UNIVERSITA' DEGLI STUDI DI PALERMO
DIPARTIMENTO DI INGEGNERIA CHIMICA, GESTIONALE, INFORMATICA, MECCANICA
Dottorato in Ingegneria Informatica

INVESTIGATING PERCEPTUAL FEATURES FOR A NATURAL HUMAN HUMANOID ROBOT INTERACTION INSIDE A SPONTANEOUS SETTING

Settore scientifico disciplinare ING-INF/05

TESI DI
ING. MARCELLO EMANUELE GIARDINA

COORDINATORE DEL DOTTORATO
PROF. SALVATORE GAGLIO

TUTOR
PROF. ROSARIO SORBELLO

XXV CICLO ANNO ACCADEMICO 2013-2014

DOTTORATO

To my family.

Sommario Breve

Dal momento che i robot umanoidi sono destinati a divenire parte integrante della vita degli esseri umani, da qualche anno a questa parte si sono resi necessari studi appositi volti ad indagare oltre che gli aspetti puramente tecnici ingegneristici anche gli aspetti sociali e collaborativi legati all'interazione uomo-robot.

Immaginiamo di avere un ambiente comune, dove un uomo e un robot possono dialogare empaticamente condividendo il "senso di co-presenza" durante lo svolgimento dei normali compiti di tutti i giorni. E, immaginiamo, dei robot dalle sembianze umane che siano in grado di eseguire dei compiti altamente complessi, che possano apprendere, giorno per giorno, dall'ambiente esterno, cosi come avviene fin dai primi mesi di vita dei bambini, dei robot che possano esibire un interazione naturale ed ampatica con gli esseri umani, sostenendo una conversazione pseudo-realistica ed esprimendo emozioni anche attraverso il linguaggio del corpo. Bene, questa tesi di dottorato vuole dare un prezioso contributo a questo ambizioso obbiettivo. E' stato condotto, pertanto, uno studio sociologico e psicologico (200 persone interessate) volto ad indagare le caratteristiche collaborative e sociali (Acceptance and Believabilty) connesse all'interazione tra uomo e umanoide. E su questo tipo di studi é stato possibile progettare un complesso sistema robotico intelligente, in grado, di sostenere un dialogo, i processi d'interazione e i meccanismi di feedback delle persone, e di incoraggiare la cooperazione tra gli uomini e le macchine supportandone le esigenze sociali e pratiche.

In particolare, stimolato dai promettenti risultati ottenuti, grazie alla collaborazioni del RoboticsLab del dip.DICGIM dell'Universitá degli Studi di Palermo, con il Centro Diagnosi e Terapia delle Sindromi Autistiche dell'ASP 6 della cittá di Palermo, il centro SLA dell'Azienda Ospedaliera Universitaria Policlinico di Palermo, e l'IshiguroLab dell'Osaka University, Osaka, in Giappone, e il Rotary Club Palermo Est, é stato possibile studiare, e condurre dei test sull'interazione tra il robot umanoide Telenoid, sia con soggetti che presentano particolari svantaggi, avendo disabilitá mentali e psichiche o motorie, come gli anziani, i bambini autistici, e i pazienti affetti da SLA, sia con persone normali che non presentano evidenti deficit cognitivi. Nei capitoli seguenti verranno, dunque, descritti, i fondamenti teorici, le metodologie, le tecnologie e gli esperimenti condotti sul campo.

Abstract

Since robots have become part of human life, several studies have been done with the aim of discovering salient social rules at the basis of the collaboration between human beings and humanoid robot (*Human Humanoid Interaction - HHI*). The purpose is to have a common environment where human and humanoid robot could engage a proficous "dialogue" in order to share the "sense of co-presence" for common empathic tasks and goals. The humanoid robot must be in a position to interact with the human, and learn day by day from external environment, exactly as it occurs in human beings in the real life. In this sense, significant progresses have been reached, having a strong impact in each aspect of every day life. In this way, is a big challenge having a humanoid robot that can exhibit a natural and empathic interaction with human beings, supporting a pseudo-realistic conversation and expressing human-like body emotions.

To reach this ambitious goal, it was essentially conducted a sociological and psychological study (Sorbello et al., 2014) oriented to investigating collaborative and social features related to human robot interaction. So, supported by the results of this study it was possible to base all assumptions that support the design of the robust robotic architecture in order to reach the purposes of this thesis. The Sociological and Psychological study aims at a descriptive analysis of the main perceptual and social features of natural conditions of agents interaction, which can be specified by agents in human-humanoid robot interaction. A principled approach to human-robot interaction may be assumed to comply with the natural conditions of agents overt perceptual and social behaviour. To validate this research a minimalistic humanoid robot Telenoid it was used and human-robot interactions test have been conducted with two hundred people with no prior interaction experience with robot.

After the definition of experimental conditions and setup, an analysis of significant variance correlation among dimensions in ordinary and goal guided contexts of interaction has been performed in order to prove that *Perception* and *Believability* are indicators of social interaction, that can increase the degree of interaction in human-humanoid interaction.

I found that the sense of a shared environment is substantive for obtaining

a satisfying interaction, where the distances between agents are constructed as perceptual-motor proxies of the regions where intentions and actions are available at a glance.

The experimental results highlight that the Perceptual and Believability, as implicit social competences, could improve the meaningfulness and the natural-like sense of human-humanoid interaction in every day-life task-driven activities and the Telenoid robot is perceived as an autonomous cooperative agent for a shared environment with human beings.

One crucial result of my doctoral research is that subjects are significantly inclined to perceive the Telenoid as cooperative and competent even though this favorableness. Since the users perceived the behaviour of the Telenoid coherent and consistent, my results suggest that they are favourably incline to accept the humanoids behaviour as tuned to theirs and moved by the commitment to meet their demands.

