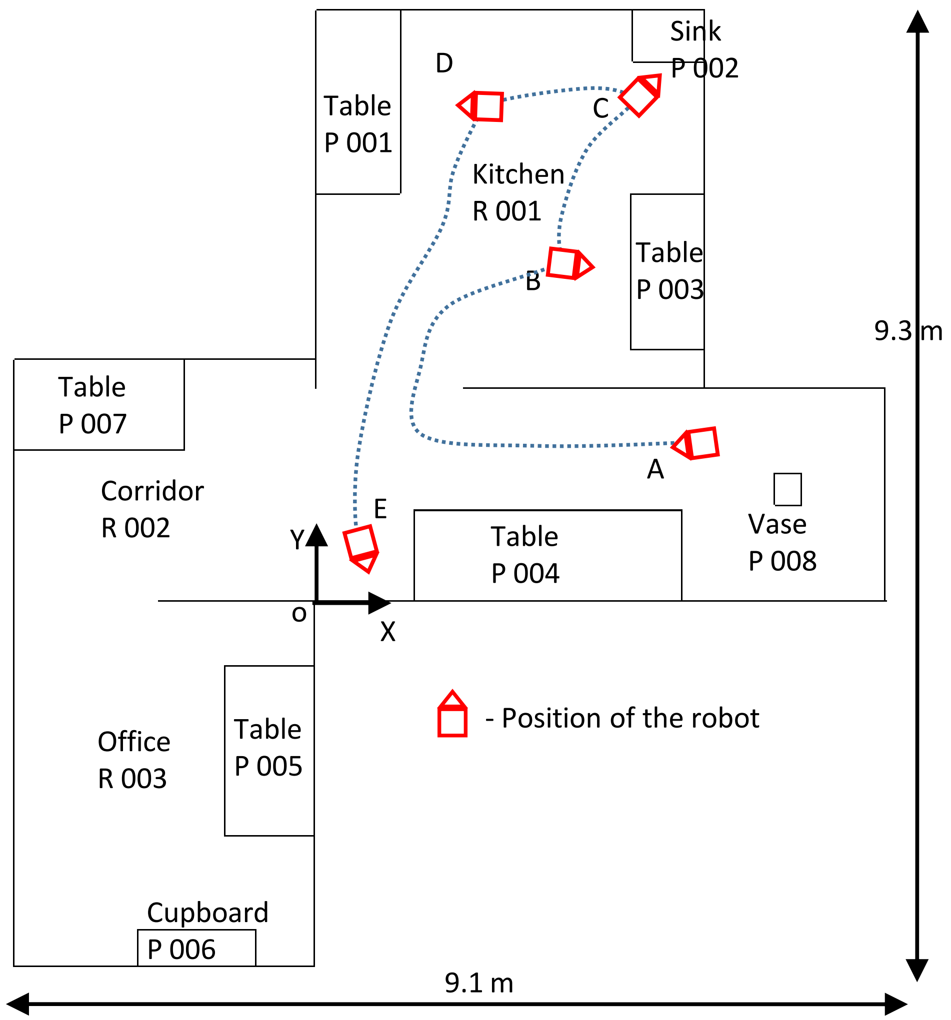


Figure 7.8: Positions of the robot recorded during the experiment 1. The map is drawn to a scale. However, the markers do not reflect the size of the robot.

size of the object ("small" is the uncertain term that distinguish the correct Ref_1 from other table). Based on the developed algorithm for interpreting the object size related relative referencing uncertain information (i.e. Algorithm 1), the robot selected 'P 003' as the Ref_1 by performing a DA-I action. The $Room_C$ (i.e. corridor) and the $Room_{Ref_1}$ (i.e. kitchen) were not the same. Therefore, first the robot had to perform an action of RA-IV to move to the kitchen and then it had to perform an action of RA-III in order to move near to the selected Ref_1 (i.e. 'P 003'). The quantitative value of the uncertain term "near" was interpreted by the sub-module 2 of the UIUM. The interpreted quantitative value for the term "near" was 65 cm and the robot moved to position 'B' where the distance between the robot and the ‘P 003’ is 65 cm.

Instruction 2: Then, the instruction 2, “move near to the sink” was issued. In this command, the Ref_1 can be identified directly since uncertain information is not associated with the object of interest. Then the robot moved to position ‘C’ by performing a RA-III action.

Instruction 3: Then the instruction 3, “move far towards the table left of the sink” was issued. Two references are included in the command (i.e, Ref_1 and Ref_2). Hence, first the robot performed a DA-II in order to identify the Ref_2. Uncertain information is not associated with the Ref_2 and hence the Ref_2 can be directly identified. After identifying the Ref_2, the relative direction related uncertain information associated with the Ref_1 was interpreted by performing a DA-I. The Ref_2 in this case was an object with a useful side and hence it has fixed directions. Therefore, the absolute front of the Ref_1 was considered along with the front of the sink. This information had been retrieved from the context layer of the REM. The chosen Ref_1 in this case was ‘P 001’. After the Ref_1 was decided, the robot moved to ‘D’ by performing a RA-II since the command was a motional command.

Instruction 4: Then the instruction 4 was issued to the robot. In here, the Ref_1 is a cupboard that is close to the sink. However, after interpreting the relative uncertainties in command there was no possible candidate for this according to the arrangement of the environment. Therefore, the robot responded with a voice response in order to notify it to the user by utilizing RA-V.

Instruction 5: In the next command, the robot was instructed to move near to the table in the corridor. According to the environment, there are two tables inside the corridor (‘P 004’ and ‘P 007’). Since no information is included in the command in order to distinguish the correct Ref_1 from possible candidates, the robot responded with a voice response that notifies it to the user.

Instruction 6: Finally, the robot was commanded to move near to the table close to the vase. First, the Ref_2 in the command was identified as 'P 008' and subsequently the Ref_1 was identified as 'P 004' by utilizing the Algorithm 2. Then the robot moved to position 'E' in order to fulfill the requirement of the instruction.

7.4.3 Comparison between the Results of the Human Study and the Proposed System

As the experiment 2, a human study has been conducted with 15 healthy human participants (age: M = 24 and SD = 3.2 years) in order to validate the performance of the proposed system. Before proceeding to the experiment, the participants were familiarized with the environment in order to build up the awareness about the experimental environment setting. During the human survey, the persons were asked to be on the positions marked on the floor of the environment. The persons were taken one by one to the study and each person was issued the set of commands given in Table 7.3 after the person settled on the given initial positions. The initial positions and the orientations of the human performers are marked on the map shown in Fig. 7.9 with corresponding command numbers. During the experiment, the person who issues the command was not in the sight of the performer except case number 11 and 13. In the case 11 and 13, the commander was directly in front of the performer. A snapshot taken when issuing the command number 11 to a performer by the commander is shown in Fig. 7.7 (b). The selected Ref_1 for a particular command was identified based on the action of the performer. Then, the same set of commands were issued to the robot after placing it on the corresponding locations used for the human survey and the selected Ref_1 in each case has been identified. The selected Ref_1 in each case is given in the Table 7.3.

Relative size interpretation: Relative size related uncertain information is associated in commands 1, 3, 5, 9 and 14. In command no 1, the performer has to select the correct table from the available two tables (P 001 and P 003) in the

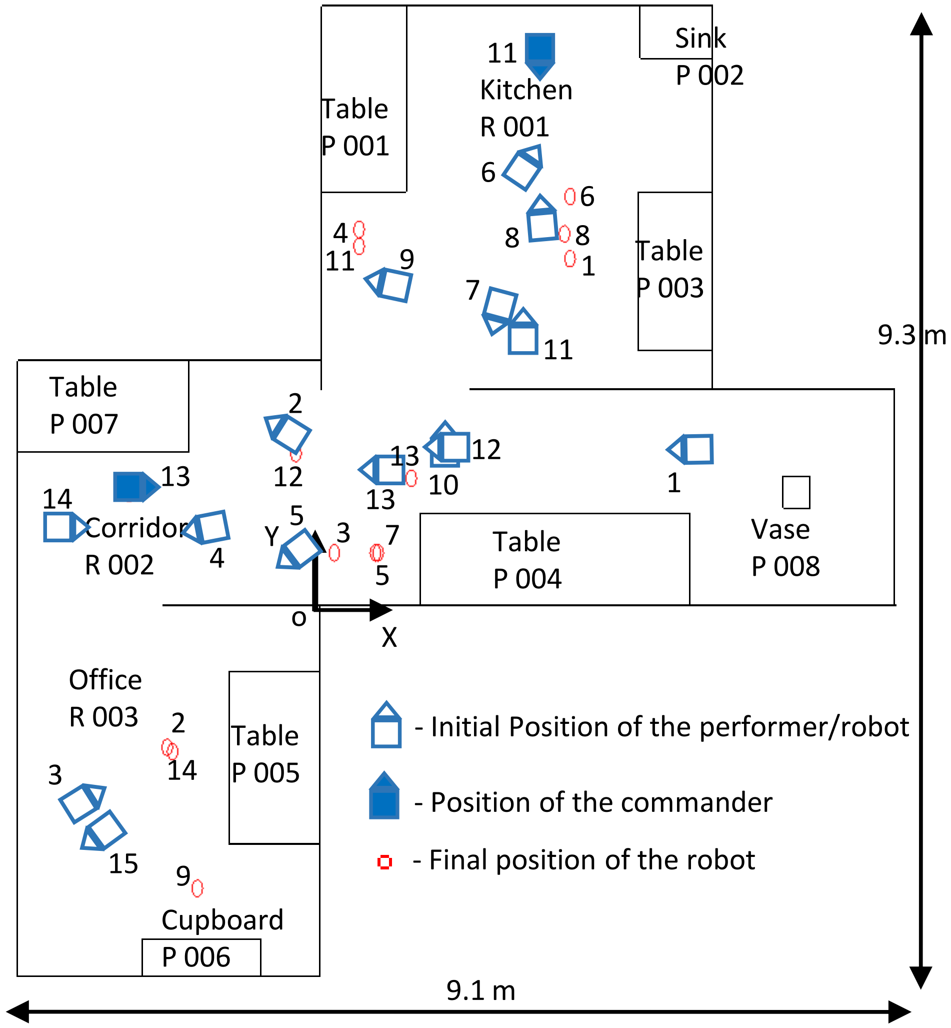


Figure 7.9: Initial and final positions of the robot and the initial positions of the performers during the experiment are marked on the map with the corresponding command numbers. It should be noted that the markers do not reflect the size of the robot or the human performers.

kitchen since the room information is given in the command. 'P 003' is relative smaller than the 'P 001' hence all the human performers and the robot have selected 'P 003' as the Ref_1 in the command. Command 14 is also similar to this. However, in that case, 'P 005' is the only candidate and all the human performers and the robot has selected it, as the Ref_1. Room information is not available in command 3, 5 and 9. In command 5, two candidates for the Ref_1 are available inside the current room of the performers (i.e. 'P 004' and 'P 007'). Hence, the performers have selected 'P 004', which is the relatively larger one. The robot has

Table 7.3: Comparison between the Ref_1 chosen by the human participants and the robot

command	Chosen Ref_1 [1] Humans	Robot
1 move near to the small table in the kitchen	P 003	P 003
2 move near to the table close to the cupboard	P 005	P 005
3 move near to the large table	P 004 - 73% P 005 - 27%	P 004
4 move near to the table left of the sink	P 001 - 93% P 003 - 7%	P 001
5 move near to the large table	P 004	P 004
6 move little towards the table in front of the sink	P 003	P 003
7 move near to the table close to the vase	P 004	P 004
8 move little towards the table on the right	P 003	P 003
9 move near to the large cupboard	P 006	P 006
10 move near to the cupboard close to the sink	N/A	N/A
11 move near to the table on the left	P 001	P 001
12 move little towards the table on the front	P 007	P 007
13 move little towards the table on the left	P 004 - 60% P 007 - 20% Rqst. - 20%	P 004
14 move near to the large table in the office	P 005	P 005
15 move near to the sink in corridor	N/A	N/A

[1] It should be noted that in the cases where the percentage values are not given the percentage is 100%. N/A means reply of non availability of an object to be selected as Ref_1. Rqst. means request of clarification.

also selected 'P 004'. In command 3, the performers were inside the 'R 003' when the command was issued. In this case, only one possible candidate was available inside the current room. 'P 004' has been selected by 73% of human performers and 'P 004' is not within 'R 003'. Therefore, this shows that when there is only one option for the section inside the current room the humans tends to search entire house instead of selecting it from the current room. 'P 005' has also been selected by 27% human performers. However, the proposed system has selected 'P 004' since the algorithm has been designed to follow only the dominant human behavior.

Relative distance interpretation: Relative distance related uncertain information is available in commands 2, 7 and 10. All the human performers and the robot have selected 'P 005' and 'P 004' in command 2 and 7 respectively since those are the closest candidates from the Ref_2 in each command. In command 10, all the human performers and the robot have identified that there is no cupboard close to the sink.

Relative direction interpretation: Relative direction related uncertain information is included in command 4, 6, 8, 11, 12 and 13. In command 4, the performers were asked to move near to the table left of the sink. ‘P 001’ has been selected as the Ref_1 by 93% of human performers. In here, the Ref_2 is a sink and it is an object with a useful side. The proposed concept has selected the same object by defining the orientation frame based on the contextual knowledge of the objects. One person has selected ‘P 003’ in this case and according to that performer, the selection was done by considering the point of view from the sink. However, the possibility of such selection is very low according to the obtained results of the human study. The command 6 also verifies that the useful side of the sink defines the orientation frame. In command 8 and 12, the human performers have decided the orientation frame with respect to them when there is no Ref_2 in the command. In command 11, the commander was in front of the human performers as shown in Fig. 7.7(b). However, the performers have decided the orientation frame based on own orientation in this situation too. The commander was in front of the performer in command 13 also. In this case, ‘P 004’ has been selected by 60% of the performers by considering own orientation frame, ‘P 007’ has been selected by 20% and 20% have requested clarifications from the commander. The orientation frame of the performer is dominant here too. In this case, ‘P 004’ is not within the sight of the human performers that will be the reason for requesting clarification or deciding the orientation frame based on the commanders point of view. Furthermore, information that reveals the intention of the commander may be conveyed through nonverbal communication such as gestures and body movements of the commander. However, the proposed system has been designed in a way that it can produce only the most likelihood results in these situations. It would be interesting for future work to consider the information conveyed through nonverbal communication for enhancing the interpretation of relative references.