The idea of social robots, able to engage users for extended periods of time, received great attention from researchers in the recent years. In particular, in HHI, empathy is a important key in order to overcome the current limitations of social robots. In facts, a main defining characteristic of human social behavior is empathy. I support the hypothesis that the robots, in order to become personal companions or partners of our daily life in the near future, they need to know how to interact with us by means of empathy.

One of the subgoal of this thesis is to overcome the state of loneliness of elderly people (understanding in which emphatic state is based the current HHI) using this minimalistic humanoid robot capable to exhibit a dialogue similar to what usually happens in real life between human beings.

To validate my research, I have designed and developed a robotic architecture using the minimalistic humanoid robot Telenoid and I have conducted human-robot interactions tests with elderly people with no prior interaction experience with robot. During the experiments, elderly persons engaged a stimulated conversation with the humanoid robot.

In order to engage a sort of empathic state in HHI, is necessary to catch the emotions that arise by conversation, so that, the humanoid robot can be able to interact socially and naturally with a human by expressing human-like body emotions in according with his interlocutor. The emotional architecture designed is based on an emotional conceptual space generated using the paradigm of Latent Semantic Analysis. The robot generates its overall affective behavior (*Latent Semantic Behavior*) taking into account the visual and phrasal stimuli of human user, the environment and his "personality", all encoded in his emotional conceptual space. The robot determines his emotion in according by all these parameters that influence and orient the generation of his behavior not predictable from the user. The goal of this approach is to obtain an affinity matching with humans.

The robot, in this way, can exhibit a smoothly natural transition in his emotion, changing during the interaction with humans taking, also, into account the previous generated emotions. To validate the system, I implemented a distribute system on an Telenoid robot and on a Google Android Phone and I tested this social emotional interaction using the phone device as intelligent interface between human and robot in a complex scenario.

The favourable acceptance of Telenoid behaviors by human users had oriented the research to the use of this android robot as:

- Therapist on the processes of Social Integration among Children with DSA and their classmates. In particular I am interested in the study related to the sense of a person to be present in a remote environment with a robot ("*Telepresence*") and to the sense of a child to be present in a common environment with a robot where humans and humanoid are "*accessible, available and subject to one another*" (Goffman, 1966).The sense of "*togetherness*" between persons and Humanoid is "*inherently social*" (Biocca, 1992) and is highly connected with the concepts of particular behavior defined "*sensible*" (Chella et al., 2011a) because capable to express cognitive functionality.

  Priority goals of the study are: to gain greater flexibility and expressivity of mime and gaze in social interaction, learning the shifts and exchange in reciprocity, the refinement of verbal interaction and non-verbal, acquisition of skills and insights to make inferences about to view of the other, the possibility of developing the semantics and pragmatics of language areas of major weakness. The individual profile, of autistic child, make possible the use of autonomous robots with a remote control by the therapist. In order to test the relevance and the goals, I conducted experiments with 2 autistic children, 23 classmates and telenoid and nao robots.

- Co-therapist for the parents of children with autism for the acceptance of diversity. Many existing research studies have shown good results relating to the important impact of Acceptance and Commitment Therapy (ACT) (Prevedini et al., 2011) applied to parents of children with autism. The ACT has in Relation Frame Theory (RFT) its theoretical foundation. The overall behaviors of the parents may potentially benefit from treatment with a humanoid robot therapist instead of a real one. In particular in the present study, Telenoid humanoid robot (Ishiguro et al., 2013) is used as therapist to achieve a specific therapeutic objective: the acceptance of diversity from the parents of children with autism. Experimental results has been conducted adapting Hexaflex model of ACT protocol to Telenoid Robot in the therapy with parents. I conducted the preliminary test with 2 parents of autistic children and the humanoid robot Telenoid.

- "Alter Ego"and "Partner" of the Locked-In patients affected by Amyotrophic Lateral Sclerosis through the use of devices BCI (Brain Computer Interface). In consideration of the undamaged cognitive status in most ALS patients and the increasing improvement of humanoid robot technology, I hypothesize that a dedicated BCI robotic system might give an unprecedented contribute to the management of these patients. In particular, I do expect that a humanoid robot may act in different dimensions as motor substitution, virtual moving, shooting and emotional interaction and communications proxy interlocutor. I conducted the preliminary test with 2 locked-in patients and the Telenoid and nao robots (Experiments sessions have been supervised by psychologists) and I want to verify with this preliminary test whether this approach might improves the quality of life and disease status acceptance of patients and caregivers.

Contents

List of Figures

Glossary

Corpus	in linguistics, a large and structured set of texts.
Lemma	In morphology and lexicography, a lemma (plural lemmas or lemmata) is the canonical form, dictionary form, or citation form of a set of words.
Morpheme	a meaningful morphological unit of a language that cannot be further divided.
Morphology	in linguistics, morphology is the identification, analysis, and description of the structure of a given language's morphemes and other linguistic units, such as root words, affixes, parts of speech, intonations and stresses, or implied context.
Part of Speech	a category to which a word is assigned in accordance with its syntactic functions. In English the main parts of speech are noun, pronoun, adjective, determiner, verb, adverb, preposition, conjunction, and interjection.
Polysemy	the coexistence of many possible meanings for a word or phrase.
Semantics	the branch of linguistics and logic concerned with meaning. There are a number of branches and subbranches of semantics, including formal semantics, which studies the logical aspects of meaning, such as sense, reference, implication, and logical form, lexical semantics, which studies word meanings and word relations, and conceptual semantics, which studies the cognitive structure of meaning..
Syntagm	a linguistic unit consisting of a set of linguistic forms (phonemes, words, or phrases) that are in a sequential relationship to one another.
Syntax	In linguistics, is the study of the principles and processes by which sentences are constructed in particular languages.
Vector space model	is an algebraic model for representing text documents (and any objects, in general) as vectors of identifiers, such as, for example, index terms. It is used in information filtering, information retrieval,indexing and relevancy rankings. Its first use was in the SMART Information Retrieval System.