7.5 Summary

In this chapter, a method has been introduced to interpret uncertain information associated with relative references. This enables the users to issue voice instructions that include relative references related uncertain information in order to navigate a mobile service robot inside heterogeneous domestic environments.

A module called Relative Uncertainty Interpreter (RUI) has been deployed into the UIUM to evaluate the relative references related uncertain information. The RUI has been designed in such a way that it can replicate the natural tendencies of humans to a greater extent.

The main improvement of the concept proposed in this chapter over the existing approach is the system is capable of interpreting user instruction that include uncertain information associated with the relative references in more human like manner. Furthermore, this facilitates the users to issue instructions that involve multiple uncertain terms in a single command. Hence, command understanding ability of the robot has been improved and subsequently human friendliness of the robot has been improved.

Chapter 8

IMPROVING ROBOT'S PERCEPTION OF UNCERTAIN INFORMATION BY EVALUATING INFLUENTIAL GESTURE NOTIONS

---

one of main downside of the existing approaches for dealing with uncertain information in language instructions is that the existing approaches are only capable of interacting with the user through voice instructions. Therefore, those systems are not capable of evaluating the information conveyed through nonverbal instructions for improving the understanding of language instructions similar to humans. According to [46, 92] the understanding of voice instructions could be improved by fusing the information conveyed from gestures with the language instructions. However, the proposed systems are not capable of quantifying distance-related uncertain information in user instructions. Hence, the proposed methods cannot be adopted in order to improve the quantification ability of uncertain information in voice instructions. Therefore, this chapter investigates a method to adapt the perception of uncertain information in navigational commands such as "little" and "far" based on the arrangement of the environment and the notions conveyed through the pointing gestures of the user.

8.1 Motivation and Insight behind the Fusing of Notions Symbolized from the Gestures

The example situation shown in Fig. 8.1 is used to explain the main downside of the work proposed in chapter 4. The adaptation of the perception of distance-related uncertain information in the system proposed in chapter 4 is entirely based on the environmental factors such as size of the room, free space and the arrangement of the environment. The arrangement of the environment is perceived by the robot though the perceptive distance, D (where $D = D_r$, D_r is the distance to the nearest obstacles in the intended moving direction and it is annotated as D_r in Fig. 8.1). In this scenario, if the robot is commanded; "move little forward", "move medium forward", and "move far forward" it will move to location '1', '2', and '3' respectively. However, if the user wants to navigate the robot to an intermediate location other than these three locations (e.g., in between robot and '1' and in between '1' and '2'), it is not possible to navigate the robot using a single instruction since the user cannot influence the interpretation done by the robot. The robot would have to be issued a series of instructions (e.g., first "move medium forward" then "move little backward"...) in order to navigate to such a location; the user cannot alter the movement as required since the entire interpretation process running on the robot considers only the environment parameters that are fixed for a particular situation. Therefore, this drawback has to be cleared in order to increase the effectiveness of quantification of uncertain descriptors by the robot.

The notions conveyed through the pointing gestures of the user can be used in order to enhance the understanding of voice instructions by the robot [28, 46]. However, the exact meaning of the notions conveyed from gestures is highly imprecise and solely the notions conveyed from the gestures are not sufficient for identifying an exact location, position or object [28, 46]. Moreover, the notions conveyed from the gestures are useful for improving the understanding of language instructions. Therefore, the influential notions conveyed from the pointing gestures accompanied with the voice instructions are fused to the inferencing process of the uncertain spatial descriptors. This will contribute towards resolving the aforementioned drawback. The example situations shown in Fig. 8.2 are used in order to explain the insight of the proposed fusing method. In this example scenario, the person 'A' issues the command, "move little forward" to the person 'B' and the person 'B' has to quantify the meaning of the term "little". The language instruction may be accompanied by gesture instructions as shown in here. If the locations pointed by the hand gesture in the cases are not the same, the person 'B' tends to differently quantify the distance meant by "little" due to the influence caused by the gesture. According to [93], the humans tend to shift the perception of goal position towards the area indicated by the pointing gesture. Moreover, when the gesture is pointed towards a location further away from the person in the intended moving direction, it influences the person to move more towards that direction and vice versa. Therefore, the person 'B' will move more

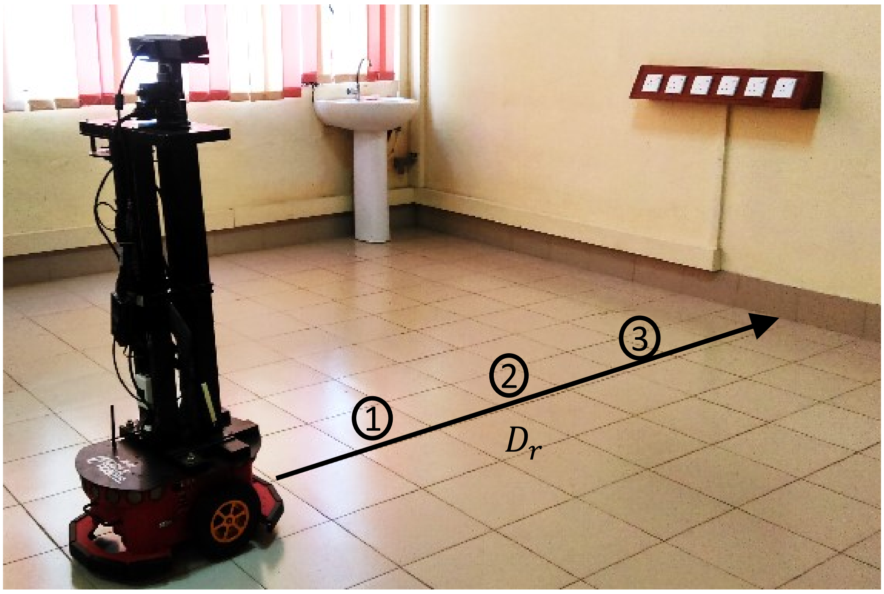
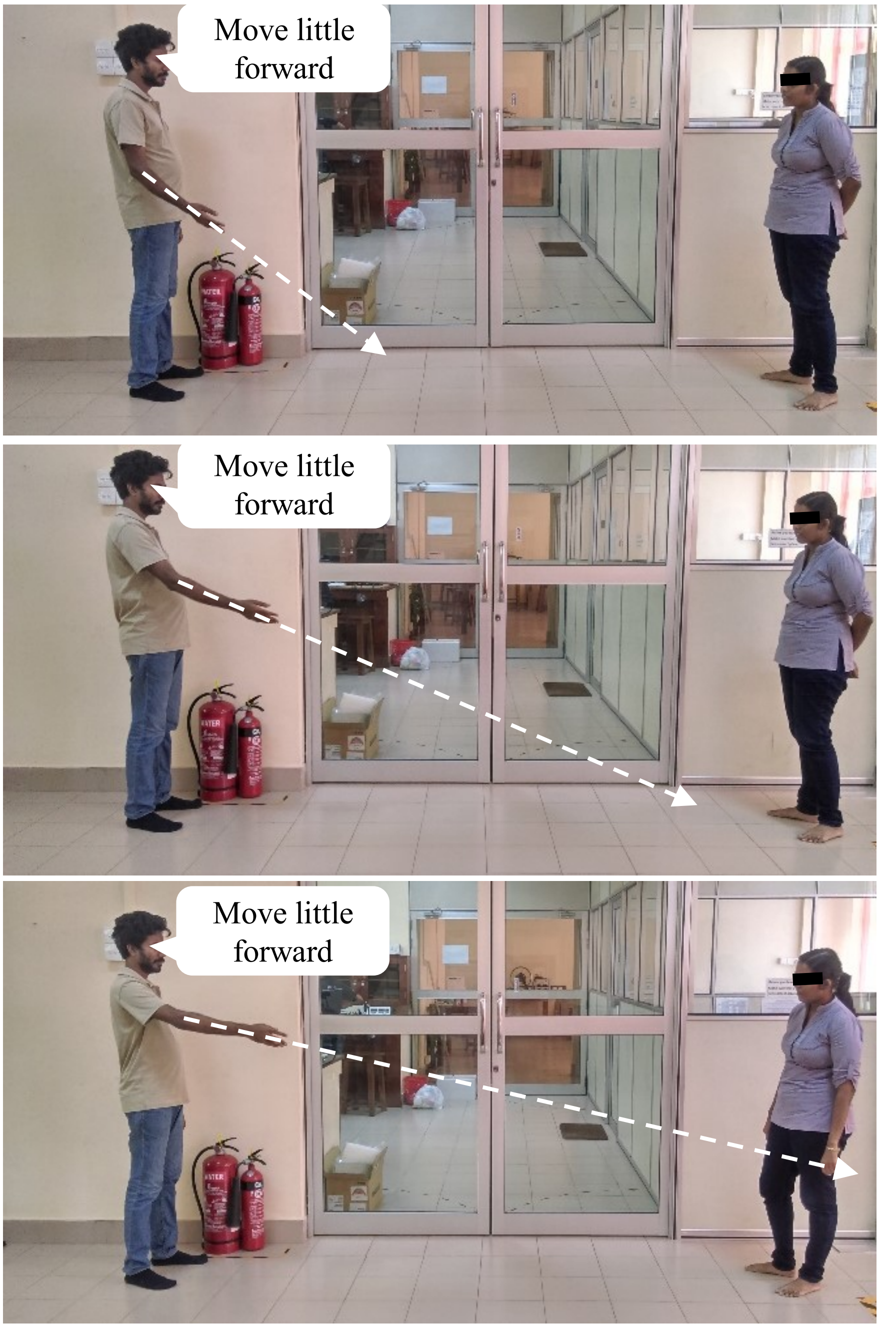


Figure 8.1: This illustrates an example scenario where the robot could be navigated with the user instructions. The robot will travel to location '1', '2' and '3' for commands "move little forward", "move medium forward" and "move far forward" respectively. It should be noted that the locations are annotated for the purpose of explanation and they do not reflect the exact positions. The perceptive distance, D_r is marked assuming the intended moving direction as forward.

distance in the case shown in Fig. 8.2(a) than the case shown in Fig. 8.2(b), even though the voice instructions and the environmental setting are the same. In case shown in Fig. 8.2(c), the person will still move towards the forward direction even though the gesture indicates a location in the back. However, the distance moved will be lesser. Furthermore, the position indicated by the gestures may be on a location where the navigation is not possible towards it (location is obstructed by an object or a wall). Therefore, the location given by the gesture cannot be solely used as the location meant by the person 'A' (experimental results are also provided in section 8.4 in support of this). Therefore, the notions conveyed from the pointing gestures are fused with the language instructions using a fuzzy inference system since fuzzy logic is infallible to such uncertainties and the approach proposed in chapter 4 is implemented with fuzzy logic. The fuzzy inference system has been designed in such a way that the system is capable of adapting the perception based on the environment and the influential notions conveyed from the gestures. This will clear up the aforementioned draw back of the system proposed in chapter 4 since this enables the user to influence the meaning of "little", "medium" and "far" by modifying the perception through gestures.

In order to realize this a fuzzy inference system based module called Voice and Gesture has been introduced into the Distance Interpreter (DisI) of the Intelligent System for Understanding Uncertain Information (ISUUI). This module is used when a valid gesture is identified by the system along with a voice instruction. When only the voice instructions are available, then the Voice module is used for the interpretation of uncertain information. The Voice module is almost similar to the submodule 1 of the system explained in chapter 4. The structure of the ISUUI with these two modules is shwon in Fig. 8.3.



(a)

(b)

(c)

Figure 8.2: (a), (b) and (c) show three situations, where the notions convoyed from the pointing gestures of person 'A' cause different influences to the distance that will be moved by the person 'B'. The elbow-wrist pointing vector is annotated in dashed-line arrow.

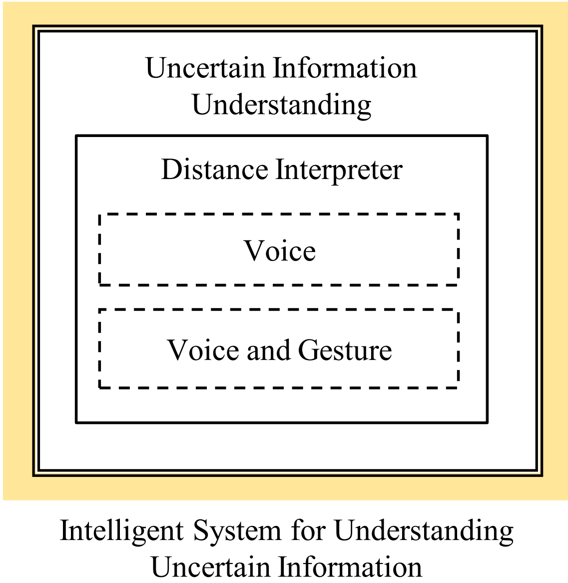


Figure 8.3: The structure of the ISUUI with the proposed Voice and Gesture module

8.2 Pointing-Gesture Evaluation

Skeletal information that can be retrieved from the Kinect motion sensor attached to the robot is used in order to identify the pointing gesture and to estimate the pointing position. The elbow-wrist vector is considered as the pointing vector (similar to the work in [46]) and it is extended until it crosses the floor plane. The point where the floor plane is crossed by the elbow-wrist vector is considered as the point that is referred by the user through the gesture (This point is calculated as explained in [93]). The displacement to the point along the intended moving direction (i.e. the direction indicated by the voice instruction) is calculated as D_g. This is illustrated on Fig. 8.4(a) and Fig. 8.4(b) for two example situations, where the direction of the movement is forward and left respectively. In order to consider a hand posture of the user as a pointing gesture, the joint positions should not be within the ranges defined for the rest positions and the elbow-wrist vector should point towards the floor plane. Furthermore, the pointing direction should be stable and the variation should be less than an experimentally decided threshold in order to consider it as a valid pointing gesture (The method used

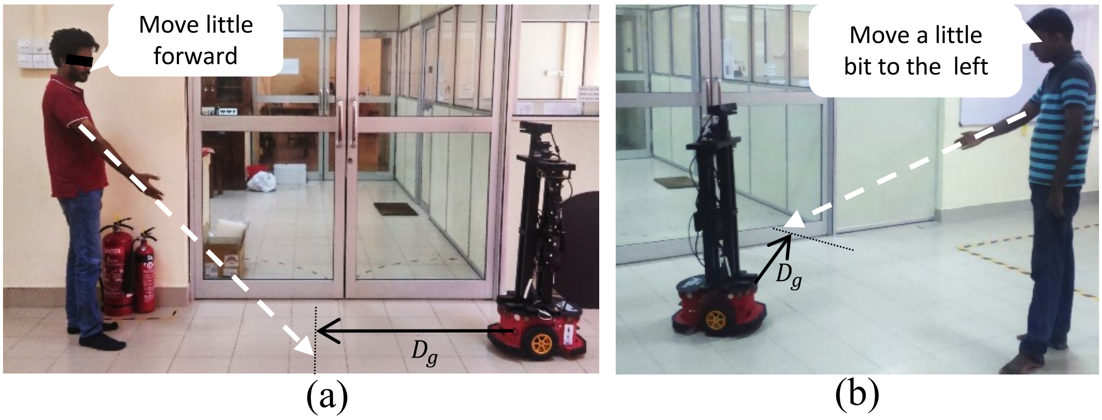


Figure 8.4: This illustrates the ways of obtaining of the parameter, D_g based on the elbow-wrist pointing vector. In (a) and (b), the explanation is done considering the indented moving direction as forward and left respectively. The same criterion is applied for other directions. The elbow-wrist pointing vector is annotated in dashed-lined arrow. It should be noted that the point referred by the gesture may not be in the intended moving path and D_g is the projection of the point on the intended moving direction.

in [93] is adopted for this). The capturing of Kinect sensor is triggered with the voice instruction and 10 consecutive frames are analyzed in order to extract the pointing gesture. It should be noted that the system has been designed and developed for single user situations and the system is only capable of detecting the gestures of a single person. If there are multiple people in the field of view of the Kinect, the system considers only the closest person. In this stage, it would be fine to consider only single user situations since the core contribution of the work is to addresses issues in interpreting navigation commands with uncertain information (interpreting phrases such as “move a little bit to the right” and “go far left”) by incorporating user gestures and spatial information of the environment. The situations with multi-users are not considered in the scope of the work presented in thesis and methods for handling such situations are proposed for future work.

8.3 Voice and Gesture Module

The voice and gesture module has been introduced in order to quantify the meaning of the uncertain spatial descriptor in a particular user instruction based on the environmental setting and the influential notions conveyed from the pointing gestures. As similar to the module for voice only instructions, the uncertain spatial descriptor in a particular command and the free space of the rooms are also taken as inputs to this module. The input membership functions for those two parameters are almost similar to the input membership functions of the voice module. In addition to that, this inference system has another input that accounts for adapting the perception based on the influential notions conveyed from the gestures. D_g obtained from the pointing gesture is considered as this input and the input membership function for D_g is modified according to the perceptive distance (D). The universe of discourse of this membership function runs from negative infinity to the positive infinity since D_g probably goes beyond D or the origin. The output of the system is the quantified distance of the uncertain spatial descriptor in a particular user instruction. The output membership function is modified according to the perceptive distance, D as similar to the voice module. The input and output membership functions of the voice and gesture module are given in Fig. 8.5. The rule base has been defined in such a way that the system can replicate the natural behavior of humans explained in section 8.1. The rule base of the fuzzy inference system is given in Table 8.1.

Table 8.1: Rule Base of the Fuzzy Module: Voice and Gesture

Input Memberships		Uncertain Descriptor								
		Little			Medium			Far		
Free Space		S	M	L	S	M	L	S	M	L
D_g	S	VVS	VS	S	S	SS	M	M	SL	L
	M	VS	S	SS	SS	M	SL	SL	L	VL
	L	S	SS	M	M	SL	L	L	VL	VVL

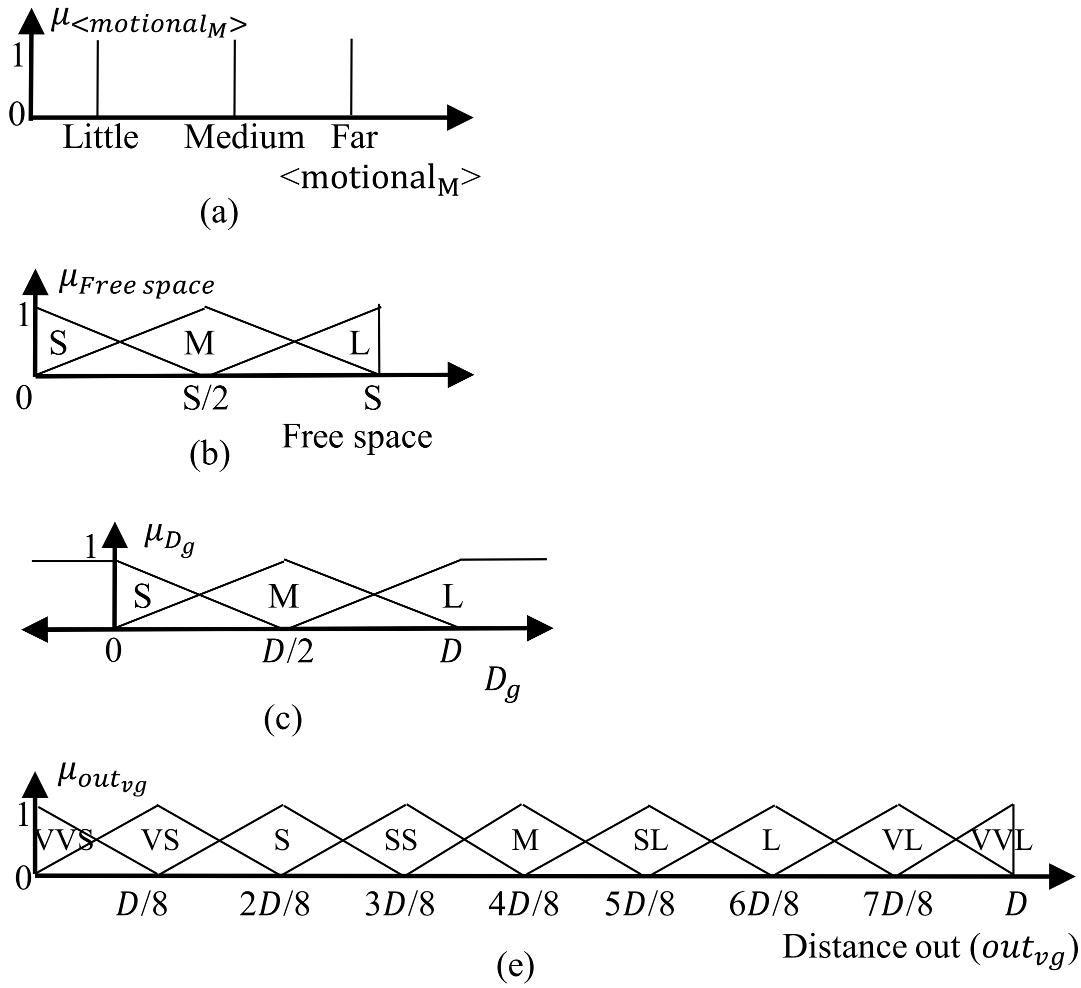


Figure 8.5: The input and output membership functions of the fuzzy inference system used in voice and gesture module are shown here. The membership functions (a), (b) and (c) are the inputs of voice and gesture module. It should be noted that the fuzzy set L and S of the membership function for D_g runs to positive and negative infinity respectively. (e) is the output membership function of the voice and gesture module.

8.4 Results and Discussion

8.4.1 Experimental setup

The proposed concept has been implemented on MIRob platform and experiments have been carried out in an artificially created domestic environment in order to validate the performance of the proposed concept in adapting the perception based on the environment and the influential notion symbolized from the pointing gestures of the user. Due attention has been paid to the recommendation given in [85] for designing and planning human studies for evaluating the human-robot interaction. The MIRob platform in few experimental scenarios are shown in Fig. 8.6. The arrangement of the environment used for the experiments can be visualized from the map shown in Fig. 8.7. Mainly, the experiments have been carried out in order to assess the performance improvement of the system due to the ability of adapting the perception based on the influential notions symbolized from the pointing gestures of the user. Therefore, the behavior of the proposed system (i.e., the system with the ability to adapt the perception based on influential notions symbolized from the pointing gestures) has been analyzed against the behavior of the system without the multimodal interaction ability (i.e., the system discussed in chapter 4). The evaluation has been carried out with participation of ten healthy persons who are in between 23-27 years (M = 25.4 years and SD = 1.71 years).

The participants were asked to navigate the robot as much as closer to a given target distance/position using only a one-step of instruction. At the start of the each operation, the target distance was clearly shown to the participants by the organizers of the experiment. Highly visible markers were not placed on the floor since they could have distracted users. Furthermore, the users could have tended to aim their hand gestures at the target, which would have detracted from the spontaneity. The participants were also instructed to be within the field of view of the robot when issuing the instructions to the robot. The same task was repeated for the robot with the proposed voice and gesture module (i.e., the system proposed in this chapter) and the system without voice and gesture module (i.e., the system proposed in chapter 4) by each user in each experimental layout. Five layout arrangements were considered for the evaluation process and each participant was given the chance to operate the robot in all the five experimental layouts. The behavior of the two systems and the internal parameters of the systems have

(a)
(b)
(c)
(d)
(e)

Figure 8.6: This shows the third person view and the view from Kinect sensor during the experimental cases given in Table 8.2. (a): case A, (b): case B, (c): case C, (d): case D and (e): case E

been recorded in each case. The sample results obtained from the experiments are given in Table 8.2. The initial and final locations of the robot and target positions in the corresponding sample cases given in Table 8.2 are marked on the map in Fig. 8.7. The corresponding third person view of the sample experimental cases along with the view of the Kinect are shown in Fig. 8.6.

In order to evaluate the performance of the two systems, the error between the given target distance and the quantified distance output of the two systems has been calculated. Absolute value of the errors has been considered for the analysis since a positive and a negative error may nullify the total error. The error of the output of the voice module (named as e_v) and the voice and gesture module (named as e_{vg}) are given in (8.1) and (8.2). where, out_v and out_{vg} are the quantified output of the voice module and voice and gesture module respectively. D_t is the target distance. Furthermore, the distance indicated by the pointing gesture (i.e., D_g) is also compared against the target distance by evaluating the error between those two (named as e_g) as given in (8.3) for analyzing the possibility of using solely the distance referred from the gesture. The errors calculated for each of the sample experimental cases are also given in Table 8.2. An explanatory video of the experiment is included in the supplementary multimedia attachment 5[1].

$$e_v = |out_v - D_t| \tag{8.1}$$

$$e_{vg} = |out_{vg} - D_t| \tag{8.2}$$

$$e_g = |D_g - D_t| \tag{8.3}$$

[1]Available in the attached CD and www.youtube.com/watch?v=dUbSdOD24SQ

Table 8.2: Sample Results of the Experiment

case	Layout	Room size (m^2)	Free space (m^2)	D_t (cm)	System with only voice module: Vocal command	D_r (cm)	out_v (cm)	Destination position	e_v (cm)	System with voice and gesture module: Vocal command	D_r (cm)	D_g (cm)	out_{vg} (cm)	Destination position	e_g (cm)	e_{vg} (cm)
A	1	15.08	12.95	40	move little left	204	86	A_v	46	move little left	203	20	51	A_{vg}	20	11
B	2	15.08	12.95	150	move medium forward	278	186	B_v	36	move medium forward	278	122	156	B_{vg}	29	6
C	3	15.08	12.95	120	move far backward	129	105	C_v	15	move far backward	131	324	116	C_{vg}	204	4
D	4	18.85	16.33	100	move little forward	429	182	D_v	82	move little forward	429	-206	91	D_{vg}	306	9
E	5	11.50	9.27	150	move medium right	288	187	E_v	37	move medium right	277	231	179	E_{vg}	81	29

In the execution of the system, the distance values are calculated in millimeters and the data given in here are rounded off to 1 cm. The parameters, Room size, free space and D_t are common for both the systems.

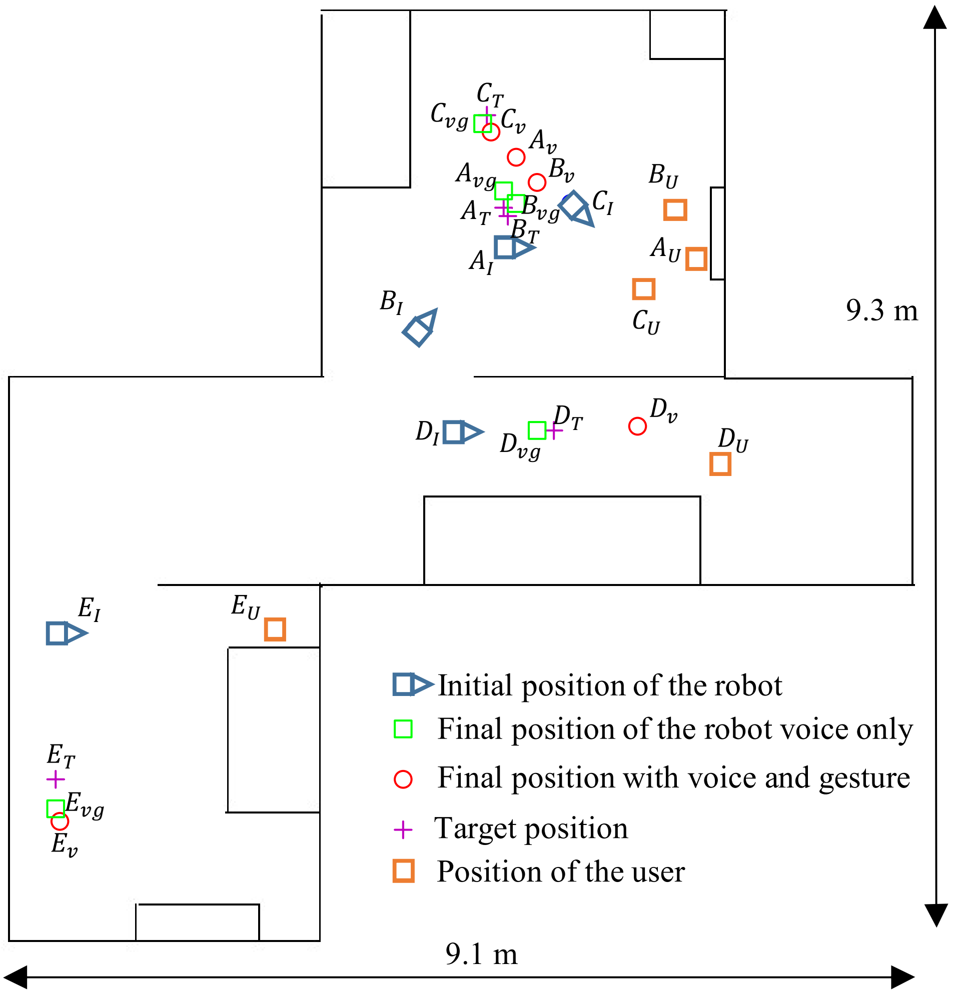


Figure 8.7: The initial and final positions of the robot during the experimental cases given in Table 8.2 are marked on the map with the corresponding indexes for both the systems. The map is drawn to a scale. However, it should be noted that the markers do not represent the actual size of the robot or the users.