Chapter 1

Introduction

The *Human-Humanoid Robot Interaction* is part of a multidisciplinary research area recently born, in constant development, full of ideas for advanced research and technology transfers from several industrial and technologists sectors. It plays a vital role in the creation of robots that can operate in open environments and can cooperate with humans. This type of tasks requires the design and the development of interactive systems that are usable, reliable, and that can support and facilitate human activities in order to help, inexperienced users, to use their robot easily and safely through natural and intuitive interfaces. The purpose of this thesis is the development and testing of innovative approaches, that can facilitate the interaction between the Robot and the Humans Beings in the complex process and tasks performed in real environment. Since humanoid robots are destined to become an integral part of the human beings lives, it is, always more, important to focus the researcher's studies, not only, on specific technical aspects of engineering, but also, on the social aspects related to the interaction and the collaboration between Human and Robot. The important questions that are posed to researchers are: The human can interact with the humanoid ? and it is possible have trust to their ? And, finally, How the robot can be partner and human avatar in the real life ? Part of this thesis work is aimed to investigate this hard themes, searching to give them, valid answers supported by theoretical assumptions and real experiments. With the advancement in the field of robotic technologies, therefore, researchers are pushing more and more in there way, trying to explore areas of robotics and artificial intelligence that can perform highly complex tasks, to the point of having robots with human features they can take the "pseudo realistic conversations" that will have in the future, now coming, an important role in the life of every day. Now, the robots are beginning to play a role so advanced in our lives, such as,in Health Care (robot to service of ALS Patients) such as Robot to service of Autistic Children, such as robot to service to siblings and parents of autistic children or in the 'context of' assistance for the disabled and elderly people, so, psychological aspects, such as the Trust human-robot and the ability Manipulation of the robot itself, begin to take on connotations increasingly important, becoming themselves object of research (NISHIO and ISHIGURO, 2011) (Ishiguro et al., 2013) (Ishiguro et al., 2012) (Sorbello et al., 2014)

1.1 Motivation and Goals

Since robots have become part of human life, several studies have been done with the aim of discovering salient social rules at the basis of the collaboration between human beings and humanoid robot (Human Humanoid Interface - HHI)(Sorbello et al., 2014), (Pereira et al., 2011), (Kozima et al., 2004), (Hegel et al., 2006), (Misselhorn, 2009), (Tapus and Mataric, 2007). HHI has the purpose of creating a hybrid common environment in which human beings and humanoids robot will be capable to share the most common operation. The Human Humanoid Interaction (HHI) needs to be more narrowed to the idea of having a common environment where human and humanoid robot could engage a proficous "dialogue" in order to share the "sense of co-presence" for common empathic tasks and goals (Balistreri et al., 2011a) (Ishiguro et al., 2012) (Ishiguro et al., 2013). It has become indispensable that the human-humanoid communication is as efficient and realistic as possible. In contrast to other robots which are in other environments and are able to perform only some basic operations in a specific reality, the humanoid must be in a position to interact with the human and to update its own knowledge day by day; Learning through natural language or through the use of texts and self-learning, as well as it happens, in the real life. Today the interaction of the robot with the human is among one of the great challenges which the Artificial Intelligence has defined for the near future (Yamazaki et al., 2012). Significant progress in this sense will have a strong impact in each aspect of everyday life (Straub et al., 2012) (Iocchi et al., 2013). In this way, is a big challenge have Humanoid Robots that can exhibith a natural and empathic interaction with human beings, supporting a pseudo-realistic conversation and expressing human-like body emotions. Right now, empathic features have represented, strong limitation of social robot, so many studies are oriented to investigating that overcome thi sort of gap. Rather than considering empathy as a mono-dimensional construct I started from idea of Davis (Davis, 1983) where empathy can be described as a set of construct consisting of four main sub-scales: Perspective-Taking(PT), Fantasy (FS), Empathic Concern(EC) and Personal Distress(PD). Using this definion of empathy, I considered, also, very interesting the role of insular cortex described by (Carr et al., 2003). In particular they underlined that the insular cortex allows us the understanding of the feeling of others using an action representation that generates empathy and tunes our emotional contents. In particular the Emotional competence is one of main key of the minimal agency of a robot where the Biologically Inspired Cognitive Architecture (BICA) community wants to invest their efforts (Chella et al., 2011a). An emotional BICA agent need to exibit the capability (Breazeal, 2003a) to capture and feels about human emotions during interaction with a person if we want to accept him in people society(Oztop et al., 2005)(Fellous, 2004). Emotional equipment is nowadays a key condition of minimal agency for an artificial agent like a humanoid robot(Menegatti et al., 2008). It is important underline, that Emotions (Thagard and Shelley, 2001),(Arnold, 1960) improves the capability of humanoid robots(Monceaux et al., 2009) and allow their behaviors (Xie et al., 2010) to be effective in the interaction with humans(Hudlicka, 2007).So, a good way it is equipping a humanoid robot with emotional competence(Chella et al., 2011c) during their interaction with humans. In order to reach the big challenge of have Humanoid Robots that can exhibit a natural and empathic interaction with human beings, supporting a pseudorealistic conversation and expressing human-like body emotions, was essentially to conduct a sociological and psicological study (Ishiguro et al., 2012) oriented to investigating collaborative and social features related to human robot interaction. So, on the results of this study it was possible to base all assumptions and the design of robotic architectures in order to reach the purposes of this thesis. This study aims at a descriptive analysis of the main perceptual and social features of natural conditions of agents interaction, which can be specified by agents in human-humanoid robot interaction. A principled approach to human-robot interaction may be assumed to comply with the natural conditions of agents overt perceptual and social behaviour. To conduct my research, in order to understanding areas of human cognition, which have not been tested or clarified until now, I used the Telenoid robot shown on figure 2.3. The telenoid robot is a humanoid robot, designed to appear and behave like a minimalistic human.It was created following processes of removing as many unnecessary features, as possible, choosing the necessary features for communication from humans and discarding the unrelated ones. The result is a robot endowed with some of perceptual and motor features of overt behaviour synchronized to speech, head and arms movement of a human agent through a Teleoperated system that allows to study such perceptually accessible features as meaningful clues for social interaction. Thanks to this kind of robot, it is possible to understand how some features of perceptual behaviour work, such as distance and relative positions of agents, face regions spotted as highly informative about emotion or intention reading, the degree at which the space of the interaction appears to be a shared environment; To assess the degree of Believability of interaction along dimensions that can be reasonably taken as meaningful indicators of social interaction, both in free and task directed conditions. In order to prove the goals, I have conducted a human robots interactions test with people who did not have prior interaction experience with humanoid robots, though it was not excluded that they possessed informations or informal notions of IA and robotics. After every stage of interaction, each subject was asked to fill up a questionnaire whose questions were finalized to retrieve information about the salient perceptual and social dimensions of the interaction. Given the data analysis performed, it is possible identify two interaction dimensions, that are the *Perceptual* and *Believability* behaviour. This two dimension can serve for either assess the perceptual and observable behaviour conditions of an humanoid agent, or to increase the natural-looking-like of interaction behaviour in human-humanoid interaction. Supported by the good results obtained by the study on the Acceptance and believability of the Telenoid robot, I was interested to investigate the possibility to have a humanoid robot that is both *avatar* and *partner*, of subject with disadvantage, having mental and psychic disability, such as Elderly people, Autistic Children, Autistic young adult and ALS patients, and with normal people but that live every day with the problematic of autism as parents of autistic children, in order to understand how the robot can help this people and how design the complex mechanism that support the robotic architecture for reach this kind of goals.