8.4.2 Analysis of the Performance and Behaviors

Case A represents the user 1 in layout arrangement 1. The robot was initially placed on the location 'A_I' and the target distance (D_t) was 40 cm away from the left side of the robot (marked as 'A_T'). Then the user was asked to navigate the robot to a position very much closer to the target distance. In the first case, the interaction ability of the robot was limited to the voice interaction and it cannot adapt the perception based on the influential notions conveyed through the gestures. The user issued the command "move little left" in order to move it closer to the target location and the robot moved to location 'A_v' by interpreting 86 cm as the distance meant by "little" based on the current environment setting (i.e., room size = 15.08 m^2, free space = 12.95 m^2 and perceptive distance, $D =$ 204 cm). The error between the target distance and the actual moment in this case (i.e., e_v) was 46 cm. Then, the robot was again placed at the same initial position (i.e., 'A_I') with the activated gesture and voice module of the DisI and again the user was asked to move the robot to the target. In this instance, the user issued the same voice instruction as similar to the previous case and a valid hand gesture was identified by the Gesture Evaluation Module (GEM). The distance meant by the gesture (D_g) was 20 cm and the robot moved to location 'A_{vg}' by interpreting the quantitative meaning as 51 cm based on the environment and the influential notions conveyed from the gesture (i.e., room size = 15.08 m^2, free space = 12.95 m^2, $D = 203$ cm and $D_g = 20$ cm). The error between the distance interpreted by the robot and the target distance in this situation (i.e., e_{vg}) is 11 cm. Therefore, the system with the voice and gesture has produced a less error compared to the system with only voice interaction capabilities. Furthermore, the error between the target and the distance directly indicated by the gesture (i.e., e_g) was 20 cm that is higher than the distance quantified by the system based on the voice and the gesture. The third person view and the image captured from Kinect motion sensor in this instance are shown in Fig. 8.6(a).

The user 2 on layout arrangement 2 is given as case B. In this situation, the initial position of the robot was B_1 and the target position was 150 cm away from the front of the robot (marked as 'B_I'). In the situation, where only the voice module of the DisI was being activated, the user commanded "move medium forward" to the robot. The voice module of the DisI quantified the distance meant by the user as 186 cm (out_v) based on the environmental parameters (i.e. room size = 15.08 m^2, free space = 12.95 m^2 and D_r = 278 cm) and moved to the location 'B_v' which resulted 36 cm for e_v. Then, the robot was again placed at the same initial position (i.e. 'B_I') with the activated voice and gesture module of the DisI. This time, the user again issued the same voice instruction similar to the previous situation. However, the distance moved by the robot was 156 cm in this situation since the system was capable of adapting the perception of uncertain spatial descriptors based on the influential notion conveyed from the gesture. This movement resulted an error of 6 cm with the target distance, which was less than the system with voice only case. Furthermore, the distance meant by the gesture (i.e., D_g) was 122 cm and if it had been alone considered for the movement, it would have resulted an error of 29 cm that was quite higher than the voice and gesture case but slightly less than the voice only case.

The user 3 on layout arrangement 3 is given as case C. In this layout arrangement, the initial position of the robot was 'C_I' and the target position 'C_T' was 120 cm away from the backside of the robot. The system with only voice module case, the robot was commanded, "move far backward" by the user and it moved to location 'C_v' by resulting a movement distance of 105 cm. The error between the target and the moved distance (in here e_v) was 15 cm. The system with the activated voice and gesture module case, the robot was commanded with the same voice instructions with a gesture that influenced a higher movement by the user. The distance indicated by the gesture (i.e. D_g) was 324 cm which was greater than the perceptive distance ($D = 131$ cm). Moreover, the position meant by the gesture was a location away from the obstructed wall and table. Furthermore, the error between the target distance and the distance indicated solely by the gesture (i.e., e_g) was 204 cm. Therefore, if solely the gesture had been considered for the movement, it would have resulted a movement that cannot be achieved due to the possible obstructions. Furthermore, it would create a huge error if the movement were possible. However, the distance quantified by the system was 116 cm since it adapts the perception based on both environment setting and the influential notions conveyed from the gesture. Therefore, in this situation the performance of the system with voice and gesture module was par above the voice only case and gesture only case (It should be noted that the system with the gesture only case was not separately considered and the gesture notions evaluated in gesture and voice situation was used for the analysis).

Case D represents the situation where the user 4 operated the robot on layout arrangement 4. In this layout arrangement, the initial position of the robot was 'D_I' and the target location was 'D_T', which is 100 cm away from the front of the robot. In the situation where only the voice module was being activated, the robot was commanded "move little forward" by the user and robot moved to the location 'D_v' by quantifying the meaning of "little" as 182 cm. In this situation, the error of the movement was 82 cm, which is quite high. In the situation where the voice and gesture module was also being activated, the distance moved by the robot was 91 cm and the error with the target distance was 9 cm, which is very small with respect to the error in the previous situation (i.e., system with only the voice module). The position indicated by the gesture was in behind the robot (Hence, D_g got a negative value. D_g = -206 cm) even though it was commanded to move towards front. This situation was similar to the situation described in section 8.1 using Fig. 8.2(c). Therefore, if solely the notion conveyed from the gesture accompanied with the voice instruction had been used to decide the goal position, the robot would have moved to a location behind the robot which is completely erroneous. Moreover, solely the notions conveyed from the gesture accompanied with the voice instruction are not suitable for deciding the destination for the movement.

The user 5 in the layout arrangement 5 is considered as the case E. In this layout arrangement, the initial position of the robot was 'E_I' and the target position was 'E_T', which is 150 cm away from the right side of the robot. The robot moved 187 cm in the situation where only the voice module of the DisI was being activated and the error between the target and the movement was 37 cm. In the second phase (i.e. with both the modules of the USDI were being activated), the distance moved by the robot was 179 cm and the error was 29 cm which is less than the earlier situation. The distance meant by the gesture, (i.e., D_g) was 231 cm and if the gesture alone had been considered for the movement, it would have caused an error of 81 cm, which is quite high.

Similarly, all the users were given the chance to operate the robot in all the layout arrangement and the data were gathered. The mean values of the calculated error values (i.e. means of e_v, e_g and e_{vg}) are plotted on the graph shown in Fig. 8.8 along with the error bars, for each layout arrangement (i.e. for layout arrangement 1 to 5).

In all the layouts except the layout 1, the error between the distance indicated by the gesture and the target (i.e., e_g) is the highest. It should be noted that e_g was calculated based on the position meant by the gesture accompanied with the voice instructions and the users were not asked to navigate the robot using solely gesture instructions since it may withdraw from the natural behavior. In layout number 2, the arrangement of the user and the target position may have influenced in order to have a less error for the distance meant by the gesture since the users may easily aim their hand towards the target position by controlling more degree of freedoms of the arm. However, the exact reasons for variation of the position meant by gestures have not been investigated in this research since the contribution of the work is to fuse the influential notions conveyed from the gestures to adapt the perception of uncertain spatial descriptors related to the navigation command for improved quantification of uncertain descriptors in voice instructions. As explained in cases D and C above, if solely the position indicated by the gesture had been used to decide the goal the robot would have moved to

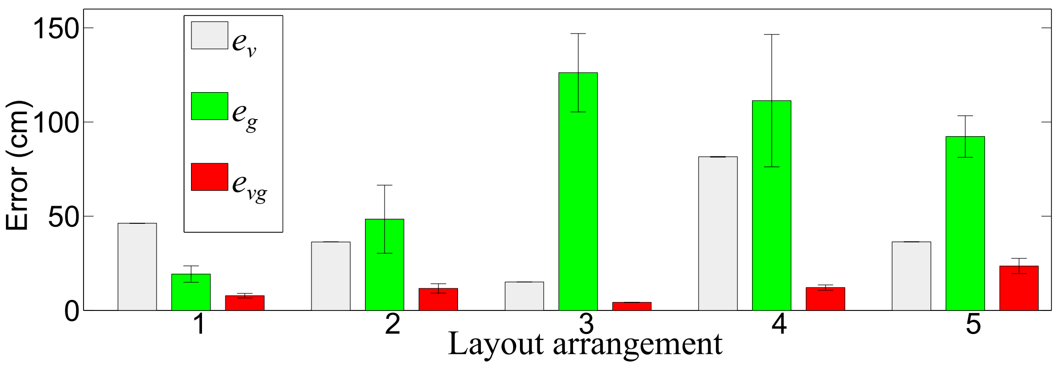


Figure 8.8: This visualizes the mean values of the error values (i.e. e_v: Voice, e_g: Gesture and e_{vg}: Voice and Gesture) for the considered 5 layout arrangements by all the 10 users. The error bars show the standard error.

a completely wrong position or the goal position would be a non-achievable one. Furthermore, as seen in the graph, e_g has a high variance compared to others and this implies that the distance indicated by the gesture has a high variation. Therefore, solely the position indicated by a gesture is not feasible for deciding the goal position when the pointing gesture is conveyed with a voice command.

In all the layout arrangements, the system with the voice and gesture has the lowest error in each layout. The significance of the results have been analyzed using multiple comparison tests for one-way ANOVA test. According to the test statistics, the system with the activated voice and gesture module of the DisI has a statistically significantly ($P<0.05$) lower error than the system with only the voice module of the DisI in all the five layout arrangements. The error reduction effected to the system due to the addition of the voice and gesture module is remarkable since the Cohen's d value greater than 0.8 can be seen (Cohen'd value greater than 0.8 implies a large effect [88]). Furthermore, the powers of the statistical tests are greater than 0.8 and it implies that the experimental results genuinely validate that e_{vg} is less than e_v (according to Cohen's four-to-one weighting of beta-to-alpha risk standard [88]). Based on these observations, it can be concluded that the system with activated voice and gesture module outperformed the system with only the voice module in quantifying the uncertain spatial descriptors. Moreover, fusing of the notions conveyed from the gesture for adapting the perception has improved the quantification ability of uncertain spatial descriptors related to navigation instructions such as "little" and "far" by the robot.

8.5 Summary

A method has been proposed in this chapter to adapt the perception of uncertain spatial descriptors related to navigation instructions such as "little" and "far" based on the current environment setting and the influential notions conveyed from the pointing gestures accompanied with the voice instructions.

The voice and gesture sub-module has been introduced to the DisI in order to quantify the meaning of uncertain spatial descriptors based on the environment and the influential notions conveyed from the pointing gestures accompanied with the voice instructions. This module has been implemented with fuzzy logic based on the natural tendencies of humans. The quantified output of the fuzzy inference system varies with the degree of influence conveyed from the pointing gesture in a particular environment setting. That is the main enhancement of the proposed concept over the existing approaches for quantifying the distance-related uncertain information in navigation commands.

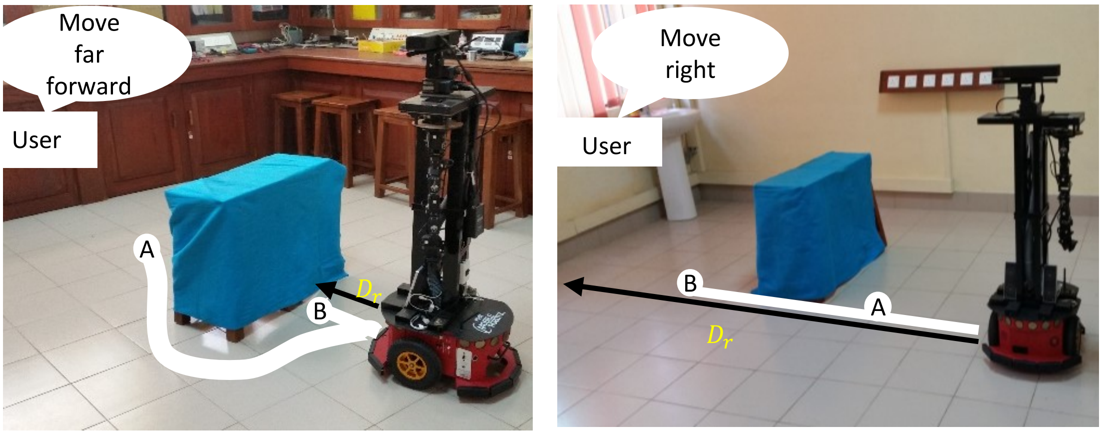
According to the obtained experimental results, the proposed concept is capable of remarkably reducing the error of quantifying the uncertain spatial descriptors than the existing approaches. Moreover, the quantification effectiveness of uncertain spatial descriptors by the robot has been improved.

Chapter 9

RESOLVING AMBIGUITIES IN NAVIGATION INSTRUCTION BASED ON MOTION INTENTION SWITCHING BY IDENTIFYING THE ACTUAL INTENTION OF THE USER

9.1 Rationale behind the Evaluation of Gestures Accompanied with User Instructions for Resolving the Ambiguities Arisen due to the Spatial Arrangement

The two example scenarios given in Fig. 9.1 are considered for investigating the limitations of the system explained in chapter 4. In scenario (a), the user issues the command, "move far forward". In this situation the maximum quantified output of the system explained in chapter 4 will be less than D_r since the perceptive distance is limited to D_r. Therefore, the robot will move to position 'B'. However, there are situations where the intention of the user is to move the robot to a position similar to location 'A' since the user expects that the robot can see beyond the obstacle. In situation (b), the user issues the command, "move right". In this situation the quantified output of the system proposed in chapter 4 will result in a movement of the robot to location 'B'. However, there are situations where the intention of the user is to move the robot to a location similar to the location 'A' since the user expects that the robot can consider the nearby obstruction for adapting the perception. Therefore, the system proposed in chapter 4 is not capable of understanding the intention of the user effectively



(a) (b)

Figure 9.1: (a) and (b) show two example situations that exhibit the limitations of the work proposed in chapter 4. The position requested by the user may be either positions 'A' or 'B'. However, the existing system considers only position 'B'. It should be noted that the annotated positions and paths are not exactly those generated from the systems and these are marked for the sake of explanation.

to resolve this ambiguity arisen due to the spatial arrangement.