The following Chapters present the case study of Applied Research in the real life and the obtained results. Then the chapters proceed presenting the domain of social interactions with a description of a protocol of interaction and then the related problem detected and dealt during all the sessions with the users.

1.2 Contributions

During the Ph.D course my contributions are summarized as :

- I focused on the analysis of two social dimensions at the base of the natural behaviors between people and robots that cover the perceptual features during a normal interaction: the Perception and Believability. In this study I have conducted a series of important experiments (about 200 participants) at the Department DICGIM University of Palermo. In particular, thanks to the many collaborations enjoyed by the robotics lab of University of Palermo, I have got the possibility to bring my research on various social categories with cognitive problems, ranging from the elderly, autistic children, young adults with autism, siblings and parents of autistic children and patients with amyotrophic lateral sclerosis.

- I collaborated in the research project " Geminoid Telenoid, Maximum and Minimun of Human Design" coordinated by prof. Hiroshi Ishiguro of Osaka University, Japan and Dr. Suishi Nishio Director of "Ishiguro Lab" at the Advanced Telecommunications Research Institute International, ATR Kyoto, Japan. In particular my research has focused on the way to have an autonomous humanoid robot that would be able to support a simple dialogue with elderly people. After a preliminary study of conversations between elders and nephews in search of some kind of repetitive pattern inside to the conversation that could be analyzed and reproducible, I worked to design and develop a system of "Topic Recognition", for the recognition of 'topic of the conversation, a system of "Emotion Recognition", in order to equip the robot of emotional behavior spontaneous and not repetitive in response to the stimuli coming from its interlocutor and the external environment and in order to make the conversation emphatically in line with a kind of emotional threshold, a system of "Learning", in order to equip the robot of a system that can permits to learn directly from its interlocutor, all the arguments not known (such as occurs in children) and a system of "Question Answering" able to hold a simple conversation. In order to achieve the goals of the research I have designed a modular robotics architecture with interfaces that allows: each modules communicate with each other and add, modify or replace them (if necessary) without that the entire infrastructure can be affected changes, continuing to operate with maximum efficiency. A modular structure so designed, made it possible to divide the overall problem into sub-problems, permitting of consider each module as a sub-project with its programming language, with its computational load and its processes of development. I attended in international meetings, taking an active part in test sessions with the robot Telenoid inside of "for Culture and Society Research Programme for Philosophy and Intellectual Ideas" at Aarhus University Institute in Denmark and in international meetings and session test at the Advanced Telecommunications Research Institute International, ATR, Kyoto, Japan, and at the Osaka University, Japan.

- I participated in the pilot project "The Humanoid Robot as Communicative Mediator for Social Integration of Autistic Children", thanks a collaboration of RoboticsLab, Dicgim, University of Palermo, the ASP 6 (local healthcare unit) Palermo, Palermo Guglielmo Marconi Secondary School and Rotary Club Palermo East, which has as purpose to trigger the interaction between the child and the other children with ASD using a humanoid robot as Socio - Communicative Mediator in order to promote the integration (long-term) in a social group. I conducted my research in the area related to the methodologies, tools and techniques at support of the objectives of the project as well as in relational on interpersonal dynamics and communication between the child with ASD and other children. I designed and developed an architecture in Cloud network for remote control of a humanoid robot through mobile devices and I designed and developed an architecture for making (very user friendly wizard) of complex robotic behaviors usable by children with ASD.

- I participated in the pilot project "Brain Computer and Robotics in disease management for patients with ALS (Amyotrophic Lateral Sclerosis)" thanks a collaboration of RoboticsLab, dip. Dicgim, and the dip. Experimental Biomedicine and Neuroscience, University of Palermo, which has as its purpose the use of humanoid robots, as alter ego of caregivers of Locked-In patients suffering from Amyotrophic Lateral Sclerosis, through the use of BCI (Brain Computer Interface) devices. I focused on the concept and design of the architecture to support the objectives of the project, the optimization of the chain of information flow signals from the BCI device to a central control system and the development of an architecture in Cloud network for remote control of a Humanoid Robot through the aid of devices BCI for the communication.

1.3 Pubblications

The work presented in this thesis resulted in the publication of the following research articles:

- Antonio Chella, Rosario Sorbello, Giovanni Pilato, Giuseppe Balistreri, Salvatore Maria Anzalone and Marcello Giardina: "An Innovative Mobile Phone Based System For Humanoid Robot Expressing Emotions And Personality", International Conference of Biologically Inspired Cognitive Architectures (BICA 2011), Proceedings of the Second Annual Meeting of the BICA Society Journal of Frontiers in Artificial Intelligence and Applications (FAIA), IOS Press Editor.

- Antonio Chella, Rosario Sorbello, Giovanni Pilato, Giorgio Vassallo,Giuseppe Balistreri and Marcello Giardina. "An Architecture with a Mobile Phone Interface for the Interaction of a Human with a Humanoid Robot Expressing Emotions and Personality", XIIth International Conference of the Italian Association for Artificial Intelligence (AI*IA 2011) Palermo, Italy, September 15-17, 2011. Proceedings of Journal Lecture Note in Artificial Intelligence, LNAI n 6934, pp. 117-126, 2011, Springer-Verlag Berlin Heidelberg Editor.

- Hiroshi Ishiguro, Shuichi Nishio, Antonio Chella, Rosario Sorbello, Giuseppe Balistreri, Marcello Giardina, Carmelo Calí: "Investigating Perceptual Features for a Natural Human - Humanoid Robot Interaction Inside a Spontaneous Setting", Biologically Inspired Cognitive Architectures 2012, Advances in Intelligent Systems and Computing Volume 196, 2013, pp 167-174, Springer Editor.

- Antonio Chella, Rosario Sorbello, Giovanni Pilato, Giorgio Vassallo, Marcello Giardina: "A New Humanoid Architecture for Social Interaction between Human and a Robot Expressing Human-Like Emotions Using an Android Mobile Device as Interface", Biologically Inspired Cognitive Architectures 2012, Advances in Intelligent Systems and Computing Volume 196, 2013, pp 95-103, Springer Editor.

- Hiroshi Ishiguro, Shuichi Nishio, Antonio Chella, Rosario Sorbello, Giuseppe Balistreri, Marcello Giardina, Carmelo Calí: "Perceptual Social Dimensions of Human - Humanoid Robot Interaction", Intelligent Autonomous Systems 12, Advances in Intelligent Systems and Computing Volume 194, 2013, pp 409-421, Springer Editor.

- Luca Iocchi, Emanuele Menegatti, Andrea Bonarini, Matteo Matteucci, Enrico Pagello, Luigia Carlucci Aiello, Daniele Nardi, Fulvio Mastrogiovanni, Antonio Sgorbissa, Renato Zaccaria, Rosario Sorbello, Antonio Chella, Marcello Giardina, Primo Zingaretti, Emanuele Frontoni, Adriano Mancini, Grazia Cicirelli, Alessandro Farinelli, Domenico G. Sorrenti, "Development of Intelligent Service Robots" Intelligenza Artificiale, Volume 7, n 2/2013, pp.139-152, DOI: 10.3233/IA-130055, IOS Press.

- Rosario Sorbello, Hiroshi Ishiguro, Antonio Chella, Shuichi Nishio, Giovan Battista Presti, Marcello Giardina: "Telenoid mediated ACT Protocol to Increase Acceptance of Disease among Siblings of Autistic Children", In HRI2013 Workshop on Design of Humanlikeness in HRI : from uncanny valley to minimal design, Tokyo, Japan, pp. 26, March, 2013.

- Rosario Sorbello, Antonio Chella, Marcello Giardina, Shuichi Nishio, Hiroshi Ishiguro: "An Architecture for Telenoid Robot as Empathic Conversational Android Companion for Elderly People", Intelligent Autonomous Systems 13, Advance in Intelligent Systems and Computing, Volume 301, ISSN 2194-5357, Springer Editor.

- Rosario Sorbello, Antonio Chella, Carmelo Calí, Marcello Giardina, Shuichi Nishio, Hiroshi Ishiguro:"Telenoid android robot as an embodied perceptual social regulation medium engaging natural human-humanoid interaction", Robotics and Autonomous Systems 13, Volume 62, ISSN 0921-8890, Elsevier Science Limited.

1.4 General Outline

The thesis is organized as follows:

- Chapter 2: presents the scientific literature and distinguishes the choices made in the present architecture from the state of the art.
- Chapter 3: presents the Experiment to Dinfo Department - Acceptance of Telenoid Robot as Perceptual Social Agent in Human Humanoid Robot Interaction; the theoretical base at support of the study, the proposed approach, the experimental results and the final observations.
- Chapter 4: presents the Telenoid Dialogue system at support of the interactions between robot and Elderly people; the methods adopted, the knowledge acquisition, the emotional, cognitive and behavioural component of the robot, the experimental results and the final observations.
- Chapter 5: presents the project of "the Humanoid Robot as Communicative Mediator for Social Integration of Autistic Children "; the methods adopted; the experimental setup of laboratory with the children involved; the results and the final observations.
- Chapter 6: presents the project "Brain Computer and Robotics in disease management for patients with ALS (Amyotrophic Lateral Sclerosis)"; the methods adopted; the proposed system; the experimental setup of the test with ALS patients involved on the project; the results and the final observations.