As similar to the work proposed in chapter 8, the information conveyed from pointing gestures is analyzed in order to identify the intention of the user effectively. The two example scenarios given in Fig. 9.2 are considered for the explanation of the gesture-based user intention identification process that can be used in order to resolve the above-mentioned ambiguity. In case (a), the user is pointing to a location that is well beyond the default perceptive distance (i.e., D_r). Therefore, if the gesture is being pointed towards a location well beyond the default perceptive distance it can be concluded that the intention of the user is to navigate the robot beyond D_r (i.e., location 'A' instead of 'B' in Fig. 9.1(a)). Similarly, in case (b), if the user is pointing to a location that is well within the default perceptive distance (D_r), then it can be concluded that the intention of the user is to move the robot to an alternative position 'A' instead of position 'B' in Fig. 9.1(b).

The Motion Intention Switcher (MIS) is proposed in this chapter to resolve the above-mentioned ambiguity in navigation instructions. Based on the intention of

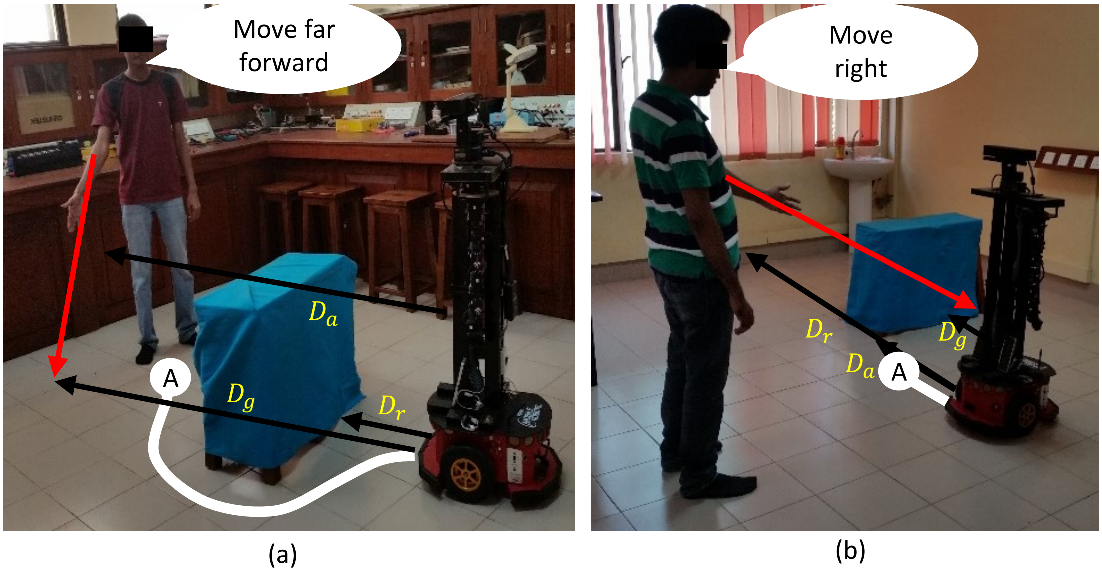


Figure 9.2: (a) and (b) show two example scenarios that explain the possibility of using pointing gestures in order to identify the intention of the user for switching the perceptive distance. It should be noted that the annotated positions, paths and vectors are not exactly those generated from the system and these are marked for the sake of explanation.

the user identified from the gesture information, the required actions for fulfilling a command may be switched by the MIS. Subsequently, the parameters required for the quantification of the uncertain information by the submodule 1 of the Distance Interpreter (DisI) (see chapter 4) will be modified by the MIS, if alterations are required. The structure of the Intelligent System for Understanding Uncertain Information (ISUUI) is depicted in Fig. 9.3.

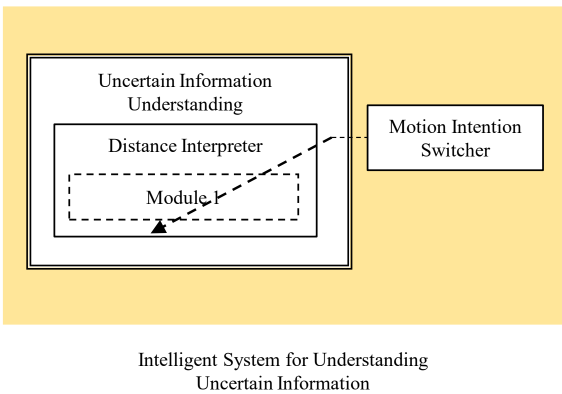


Figure 9.3: The structure of the ISUUI with the proposed MIS.

9.2 Motion Intention Switcher (MIS)

The desired position for the movement cannot be directly taken as the position referred from the gesture since the point referred from the gesture is not very accurate and typically it would not be the exact location that the user wants to navigate the robot (see chapter 8). Moreover, the gesture instructions are often useful in enhancing the meaning of vocal instructions in humanrobot interaction [28,46]. Therefore, it is only used for altering the perceptive distance (D) to an alternative perceptive distance (indicated as D_a in Fig. 9.2) from the default (i.e., D_r) by identifying the actual intention of the user. The assigning of alternative perceptive distance (D_a) for the perceptive distance (D) is done by MIS if required. The decision as to whether the perceptive distance has to be altered to an alternative (D_a) is decided based on a rule-based approach that evaluates D_g and D_r.

The procedure of assigning the perceptive distance D, is given in Algorithm 3. δ_{max} and δ_{min} are scalar constants used in order to avoid the false triggering of the intention switching due to the less accurate D_g. The alternative perceptive distance, D_a, has two cases where the $D_a > D_r$ and $D_a < D_r$. If $D_a > D_r$, then it is considered as $D_{a,max}$ and if $D_a < D_r$, then it is considered as $D_{a,min}$. Moreover, the MIS shifts the perception of robot between the alternative and default hypotheses based on the defined thresholds that depend on the pointing gesture issued by the user and the layout of the surrounding environment.

9.2.1 Estimation of Alternative Perceptive Distance (D_a)

The estimation of alternative perceptive distance, D_a is illustrated in Fig. 9.4 considering the two possible cases where $D_g > \delta_{max} D_r$ and $D_g < \delta_{min} D_r$. A field angle of α in the intended moving direction is considered for estimating the D_a. The field angle, α, is considered as 30° since according to chapter 4, the objects in

Algorithm 3 Assigning perceptive distance (D)

INPUT: D_r, D_g, D_a

OUTPUT: D

if $D_g > \delta_{max} D_r$ **then**
    $D = D_{a,max}$
else if $D_g < \delta_{min} D_r$ **then**
    $D = D_{a,min}$
else
    $D = D_r$
end if

that region have a higher impact for the human mobility. In case (a), D_a should be a value greater than D_r since $D_g > \delta_{max} D_r$. Therefore, $D_{a,max}$ exists and in order to estimate that, a vector parallel to the direction of the intended moving direction (i.e., the vector parallel to D_r) is extended until it reaches another obstruction for the movement inside the considered field. The magnitude of this vector is considered as $D_{a,max}$ in such cases (i.e., cases where $D_{a,max}$ is required as a result of $D_g > \delta_{max} D_r$). In case (b), D_a should be a value less than D_r since $D_g < \delta_{min} D_r$. Therefore, $D_{a,min}$ is required. The distance along a path parallel to the default intended moving path (i.e., parallel to D_r) to an obstacle within in the considered field from the robot is taken as the $D_{a,min}$ in such cases. If $\delta_{main} D_r \leq D_g \leq \delta_{max} D_r$ or a valid gesture is not detected (i.e., $D_g = null$) , the default perceptive distance (D_r) is considered as the perceptive distance (D) and hence the intention of the robot will not be switched from the default to an alternative intention in such cases.

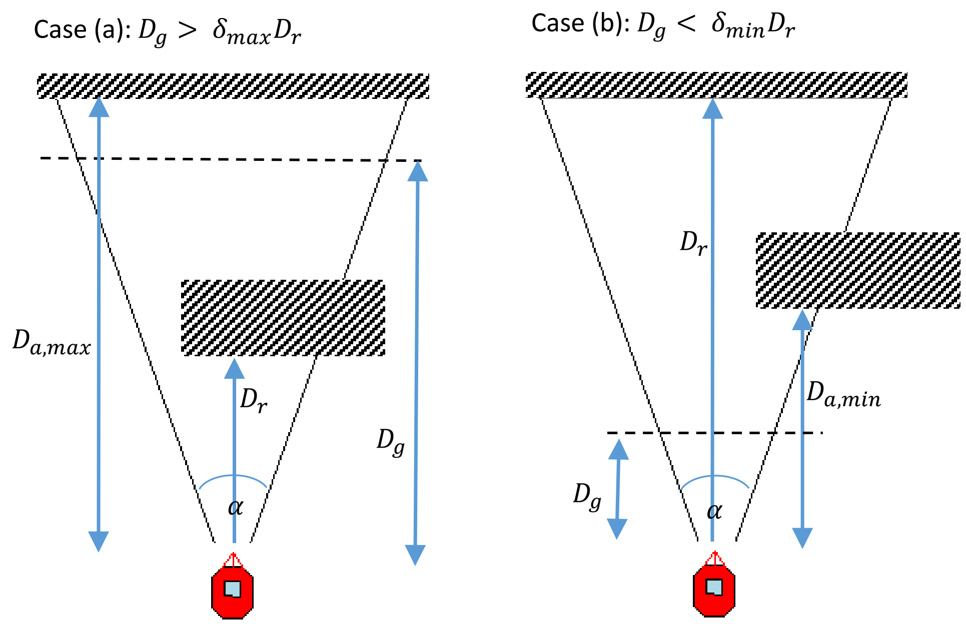


Figure 9.4: The ways to estimate the alternative perceptive distances are illustrated for the possible two scenarios. The shaded areas represent the obstacles/objects in the environment that are in near vicinity of the considered field view. The field angle is denoted as α. The dashed-line represents the perpendicular drawn to the intended moving path from the evaluated gesture pointing position in each scenario. D_g is calculated based on the point referred by the gesture as explained in section 8.2. D_r, $D_{a,min}$ and $D_{a,max}$ are computed based on the data of navigation map. This illustrates the parameter estimation considering the indented moving direction as forward. The same is applied for other directions too.

9.3 Results and Discussion

9.3.1 Experimental Setup

The proposed concept has been implemented on the MIRob platform and experiments have been carried out in an artificially created domestic environment in order to validate the behavior of the proposed system in switching the perceptive distance according to the intention of the user based on the pointing gestures accompanied with verbal instructions. Furthermore, another set of experiments has been carried out in order to evaluate the performance gain of the proposed method over the work explained in chapter 4 (i.e., system without the intention switching ability) which is not capable of analyzing the information conveyed through gestures. The evaluation was carried out with five healthy participants (average and standard deviation of the age of the participant are 25.2 years and 1.7 years, respectively) and they were graduate students in the university. The experiments have been carried out based on the guidelines suggested in [85] for designing, planning and executing human studies for humanrobot interactions in order to avoid the subjectivity of the experimental results. The scalar constants δ_{max} and δ_{min} are chosen experimentally as 1.5 and 0.75, respectively, in for achieving the desired characteristics.

9.3.2 Validation of the Behavior of the Motion Intention Switcher (MIS)

In order to validate the behavior of MIS in switching the intention based on the pointing gestures, experiments have been carried out in 10 different layout scenarios where such intention switching may be required in order to effectively evaluate the user instructions. Each participant was given the chance to perform the evaluation in any two of the previously unused arrangements among these 10 scenarios. The behavior of the proposed method (i.e., the system with MIS) and the system without the intention-switching ability (i.e., the system presented in chapter 4) have been analyzed in those situations. The sample results obtained from the experiment are given in Table 9.1. The views from the robot with tracked skeletons of the users in the sample cases are shown in Fig. 9.5 along with the third person view of the scenarios. The corresponding positions of the robot during the execution of each case are marked on the map shown in Fig. 9.6.

In case (a), the robot was initially placed on the location 'a_I' without deploying the MIS to the system. Then, the robot was commanded, "move far forward". The uncertain term in the command is "far" and the robot had to quantify the meaning of "far" for fulfilling the user command by navigating to the desired location. In this case, D_r was 33 cm since the robot only considers the immediate obstruction in its intended straight moving path. Therefore, the perceptive distance D was 33 cm and the quantified output generated from the UIUM was 29 cm, resulting a destination position in between the robot and the obstacle as explained in section 9.1. Therefore, the robot moved to location 'a_B'. Then, the MIS was activated and the robot was again placed at the initial position (i.e., location 'a_I'). The robot was again commanded with the same voice instruction accompanied with a pointing gesture that expresses that the intention of the instruction is to navigate the robot to a position that is beyond the obstacle. The gesture evaluation system interpreted the gesture and calculated D_g was 121 cm. In this situation, the perceptive distance was altered by MIS to alternative perceptive distance $D_{a,max}$ since $D_g > \delta_{max} D_r$. $D_{a,max}$ was evaluated as 252 cm and it was assigned to the perceptive distance (D). Therefore, the output of the UIUM was 199 cm that resulted a destination position beyond the obstacle and then robot moved to location 'a_A' by taking a curvy path generated by the navigation controller for avoiding the obstacle.

In case (b), the robot was initially placed in location 'b_I' with disabled MIS. Then it was commanded, "move medium right". The robot had to quantify the meaning of the uncertain term "medium" in order to move to the destination position requested by the user. Here, D_r was 272 cm. Subsequently, D and the quantified outputs were 272 cm and 181 cm, respectively. Therefore, the robot moved to location 'b_B' that is located well past the nearby obstacle. Then, the robot was again placed in the same initial position (i.e., 'b_I') with enabled MIS. This time the robot was commanded with the same voice instruction accompanied with a pointing gesture that expresses the intention of the user is not to move the robot to a location well past the nearby obstacle. Here, the D_g was 57 cm, that lead to assigning of $D_{a,min}$ to D since $D_g < \delta_{min} D_r$. The evaluated $D_{a,min}$ was 72 cm since the robot considers the distance to the nearby obstacle in the considered field along the intended moving direction. Therefore, the quantified output was 48 cm, which resulted the movement of the robot to location 'b_A' where the robot is not required to move beyond the nearby obstacle.