Chapter 2

State of the Art

This chapter presents a review of relevant works on Human Humanoid Interaction, particularly those with biological inspiration.

Since humanoid robots are going to be part of the lives of human beings, specific studies are oriented to investigating collaborative and social features related to human-humanoid interaction (HHI) (Balistreri et al., 2011b),(Ishiguro et al., 2012),(Balistreri et al., 2011a). The HHI is oriented toward a cohabitation environment where human and humanoid will share common tasks and goals (Chella et al., 2011a).

In particular (Kanda et al., 2008) focused their attention to the concept of "communication" humanoid robot thinking as partner to help the human activities. (Oztop et al., 2005) put their attention to understand the perceptive relation between human and humanoid robots. The minimal agency includes a key aspect that is defined as "*sense of co-presence*" (Durlach and Slater, 2000),(Zhao, 2003). The research in the HHI field is oriented in the direction of "*sense of being together with other people in a shared virtual environment*" (Slater et al., 2000). In particular I am interested in the study related to the sense of a person to be present in a remote environment with a robot ("*Telepresence*") and to the sense of a person to be present in a common environment with a robot where humans and humanoid are "*accessible, available and subject to one another*" (Goffman, 1966)(Mutlu et al., 2009) (Shimada and Kanda, 2012). Based to other study is The iCat, (Poel et al., 2009), where it is used as a user-interface robot able to exhibit a range of emotions through its facial features and it is generally controlled by predefined animations. ICub (Fig. 2.1), (Metta et al., 2010), is a a child humanoid robot used in embodied cognition research. In contrast to these typical humanoid robots, (Kanda et al., 2004), Geminoid HI-1 (Fig. 2.2) is a humanoid robot with the external appearance of its ideator, Prof. Hiroshi Ishiguro and it is thought of being indistinguishable from real humans at first sight. Repliee R1 and Repliee Q2, (Minato et al., 2004) and(Shimada et al., 2006), are designed with the same aim of Geminoid, and the

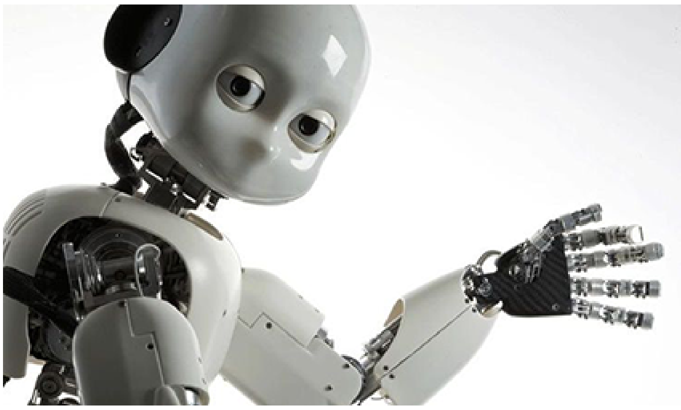

Figure 2.1: Icub Robot

use of these kinds of robot points towards Android Science, (Ishiguro, 2007). The uncanny sensation using android, as (Shimada et al., 2006) shown in their work, is reduced when their behavior's complexity is improved.

A number of researches have investigated how human-robot interaction may be assumed to comply with the natural conditions of agents overt perceptual and social behavior. In this thesis I use Telenoid as testbed for studying the social and perceptual abilities of robots copying with environment. In particular, much relevant literature appeared on the features of natural character of agents interaction. Since 1970, the research of (Argyle et al., 1973) and(Argyle and Cook, 1976), was devoted to specify the different functions of gaze, and to highlight the proxemic indicators that contribute to the organization and to the dynamics of the interpersonal space that subserves social cognition and behavior (the so-called equilibrium theory of Argyle and Dean,1965(Argyle and Dean, 1965)). (Torres et al., 1997) apply the empirical analysis of gaze behaviour in a dyadic-conversation paradigm to show a meaningful relationship between gaze and information retrieval of discourse content in the communicative humanoid agent proposed by (Thórisson, 1997). (Vertegaal et al., 2001) argue that evidence of gaze function in coordinative behaviour comes from research on gaze directional clues as reliable non-verbal predictors of conversations in multi-agent, multi-user environments. (Mutlu et al., 2006) study the extent at which gaze contact frequency between a storytelling robot and its human listeners is correlated with story understanding and recall. They argue that results highlight meaningful commonalities between human-human and human-robot communication. (Jackson and Decety, 2004) and (Pietroni et al., 2008) drew attention to other features of overt behavior that play an informative role in understanding other human agents purposive and intentional behavior. Furthermore neurophysiological research (Anderson et al., 1991) achieved consensus findings about the crucial link among the perception of emotional features,

Figure 2.2: Prof. Ishiguro and the robotic copy of himself (Geminoid Robot HI-1)

gaze of interacting agents, and the representation of emotional and social significance of salient stimuli by tracking the coordinated activation in specialized cortical and subcortical systems (Adolphs, 2002); (Phan et al., 2002); (Phillips et al., 2003)).Therefore I reasoned that the emerging picture of natural interaction condition requires a general description of the environment, where cognitive and social interaction of agents with their surroundings obtains, which can be carved up at the level at which environment look somehow like to agents. Indeed since (Koffka, 1955), (Köhler, 1992), (Heider, 1982), and (Lewin et al., 1936), cognition and behavior is proposed to be analyzed at the scale of what agents themselves take as meaningful units. Accordingly their environment can be decomposed in what they see as directly or indirectly accessible objects sense properties, affordances, scaffolds and proxies of other agents intentions and behaviors. Analytical treatment of the behavioral environment as it looks like from an agents standpoint allows to recover its qualitative structure that support cognition, agency and interaction with other agents. As (Chrisley, 2009) pointed out, there is no hindrance to the definition of a synthetic phenomenology devoted to the research of perceptual qualities that carry out cognitive functions of an artifact such as a robot. Agents behaviour is to be explained as organized and regulated by the cognitive frames of reference that build up their environment.

In the recent past, the researchers are focused their studies on the Humanoid robots as systems which are able to communicate and interact with humans as if they were a real partner (Breazeal, 2003b). The communication between human beings could be classified in two different classes: verbal and non verbal. While the nonverbal communication is based on gazes, gestures or facial expressions, the verbal communication is completely based on the language (Breazeal, 2004). Human language is considered to be a natural and intuitive interface for communication with the robots. Despite this, it is very complicated to obtain good interaction with humanoid robots using verbal communication because of the nature of the auditory patterns and the nature of the human language (Benesty et al., 2008). The auditory stream contains much information which is difficult to manage: environmental conditions such as noises and echo are for instance the first problems to solve. A more complex problem to overcome is to focus the attention of the system on a single speaker among a series of conversation and the background noise. This problem is described in literature as the "cocktail party effect". The human language contains a series of information about the speaker, his identity, what the speaker is saying, how he is saying it and his prosody.

In addition to these problems, which are related to the auditory recognition, another problem concerns the natural human language (Jackson and Moulinier, 2007): scientists of language have not really understood all the basis rules of speech because it is impossible to describe it in term of syntax, semantic or phonetic rules in contrast to what happens with the written language. The comprehension of the real meaning of a conversation becomes a very difficult operation because of its incompleteness, of its ambiguity and of its few definite structures. Several studies have tried to overcome these hurdles both from the point of view of speech recognition and the point of view of natural language. Researchers in this field have demonstrated that in order to increase the accuracy of systems of speech recognition it is necessary to use a more detailed description of the phonemes, either using a sequence of three phonemes (Darjaa et al., 2011) or using a larger vocabulary or by providing additional constraints (Lee and Kawahara, 2009). From the point of view of research about communication of humanoid robots, many experiments have been done.

Systems based on simple commands have given important results, if the users were familiar with the command or if they had been previously educated about the behaviours of the robot (Shim et al., 2011).

Systems based on dialogue has been used successfully (Toptsis et al., 2004), but also in this case it is very difficult to reach a free conversation because the system would need to cover a wide range of possible conversations. Some systems have tried to use ontologies in order to recover a full comprehension of the conversation (Kobayashi et al., 2011). Other systems have tried to use custom algorithms, such as Latent Semantic Analysis, to extract from the speech some important characterizations, such as emotions (Anzalone et al., 2012), (Chella et al., 2011c). All these systems have the problem of having errors with speech recognition in input due to the speech recognition systems (Kraft et al., 2010).

Figure 2.3: Telenoid Robot

2.1 Telenoid Humanoid Robot R2

The Telenoid, as shown in figure (2.3), is an ageless android robot specifically designed for cognitive *human humanoid* communication. It is designed to appear and behave like a minimalistic human. At first glance, is it possible to recognize the human features in the robot, but it can be interpreted as being ageless either male or female. At first glance, is it possible to recognize the human features in the robot, but it can be interpreted as being either male or female, old or young. Due to this minimal design, the Telenoid allows people to feel as if a faraway acquaintance were close to them. The Telenoid has nine degrees of freedom. The provided DOF allow horizontal and vertical motion for the left and right eyes, opening and closing the mouth, yaw, pitch and roll rotations for the neck, as well as motion for the right and left hand. The Telenoid length is approximately seventy centimetres, and its weight is around six kilograms. The skin is made of silicon, and the touch is similar to the human one. The Telenoid is a teleoperated robot. The operator's face directions, mouth movements and facial expressions are captured by a face recognition system. Some spontaneous behaviors, such as 'bye-bye', 'happy' or 'hug' can be inducted by a graphical GUI implemented in a tablet. Some spontaneous behaviors, such as breathing, are generated automatically to create the sense that the robot is alive.

Telenoid, as a minimalistic human, was created following a strategy to remove as many unnecessary features as possible. Essential features that remained after this pruning process might be helpful to create efficient social cognitive robot which can be used by all the types of people. The aim for the Telenoid was to create a minimalistic human's appearance, as such an appearance might allow any kind of person to transfer their own presence to distant locations.

Chapter 3

Case Study 1: Experiment to Dinfo Department - Acceptance of Telenoid Robot as Perceptual Social Agent in HHI

The main goal of the proposed research is the analysis of the two social dimension (Perception and Believability) useful for improving the natural behavior between users and Telenoid.

I oriented my research in the HHI field in the direction of "*sense of being together with other people in a shared virtual environment*" (Slater et al., 2000). In particular I am interested to the study related to the sense of a person to be present in a remote environment with a robot ("*Telepresence*") and to the sense of a person to be present in a common environment with a robot where humans and humanoid are "*accessible, available and subject to one another*" (Goffman, 1966). In this direction, in order to understanding areas of human cognition, which have not been tested or clarified until now, I used the Telenoid robot.

In order to understand how some features of perceptual behavior work during the HHI I want to assess the degree of Believability of interaction along dimensions that can be reasonably taken as meaningful indicators of social interaction, both in free and task directed conditions.

3.0.1 The Proposed Approach

From the relevant literature, the perceptual and cognitive features of overt behaviour are used as parameters for the evaluation of natural HRI. Students of the Faculty of Architecture and Engineering (University of Palermo) were recruited for the tests and they did not have prior interaction experience with humanoid robots.