Figure 9.5: The view of the robot and the third person view of sample scenarios are shown with the corresponding case (**a**–**e**) given in Table 9.1. The tracked skeletons of the users are also superimposed with the RGB view of the robot for better clarity.

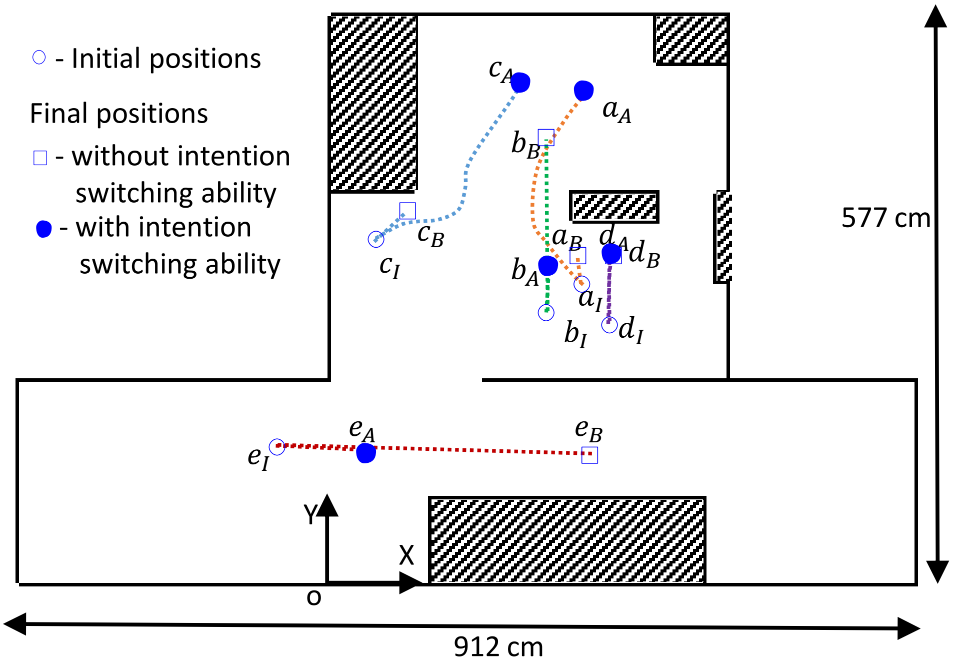


Figure 9.6: The initial and final positions of the robot during the experiment for identifying the behaviors of the proposed method are marked on the map with corresponding case letters. The shaded areas represent the objects in the environment. The map is drawn to a scale. However, it should be noted that the markers do not represent the actual size of the robot.

In case (c), the initial position is location 'c_I' and it was commanded, "move far forward". The system without the MIS quantified the meaning of "far" as 42 cm by considering the default perceptive distance and the robot moved to location 'c_B'. The quantified output of the system with MIS was 218 cm since it considered the $D_{a,max}$ as D since the evaluated gesture indicated a request to change the default intention.

In case (d), the initial location was 'd_I' and it was commanded, "move far forward". The quantified output of the system without the MIS was 70 cm and the robot moved to location 'd_B'. In the system with the MIS case, D should be altered to $D_{a,min}$ since $D_g < \delta_{min} D_r$. However, in this situation $D_{a,min}$ and D_r were the same. Therefore, D was not altered and the quantified output is the same as the system without MIS. Therefore, the robot moved to location 'd_A' which was almost the same as 'd_B' (due to navigational errors there is a very small different in position coordinates). In this case, the intention of the user was to express his intention of navigating the robot to a location that is in between the obstacle and the robot without altering the default intention. Moreover, the proposed system is capable of successfully handling such situations.

In case (e), similarly to the case (b) the robot with MIS switched the intention by identifying the actual intention of the user by analyzing the instructions conveyed from pointing gestures given along with voice instructions. Similarly, the behavior of the MIS was found to be capable of effectively switching the intention of the robot according to the actual intention of the user in all the test cases. An explanatory video that shows the behaviors of the two systems in a similar kind of experimental scenario is provided as a supplementary material in the multimedia attachment 6[1]. It shows the video footage from a third person's view along with the traced location of the robot within the navigation map. Furthermore, parameters used in the interpretation process of the commands are also given with annotated explanations.

[1] Available in the attached CD and www.youtube.com/watch?v=bBSouHJhTzY

Table 9.1: Sample Results of the Experiment for Validating the Behaviours of the MIS

User command	**Initial position** X, Y, θ	**Uncertain term**	**Room size** (m^2)	**Free space** (m^2)	**without Motion Intention Switcher** D_r (cm)	D (cm)	**Output** (cm)	**Final position** X, Y, θ	**with Motion Intention Switcher** D_r (cm)	$D_{gesture}$ (cm)	D (cm)	**Output** (cm)	**Final position** X, Y, θ
a move far forward	254, 302, 88	far	15.08	12.77	33	33	29	252, 329, 95	33	121	252	199	254, 500, 90
b move medium right	220, 272, 179	medium	15.08	12.77	272	272	181	218, 452, 87	274	57	72	48	217, 319, 93
c move far forward	46, 344, 49	far	15.08	12.77	64	64	42	78, 375, 50	66	180	269	218	189, 509, 49
d move far forward	285, 260, 87	far	15.08	12.77	86	86	70	289, 330, 87	85	-140	85	70	283, 334, 88
e move medium forward	-53, 135, -5	medium	18.55	16.33	470	470	313	262, 125, -3	470	100	130	86	33, 127, -5

9.3.3 Evaluation of Performance Gain of the Proposed Method

A set of experiments has been carried out in order to compare the performance gain of the system with MIS (i.e., the proposed system) over the system without MIS (i.e., the system explained in chapter 4). For this experiment, the users were asked to navigate the robot from a given initial position to a given goal position marked on the floor as shown in Fig. 9.7. The number of steps taken for navigating the robot towards the goal has been considered as the parameter for the evaluation work based on the experimental evaluation carried out in the work presented in [94]. The same task was repeated for both systems and the information related to the systems was recorded. Ten different layout arrangements (i.e., with different initial and goal positions) have been selected by randomly choosing the initial and goal positions. The initial and goal positions for a particular layout scenario have been kept within the same room since it is impractical to navigate the robot from one room to another room using only this kind of simple motion command. Furthermore, such navigation tasks could be deduced into this kind of problem by using the ability of the robot to understand a command like "move to the kitchen" as explained in chapter 3. All the participants have been given the chance to perform one by one in all 10 layout arrangements and the results have been analyzed in order to evaluate the value addition of the proposed MIS. It should be noted that this experimental scenarios are independent of the experimental scenarios discussed in experiment 1 (i.e., in section 9.3.2).

The data of the experiments for user 1 in layout arrangement 1 (i.e., named as case 1) and user 1 in layout arrangement 2 (i.e., named as case 2) are given in Table 9.2 as sample results. The corresponding positions of the robot after executing each user instruction are marked on the map shown in Fig. 9.8. The positions are annotated with the corresponding indexes given in Table 9.2.

Figure 9.7: This shows the experimental scenario of the case 1 of the experiment for comparing the performance of the system with the MIS and the system without the MIS. The user was asked to navigate the robot to the goal position marked on the floor by implementing both the system in the robot. The goal area is annotated as "goal" in here.

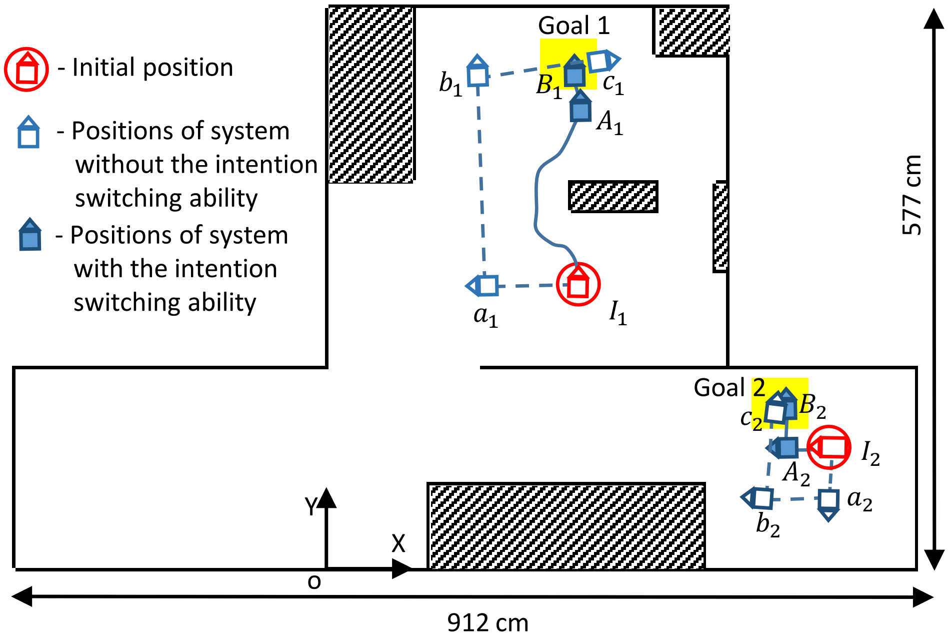


Figure 9.8: The positions of the robot after executing each user instructions given in Table 9.2 for cases 1 and 2 are marked on the map with the corresponding indexes. The shaded areas represents the objects in the environment. The light color solid areas represent the positions of the goals. The map is drawn to a scale. However, it should be noted that the markers do not represent the actual size of the robot.

Table 9.2: Sample Results of the Experiment for Evaluating the Performance Gain of the System with the MIS

	User command	**Uncertain term**	**Room size** (m^2)	**Free space** (m^2)	D_r (cm)	$D_{gesture}$ (cm)	**Intention switched**	D (cm)	**Distance moved** (cm)	**Position** (X,Y,θ)
Case 1	Initial position									I_1 (247,283,89)
with MIS	A. move medium forward	medium	15.08	12.77	57	128	True	275	183	A_1 (250,466,89)
	B. move little forward	little	15.08	12.77	87	Not detected	False	87	36	B_1 (249,502,89)
without MIS	a. move little left	little	15.08	12.77	206	-	-	206	86	a_1 (160,283,-179)
	b. move far right	far	15.08	12.77	270	-	-	270	219	b_1 (149,502,92)
	c. move medium right	medium	15.08	12.77	149	-	-	179	117	c_1 (270,519,8)
Case 2	Initial position									I_2 (504,117,179)
with MIS	A. move little forward	little	18.55	16.33	470	60	True	102	42	A_2 (462,118,179)
	B. move medium right	medium	18.55	16.33	63	Not detected	False	63	42	B_2 (460,159,87)
without MIS	a. move medium left	medium	18.55	16.33	98	-	-	98	64	a_2 (504,57,-89)
	b. move medium right	medium	18.55	16.33	106	-	-	106	71	b_2 (434,68,175)
	c. move far right	far	18.55	16.33	110	-	-	110	89	c_3 (447,152,83)

In this case, the initial position of the robot was 'I_1' and the goal position is annotated as 'goal 1' in the map. In the system with the MIS event, first the robot was commanded, "move medium forward" while being shown a gesture that expresses the requirement of switching the intention to navigate the robot beyond the obstacle in the front. D_r and D_g were 57 cm and 128 cm, respectively. The intention of the robot was switched by the MIS since $D_g > \delta_{max}D_r$ and $D_{a,max}$ was assigned to the perceptive distance (D). Therefore, D was 275 cm and subsequently the quantified distance output was 183 cm which resulted the movement of the robot to location 'A_1'. Then the robot was commanded, "move little forward" and a pointing gesture was not detected by the system since a pointing gesture was not issued by the user. Therefore, the intention of the robot was not switched and the robot moved 36 cm by considering D_r as perceptive distance (D). The moved position was 'B_1' that was inside the given goal area. Therefore, this was considered as the completion of the task. Then, the robot was placed on the same initial position (i.e., 'I_1') after disabling the MIS (i.e., system similar to the system explained in chapter 4) and again the user was asked to navigate the robot to the goal. In this event, if the user had commanded the robot "move medium forward" similar to the earlier event, the robot would have moved to a point between the obstacle and the robot (due to the limitation of the system without MIS discussed in section 9.1). However, that movement would be a waste since the user cannot navigate the robot beyond the obstacle without changing the moving direction. Therefore, with this in mind, the user first issued the command "move little left" in order to take away the robot from the barrier. The robot quantified the distance meant by "little" as 86 cm by considering the default perceptive distance and moved to position 'a_1'. Then the robot was commanded "move far right" and robot moved to position 'b_1' in order to fulfill the request of the user. Then, the robot was commanded "move medium right" and the robot moved to position 'c_1' which was inside the goal area. Therefore, the task was completed. In order to complete the task with the system with the MIS, the user had to issue only two user instructions while with the system without the MIS, the user had to issues three instructions in order to complete the tasks. Moreover, the work overhead of the user is comparative less when the MIS is deployed into the robot.

In case 2, the initial position of the robot was 'I_2' and the goal is annotated as 'goal 2' in the map. In the system with the MIS event, the user first issued the command "move little forward" accompanied with a pointing gesture that express the requirement for the intention switching. If such a gesture had not been issued, the robot would have moved to a location that is well past the nearby table. Therefore, the robot moved to position 'A_2' by switching the perception to the alternative perception. Then the robot was commanded, "move medium right" without giving a pointing gesture. Therefore, the robot moved to position 'B_2' considering the default intention. Therefore, the task was completed with 2 user instructions. In the event of the system without the MIS, first the command, "move medium left" was issued by the user and the robot moved to location 'a_2'. If the command "move little forward" had been issued in this case, the robot would have moved to a location that is well past the intended moving position due to the limitation of the system (without MIS) and the user already knew this from his past experience. That is the reason for issuing the command "move medium right" instead of "move little forward" similar to the system with the MIS case. Then with the next voice instruction, the robot moved to position 'b_2'. After the next instruction, the robot moved to 'c_2' that is inside the goal area . Therefore, in order to navigate the robot in this situation, three user instructions were required which is higher than for the event with the MIS.

Similarly, the experiments have been carried out in all the layout arrangements by all the participants. The average number of steps required for fulfilling the navigation task in each layout arrangement for the system with the MIS and without the MIS is given in the graph shown in Fig. 9.9.

In all the layout arrangements except 6 and 9, the robot with the MIS was able to be navigated to the goal positions with a fewer number of voice instructions

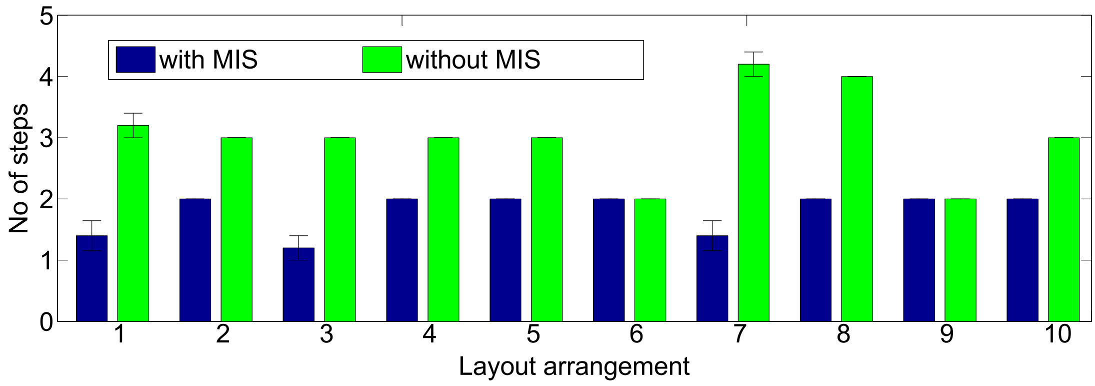


Figure 9.9: This graph shows the average number of steps/instructions taken in order to navigate the robot to the goal positions in different experimental layout arrangements during the experiment for evaluating the performance gain of the proposed MIS. The error bars represent the standard error.

compared to the robot without the MIS and the difference is statistically significant ($P < 0.05$) according to the results of two sample t tests. Moreover, the system with MIS has better abilities in understanding the intention of the user over the system without the MIS. Therefore, the deployment of the MIS enhances the evaluating ability of the ambiguous language instructions by the robot. However, in layout arrangements 6 and 9, the number of steps taken by both the system are the same. The reason behind this was in those two arrangements, the ability of the MIS was not required and the robot was navigated without switching the perception from the default perception. In all other layout arrangements, the intention of the robot was changed only once in each case which leads to a reduction of required total number of steps. Therefore, the number of user instructions or steps required to navigate the robot to a desired location in this kind of situations can be reduced by deploying the MIS. Even though the step number reduction in this kind of task is small (about 1–3 steps), a robot that is used as supportive aid in a caring facility such as a nursing home would be required to perform this sort of navigation task a large number of times per a day and hence there would be a noticeable reduction of the work load in real-world applications. Moreover, this validates the potential of the MIS in enhancing the human-friendliness of the robot and interpretation of ambiguous voice instructions.

9.4 Summary

A method has been introduced in this chapter to enhance the effectiveness of interpretation of verbal instructions with uncertain information such as "move far forward" by identifying the actual intention of the user. The ability for effectively interpreting such voice instructions by a service robot is useful in accomplishing typical daily activities and humanrobot collaborative tasks that involve navigation of the robot. Therefore, the proposed method will improve the abilities of human-friendly service robots.

The main improvement of the proposed method over the existing approaches is that the system is capable of switching the intention of the robot by identifying the actual intention of the user. The actual intention of the user is identified by analyzing the information conveyed from pointing gestures that can be accompanied with voice instructions. Moreover, the interaction ability has been improved by integrating multimodal interaction ability in order to guess the intention of the user for improved interpretation of uncertain information in user instructions.

The intention of the robot is switched by the proposed motion intention switcher (MIS) by altering the perceptive distance from the default to an alternative. The position referred from the pointing gesture and the arrangement of the environment in that scenario are analyzed by the MIS in order to decide the alternative perceptive distance. Moreover the MIS shifts the perception of the robot between the default and the alternative hypotheses based on a set of predefined rules. It would be interesting for future work to consider a probabilistic approach instead of this rule-based approach for intention switching.

Experiments have been carried out in an artificially-created domestic environment in order to analyze the behavior of the proposed MIS. The behavior of the MIS has been found to be effective according to the experimental results. Furthermore, experiments have been carried out in order to evaluate the performance gain of the proposed concept. The experimental results validates the potential of the proposed concept in enhancing the human-friendliness of service robots by effective interpretation of ambiguous voice instructions.

Chapter 10

CONCLUSIONS AND FUTURE WORK

10.1 Conclusions

Voice instructions are often used to convey information between peers in human-human interactions. Accordingly, the capability for human-like voice communication between robots and humans would enhance the overall interaction quality between robots and their users. Systems equipped with human-like voice communication would be able to support users in a friendlier manner. Typically, precise quantitative information is not conveyed through voice instructions, and such voice instructions tend to involuntarily contain imprecise and uncertain terms, lexical symbols and notions, which must be interpreted correctly for a command to be understood. As an example, humans tend to issue commands such as "move a little bit toward the TV" instead of "move 0.5 meters toward the TV". The actual quantitative meanings of uncertain terms such as "close", "near", "little", "far", "small", "large" and "few" are related to spatial information such as the size /length of an item and depend on the environment, the overall context and the perception of the user. Therefore, the ability of a robot assistant to interpret such uncertain information in voice commands and respond appropriately to those commands is crucial for improving human-robot interaction.

A service robot can cope with distance-related uncertain information contained in navigation instructions when the robot's perception of distance-related uncertain information is adapted based on the environmental parameters. Size of the room, available free space of the room, and the arrangement of the environment are the vital environmental parameters for effectively evaluating the uncertain information in motional navigation instructions. For positional commands, size of the room, available free space of the room, size of the reference, and the arrangement of the environment are the vital parameters. Fuzzy inference systems that evaluate these parameters can be used to quantify the uncertain information in language instructions. The robot's perception of distance-related uncertain information can be further improved by establishing human-like abilities in perceiving and interpreting the environment.

A robot's interpretation of uncertain information in navigation instructions can be improved by adapting perception of both direction-related and distance-related uncertain notions in navigation instructions. The directional perception of the robot can be effectively adapted by considering the free space around the robot or the reference. Fuzzy logic can be used to realize this and the output membership function of the fuzzy inference system should be modified according to the distribution of the free space. The navigational command understanding ability of a robot with an adaptable directional perception surpasses that of a robot with fixed directional perception with a significant margin. The user agreement with the actions of the robot is remarkably improved when the directional perception is adapted according to the current environment setting instead of fixed directional perception.

A robot's perception of uncertain information can be adapted toward the user by evaluating the user feedback. This can be realized by fuzzy neural networks that enable the robot to learn from user feedback while simultaneously adapting its perception based on information related to the environment. The performance of a robot with the learning ability surpasses that of a system with no learning ability with a significant margin. The user satisfaction toward the robot's quantification ability of uncertain information vastly enhanced in a robot with learning ability with respect to a robot with no learning ability.

Adapting a robot's perception of uncertain information according to the information conveyed non-verbally can improve the robot's interpretation of uncertain information. Moreover, the quantification effectiveness of uncertain spatial descriptors by the robot can be improved by establishing a multimodal interaction ability for the robot. The error of quantifying the uncertain information can be significantly and remarkably reduced by fusing the information conveyed through pointing gestures. Fuzzy logic can be used to fuse the information conveyed through pointing gestures and the parameters related to the environment. Furthermore, the information conveyed through pointing gestures can be used to resolve the spatial ambiguities and it can significantly reduced the number of steps required to navigate a robot toward a goal.

Commands understanding ability of a robot can be improved by deploying methods to interpret uncertain information related to relative references, since it enables the robot to understand more complex user instructions. Utilization of natural human tendencies improves a robot's perception of uncertain information and subsequently it enhances the human-robot interaction. Fuzzy inference systems and fuzzy neural networks are capable of quantifying the uncertain information in navigation instructions while replicating the natural human tendencies related to the perception of uncertain information.

The effectiveness of interpretation of uncertain information enhances the human-robot interaction. However, the capabilities of the prevailing systems (including previous approaches and the methods proposed in this thesis) are far below the cognitive capabilities of human beings with regard to understanding uncertain information. Therefore, a vast research gap is still remaining in this particular research niche for future developments.

10.2 Future Work

The limitations of the methods proposed in this thesis and the possible improvements are suggested based on the following three aspects; scope, interaction, and adaptation (the same definitions used in chapter 2 are used).

10.2.1 Scope

This thesis addressed the issues in interpreting uncertain information in navigation instructions in relation to distance, direction and relative references. Conversely, the navigation instructions may include uncertain information related to other aspect such as time, speed, shape, and frequency. Ways for interpreting uncertain information in relation such entities have not been addressed in this work. Therefore, the possible extensions of the capabilities of a robot to understand uncertain information in relation to such entities are proposed for the future work.

10.2.2 Interaction

The information conveyed non-verbally is used by the some of the proposed concepts to enhance the interpretation of uncertain information contained in navigation commands. However, only the pointing gestures are analyzed as the sole source of information conveyed non-verbally; the information conveyed from other means such as gaze, facial expressions, and head nodding are not analyzed in the proposed methods of this thesis. The information conveyed from other than pointing gestures could also be used to improve the understanding of uncertain information contained in navigational instructions and subsequently the interaction between the robot and users. For example, the facial expression could be used as a substitute for voice feedback for the learning of the work presented in chapter6. Such usage of notions conveyed non-verbally would reduce the overhead burden on the user and hence improve interaction. Therefore, usage of other notions conveyed non-verbally for enhancing the interaction is proposed for future work.

In all the experimental scenarios presented in this thesis, it is assumed that there is only a single user at a particular instance. Furthermore, it is assumed that the users are in the field of view of the robot when issuing the gesture instructions. If there are multiple users in a given time or the user is not within the field of view, there is an uncertainty in interaction with the user; (e.g., deciding the person who issues the commands or detecting the gesture accurately). However, the scope of the research is limited to addressing the issues related to the uncertain information in language instructions (i.e., interpreting terms like "far" and "little"). Dedicated research wok on handling such issues (e.g., which address only issues such as how to maintaining human-robot proxemics [95]) can be found in the literature and outcomes of such research could be integrated into the work presented in this thesis to resolve these issues.

10.2.3 Adaptation

The work proposed in this thesis is capable of adapting a robot's perception of uncertain information contained in navigation command based on the environment. The robot perceives the environment through available navigation maps and its low-level sensors such as range sensors. Therefore, the robot perceives the environment in 2-dimensions. Moreover, the plan views of the surrounding and the footprints of the objects. The robot calculates the required environmental parameters for adapting the perception based on these inputs. Therefore, there are limitations of the proposed concepts in perceiving the environment effectively in a human-like manner. For example, the DisI uses the footprint size of an object as the size of the reference object when interpreting positional information without considering the height of the object for the estimation of the size of it. In order to rectify these issues and enhance the performance, the robot should be capable of perceiving the environment 3-dimensionally. Moreover, human-like perceiving ability would enhance the performance. This could be achieved by using stereoscopic vision systems and laser scanners to construct 3-dimensional models of the environments that could be used by the robot to extract the environmental parameters. It would be interesting for future work to integrate such abilities for the improvement of the proposed system.

The meaning of uncertain information may depend on the specific awareness about a particular task. As an example, the quantified meaning of the term “near” in a situation where a plastic bottle is moved near to a lighted candle will be different from moving the same bottle near to a glass on top of a dinner table. Since it involves the specific knowledge of the human that the closing the plastic near to a flame is unsafe. However, the work proposed in this thesis is not capable of adapting the perception of uncertain information based on such specific awareness in relation to different objects or tasks. However, the effects caused to the perception due to such specific knowledge about different contexts are minor for navigation. Therefore, further improvement for adaptation of the perception according to such specific knowledge of different tasks is proposed for the future work.

LIST OF PUBLICATIONS ORIGINATED FROM THE THESIS

Refereed Journal Publications

1. **M. A. V. J. Muthugala**, and A. G. B. P. Jayasekara, "A Review of Service Robots Coping With Uncertain Information in Natural Language Instructions," *IEEE Access*, vol. 6, pp. 12913-12928, 2018.

2. **M. A. V. J. Muthugala**, and A. G. B. P. Jayasekara, "Enhancing User Satisfaction by Adapting Robot's Perception of Uncertain Information Based on Environment and User Feedback," *IEEE Access*, vol. 5, pp. 26435–26447, 2017.

3. **M. A. V. J. Muthugala**, P. H. D. A. S. Srimal and A. G. B. P. Jayasekara, "Enhancing Interpretation of Ambiguous Voice Instructions based on Environment and User's Intention for Improved Human Friendly Robot Navigation," *Applied Sciences special issue on Social Robotics*, vol. 7, no. 8, 2017.

4. **M. A. V. J. Muthugala**, and A. G. B. P. Jayasekara, "Improving understanding of navigational commands by adapting robot's directional perception based on environment," [under review]

5. **M. A. V. J. Muthugala**, P. H. D. A. S. Srimal, and A. G. B. P. Jayasekara, "Improving Robot's Perception of Uncertain Spatial Descriptors in Navigational Instructions by Evaluating Influential Gesture Notions," [under review]

Refereed International Conference Publications[1]

1. **M. A. V. J. Muthugala** and A. G. B. P. Jayasekara, "Interpreting Uncertain Information Related to Relative References for Improved Navigational Command Understanding of Service Robots," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vancouver, IEEE, 2017, pp. 6567–6574.

2. **M. A. V. J. Muthugala** and A. G. B. P. Jayasekara, "Interpreting Fuzzy Directional Information in Navigational Commands Based on Arrangement of the Surrounding Environment," in *2017 IEEE International Conference on Fuzzy Systems (FUZZ IEEE)*, Naples, IEEE, 2017, pp. 1-7.

3. **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 *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Daejeon, IEEE, 2016, pp. 1517-1523.

4. **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 *2016 IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, IEEE, 2016, pp. 2915-2921.

5. **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 *2016 Moratuwa Engineering Research Conference (MERCon)*, Moratuwa, IEEE, 2016, pp. 397-402. [**Best paper award**]

[1]Only the first author and accepted/published papers at the time of submission are listed here.

6. **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 *2015 Moratuwa Engineering Research Conference (MERCon)*, Moratuwa, IEEE, 2015, pp. 124-129.

REFERENCES

[1] R. C. Arkin, *Behavior-based robotics.* MIT press, 1998.

[2] S. Koceski and N. Koceska, "Evaluation of an assistive telepresence robot for elderly healthcare," *Journal of Medical Systems*, vol. 40, no. 5, pp. 1–7, 2016. [Online]. Available: http://dx.doi.org/10.1007/s10916-016-0481-x

[3] C. Jayawardena, I. Kuo, E. Broadbent, and B. A. MacDonald, "Socially assistive robot healthbot: Design, implementation, and field trials," *IEEE Syst. J.*, no. 99, pp. 1–12, 2014.

[4] H.-M. Gross, A. Scheidig, K. Debes, E. Einhorn, M. Eisenbach, S. Mueller, T. Schmiedel, T. Q. Trinh, C. Weinrich, T. Wengefeld, A. Bley, and C. Martin, "Roreas: robot coach for walking and orientation training in clinical post-stroke rehabilitation—prototype implementation and evaluation in field trials," *Autonomous Robots*, pp. 1–20, 2016. [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," *Intelligent Service Robotics*, vol. 1, no. 2, pp. 169–183, 2008.

[6] R. Liu and X. Zhang, "Fuzzy context-specific intention inference for robotic caregiving," *International Journal of Advanced Robotic Systems*, 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," *International Journal of Human Computer Interaction (IJHCI)*, 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," *International Journal of Social Robotics*, 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*. 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 *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2017, pp. 53–54.

[13] M. E. Foster, R. Alami, O. Gestranius, O. Lemon, M. Niemelä, J.-M. Odobez, and A. K. Pandey, "The mummer project: Engaging human-robot interaction in real-world public spaces," in *International Conference on Social Robotics*. Springer, 2016, pp. 753–763.

[14] Y. Kim and W. C. Yoon, "Generating task-oriented interactions of service robots," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 8, pp. 981–994, 2014.

[15] M. Mast, M. Burmester, B. Graf, F. Weisshardt, G. Arbeiter, M. Španěl, Z. Materna, P. Smrž, and G. Kronreif, "Design of the human-robot interaction for a semi-autonomous service robot to assist elderly people," in *Ambient Assisted Living*. Springer, 2015, pp. 15–29.

[16] K. Swift-Spong, E. Short, E. Wade, and M. J. Mataric, "Effects of comparative feedback from a socially assistive robot on self-efficacy in post-stroke rehabilitation," in *2015 IEEE International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 2015, pp. 764–769.

[17] S. Schiffer, A. Ferrein, and G. Lakemeyer, "Caesar: an intelligent domestic service robot," *Intelligent Service Robotics*, vol. 5, no. 4, pp. 259–273, 2012.

[18] M. K. Murray, "The nursing shortage: past, present, and future," *Journal of Nursing Administration*, 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." National Health Policy Forum Washington, DC, 2002.

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

[21] P. Vogt, M. De Haas, C. De Jong, P. Baxter, and E. Krahmer, "Child-robot interactions for second language tutoring to preschool children," *Frontiers in Human Neuroscience*, vol. 11, 2017.

[22] C. Piezzo and K. Suzuki, "Feasibility study of a socially assistive humanoid robot for guiding elderly individuals during walking," *Future Internet*, vol. 9, no. 3, p. 30, 2017.

[23] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and autonomous systems*, vol. 42, no. 3, pp. 143–166, 2003.

[24] P. Menzel and F. d'Aluisio, *Robo sapiens: Evolution of a new species.* Mit Press, 2001.

[25] S. Frennert and B. Östlund, "Review: seven matters of concern of social robots and older people," *International Journal of Social Robotics*, vol. 6, no. 2, pp. 299–310, 2014.

[26] S. Kleanthous, C. Christophorou, C. Tsiourti, C. Dantas, R. Wintjens, G. Samaras, and E. Christodoulou, "Analysis of elderly users preferences and expectations on service robots personality, appearance and interaction," in *International Conference on Human Aspects of IT for the Aged Population.* Springer, 2016, pp. 35–44.

[27] S. Huang, T. Tanioka, R. Locsin, M. Parker, and O. Masory, "Functions of a caring robot in nursing," in *2011 7th Int. Conf. Natural Language Processing and Knowledge Engineering*, 2011, pp. 425–429.

[28] N. Mavridis, "A review of verbal and non-verbal human–robot interactive communication," *Robotics and Autonomous Systems*, vol. 63, pp. 22–35, 2015.

[29] 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," *Personal and Ubiquitous Computing*, vol. 17, no. 1, pp. 127–144, 2013.

[30] J. A. Ansari, A. Sathyamurthy, and R. Balasubramanyam, "An open voice command interface kit," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 3, pp. 467–473, 2016.

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

[32] K. Sugiura and K. Zettsu, "Rospeex: A cloud robotics platform for human-robot spoken dialogues," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sept 2015, pp. 6155–6160.

[33] M. Mazo, F. J. Rodriguez, J. L. Lázaro, J. Ureña, J. C. Garcia, E. Santiso, P. Revenga, and J. J. Garcia, "Wheelchair for physically disabled people with voice, ultrasonic and infrared sensor control," *Autonomous Robots*, vol. 2, no. 3, pp. 203–224, 1995.

[34] J. Clark and R. Roemer, "Voice controlled wheelchair." *Archives of physical medicine and rehabilitation*, vol. 58, no. 4, pp. 169–175, 1977.

[35] K. B. Stanton, P. R. Sherman, M. L. Rohwedder, C. P. Fleskes, D. R. Gray, D. T. Minh, C. Espinoza, D. Mayui, M. Ishaque, and M. A. Perkowski, "Psubot-a voice-controlled wheelchair for the handicapped," in *Circuits and Systems, 1990., Proceedings of the 33rd Midwest Symposium on.* IEEE, 1990, pp. 669–672.

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

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

[38] 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.* Cham: Springer International Publishing, 2013, pp. 501–510.

[39] E. Bastianelli, G. Castellucci, D. Croce, R. Basili, and D. Nardi, "Effective and robust natural language understanding for human-robot interaction." in *ECAI*, 2014, pp. 57–62.

[40] K.-P. Gapp, "Basic meanings of spatial computation and evaluation in 3d space," in *Proceedings AAAI-94*, 1994, pp. 1393–1398.

[41] 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 *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics.* Association for Computational Linguistics, 2006, pp. 745–752.

[42] S. Guadarrama, L. Riano, D. Golland, D. Go, Y. Jia, D. Klein, P. Abbeel, T. Darrell, *et al.*, "Grounding spatial relations for human-robot interaction," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems.* IEEE, 2013, pp. 1640–1647.

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

[44] B. Talbot, O. Lam, R. Schulz, F. Dayoub, B. Upcroft, and G. Wyeth, "Find my office: Navigating real space from semantic descriptions," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, May 2016, pp. 5782–5787.

[45] I. Zukerman, S. N. Kim, T. Kleinbauer, and M. Moshtaghi, "Employing distance-based semantics to interpret spoken referring expressions," *Computer Speech & Language*, 2015.

[46] D. Whitney, M. Eldon, J. Oberlin, and S. Tellex, "Interpreting multimodal referring expressions in real time," in *2016 IEEE International Conference on Robotics and Automation (ICRA).* IEEE, 2016, pp. 3331–3338.

[47] H. Admoni, T. Weng, and B. Scassellati, "Modeling communicative behaviors for object references in human-robot interaction," in *2016 IEEE Inter-*

national Conference on Robotics and Automation (ICRA). IEEE, 2016, pp. 3352–3359.

[48] N. Kubota and Y. Toda, "Multimodal communication for human-friendly robot partners in informationally structured space," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1142–1151, 2012.

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

[50] K. Lynch, *The image of the city*. MIT press, 1960, vol. 11.

[51] R. Briggs, "Urban cognitive distance," *Image and environment: Cognitive mapping and spatial behavior*, pp. 361–388, 1973.

[52] D. R. Montello, "The measurement of cognitive distance: Methods and construct validity," *Journal of Environmental Psychology*, vol. 11, no. 2, pp. 101–122, 1991.

[53] S. Dutta, "Approximate spatial reasoning: integrating qualitative and quantitative constraints," *International journal of approximate reasoning*, vol. 5, no. 3, pp. 307–330, 1991.

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

[55] M. Skubic, D. Perzanowski, S. Blisard, A. Schultz, W. Adams, M. Bugajska, and D. Brock, "Spatial language for human-robot dialogs," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 34, no. 2, pp. 154–167, 2004.

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

[57] 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, 2004.

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

[59] K. Kawamura, S. Bagchi, and T. Park, "An intelligent robotic aid system for human services," in *NASA CONFERENCE PUBLICATION*. NASA, 1994, pp. 413–413.

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

[61] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE transactions on information theory*, vol. 8, no. 2, pp. 179–187, 1962.

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

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

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

[65] S. Schiffer, A. Ferrein, and G. Lakemeyer, "Reasoning with qualitative positional information for domestic domains in the situation calculus," *Journal of Intelligent & Robotic Systems*, vol. 66, no. 1-2, pp. 273–300, 2012.

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

[67] C. Jayawardena, K. Watanabe, and K. Izumi, "Posture control of robot manipulators with fuzzy voice commands using a fuzzy coach–player system," *Advanced Robotics*, vol. 21, no. 3-4, pp. 293–328, 2007.

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

[69] K. Watanabe, C. Jayawardena, and K. Izumi, "Approximate decision making by natural language commands for robots," in *32nd Annual Conference on IEEE Industrial Electronics (IECON 2006)*, Nov 2006, pp. 4480–4485.

[70] 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," *Journal of System Design and Dynamics*, vol. 4, no. 2, pp. 297–313, 2010.

[71] 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 *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2008, pp. 2016–2021.

[72] G. Hesslow, "Conscious thought as simulation of behaviour and perception," *Trends in cognitive sciences*, vol. 6, no. 6, pp. 242–247, 2002.

[73] C.-T. Lin and M.-C. Kan, "Adaptive fuzzy command acquisition with reinforcement learning," *Fuzzy Systems, IEEE Transactions on*, vol. 6, no. 1, pp. 102–121, 1998.

[74] A. B. P. Jayasekara, K. Watanabe, K. Kiguchi, and K. Izumi, "Adaptation of robots perception of fuzzy linguistic information by evaluating vocal cues for controlling a robot manipulator," *Artificial Life and Robotics*, vol. 15, no. 1, pp. 5–9, 2010.

[75] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, "Neuro-fuzzy and soft computing; a computational approach to learning and machine intelligence," 1997.

[76] 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 *2010 IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2010, pp. 778–783.

[77] M. W. Eysenck and M. T. Keane, *Cognitive psychology: A student's handbook*. Taylor & Francis, 2000.

[78] K. Izumi, A. Jayasekara, K. Watanabe, and K. Kiguchi, "Attentive and corrective feedback for adapting robot's perception on fuzzy linguistic information," in *SICE Annual Conference 2010, Proceedings of*. IEEE, 2010, pp. 1963–1968.

[79] 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 *7th International Conference on Information and Automation for Sustainability*, Dec 2014, pp. 1–6.

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

[81] I.-H. Kuo and C. Jayawardena, "Application of fuzzy techniques in human-robot interaction-a review," in *Social Robotics*. Springer, 2014, pp. 249–255.

[82] E. Dadios, A. Valencia, D. Johnson, J. Biliran, and R. Garcia, *Humanoid Robot: Design and Fuzzy Logic Control Technique for Its Intelligent Behaviors.* INTECH Open Access Publisher, 2012. [Online]. Available: https://books.google.lk/books?id=ILmHoAEACAAJ

[83] J. M. Mendel, "Computing with words: Zadeh, turing, popper and occam," *IEEE Computational Intelligence Magazine*, vol. 2, no. 4, pp. 10–17, 2007.

[84] J. M. Mendel, "General type-2 fuzzy logic systems made simple: a tutorial," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 5, pp. 1162–1182, 2014.

[85] C. L. Bethel and R. R. Murphy, "Review of human studies methods in hri and recommendations," *International Journal of Social Robotics*, vol. 2, no. 4, pp. 347–359, 2010.

[86] A. Hunt and A. Hunt, "Jspeech grammar format," *W3C Note, June*, 2000.

[87] J. Tan, Z. Ju, and H. Liu, "Grounding spatial relations in natural language by fuzzy representation for human-robot interaction," in *2014 IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE).* IEEE, 2014, pp. 1743–1750.

[88] P. D. Ellis, *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results.* Cambridge University Press, 2010.

[89] M. Pfadenhauer and C. Dukat, "Robot caregiver or robot-supported caregiving?" *International Journal of Social Robotics*, vol. 7, no. 3, pp. 393–406, 2015.

[90] A. G. B. P. Jayasekara, K. Watanabe, K. Kiguchi, and K. Izumi, "Adaptation of robot behaviors toward user perception on fuzzy linguistic information by fuzzy voice feedback," in *18th IEEE Int. Symp. Robot and Human Interactive Communication*, 2009, pp. 395–400.

[91] P. H. D. A. S. Srimal, M. A. V. J. Muthugala, and A. G. B. P. Jayasekara, "Identifying spatial terminology and boundaries for human robot interaction:

A human study," in *2017 Moratuwa Engineering Research Conf. (MERCon)*, 2017, pp. 46–51.

[92] C. Matuszek, L. Bo, L. Zettlemoyer, and D. Fox, "Learning from unscripted deictic gesture and language for human-robot interactions." in *28th AAAI Conf. Artificial Intelligence*, 2014, pp. 2556–2563.

[93] P. H. D. A. S. Srimal, "Development of an intelligent vision enhanced multi-modal human-robot interaction for service robots in domestic environment," Master's thesis, University of Moratuwa, 2017.

[94] A. B. P. Jayasekara, K. Watanabe, M. K. Habib, and K. Izumi, "Visual evaluation and fuzzy voice commands for controlling a robot manipulator," *International Journal of Mechatronics and Manufacturing Systems*, vol. 3, no. 3-4, pp. 244–260, 2010.

[95] R. Mead and M. J. Matarić, "Autonomous human–robot proxemics: socially aware navigation based on interaction potential," *Autonomous Robots*, vol. 41, no. 5, pp. 1189–1201, 2017.