All participants (200 total, 124 male and 76 female with average age 22) has been introduced to the Telenoid, to the interaction setting structure that required a two stage interaction with the robot and to fill up a questionnaire. All interactions were videotaped. A first free interaction stage, meant to allow subjects to adapt either to interact with the humanoid robot or to acquire, as early as possible, the skills for operating the robot through the teleoperation system. A second interaction stage was instead task driven. Participants were allowed to choose an interactions scenario among a proposed range that are related to: booking a hotel reservation, making a phone call to a mobile company to get a contract or services information, to matriculate or to enter his/her name or one of his/her fellows ones for a course examination by talking directly with the robot.

The questionnaire was meant to recover some perceptual and social aspects of natural conditions of agents interaction. These emerged aspects mirrored some salient ordinary cognitive abilities, which agents could specify in such cases to improve the efficacy of interaction. The first construct (*Perception*) is intended to cover perceptual features of overt interactive behavior. The perceptual awareness of sharing a common environment with the robot is assumed to be part of momentous importance for the interacting subjects that ascribe intentions and actions to the robot itself. This aspect of interaction can prove to be the perceptual link with the second construct: *believability*. The concept is defined and operationalised by (Dautenhahn, 1998), and (Poel et al., 2009) and it is represented by items grouped according to the indicators of personality, emotion, responsiveness, and self-motivation. I treated the perceptual features of interaction as sensory qualities that appear to have a meaning in themselves only when appropriately fit to one another.

3.0.2 Experimental Results

The items of the questionnaire were designed to represent the two construct of the questionnaire: "*Perception*" and "*Believability*". Standard item analysis has been performed (Ishiguro et al., 2012) on the codified data using Split-half Spearmann Brown coefficient, Pearson Coefficient, and Cronbach alphas. These coefficients were calculated to test the reliability and internal consistency of the scales and the correlation among the multiple items of each single construct of the questionnaire.

We surveyed the judgement of a group of 200 young people about the degree of acceptability of these artificial robots in every-day life.

The alpha value is 0.88 for the items grouped under the construct Perception. The alpha value is 0. 77 for the items grouped under the construct Believability. As an additional indicator of internal consistency, we measured the Pearson coefficient of correlation by means of the split-half method as a gauge for the inter-correlation of items in either sections of the questionnaire. For the items of the construct

Figure 3.1: Human Telenoid Robot Interaction

Perception the Pearson coefficient value is 0,97, and its value corrected by the Spearman-Brown formula is 0,98. For the items of the construct Believability the Pearson coefficient value is 0.59, and its value corrected by the Spearman-Brown formula is 0.74.

Asking to the users, that had either none or mixed previous experience and knowledge of robotics, their reaction to the possibility to have the robot in the next 5 years inside the working and home space I registered the high degree of the acceptability of this artificial robotic agent in every-day life.

Information was taken on the general knowledge and attitude of subjects on issues concerning robotics and artificial intelligence. Subjects were asked to assess their extent of knowledge on robotics and autonomous artificial systems. Only 9.38% consider their knowledge significant in one's own culture, whereas most subjects claim to be only interested in it. About 85% of subjects claim to have acquired at least informal information on robotics and a really high fraction of them is likely due to casual circumstances where this information was acquired.

In order to have a preliminary assessment of subjects general attitude to robots, they were asked to rate their agreement or disagreement in a five point Likert question with items on five scenarios of robotics devices in the near future. Items were meant to give some proxies of the degree of acceptability of robotics, by combining distinct conditions: their use in one's own daily life or only for one's own works, their use in other's daily life or works along with the possibility of expressing directly the lack of approval or liking for any of these scenarios. Subjects responses are quite mixed and not easily interpretable. Most subjects show to have a likely favourable attitude towards robots and their use both in one's daily and work, with about 17% that claim to be indifferent.

The rate of agreement on a favourable attitude seems to decrease when other's life and works are concerned, but this can be due to a redistribution of responses among subjects who overall have a positive attitude and subjects who claim to be indifferent. The measures of the attitudes towards non-human agents were concrete measured with the following dimensions: (1) the varying degree of acceptance in first person or on the ground of others choices; (2) the degree of diffusion weighted by the artificial tasks of the robotic device either in household or at work; (3) the range of attitudes from favourable reception to interest, distrust, annoyance.

For a first descriptive survey of the data collected, it emerged that subjects consider robots as potential useful device for many day-life activities. The percentage of skeptical subjects is quite low with a value of 20-25% across all contexts.

3.0.3 Final Observations

The sense of "*togetherness*" between persons and Humanoid is "*inherently social*" (Biocca, 1992) and is highly connected with the concepts of particular behavior defined "*sensible*" because (Chella et al., 2011a) capable to express cognitive functionality. The present research aims at a descriptive analysis of the main perceptual and social and cognitive features of natural conditions of agent interaction, which can be specified by agent in human-humanoid robot interaction. These perceptual and cognitive features underlie the ordinary human-human interaction in daily life contexts and therefore can be exploited in the human-humanoid interaction because they could be seen as a sort of specialized implicit competences.

The research was focused on the social dimensions that allow to obtain in general a successful interaction. Some of these dimensions were extracted to be embedded in a controlled interaction set up and then they were studied by representing them as items of a questionnaire.

I found that the sense of a shared environment is substantive for obtaining a satisfying interaction, where the distances between agents are construed as perceptual-motor proxies of the regions where intentions and actions are available at a glance. The characteristics of the present research aims at specifying the cognitive contribution of features that arise during the interaction to the perception of its consistent purposed driven behaviour. One crucial result is that subjects are significantly inclined to perceive the Telenoid as cooperative and competent even though this favourableness, is not associated with a clear cut implicit non-perceptual criterion for ascribing such behaviour to the Telenoid. Since the users perceived the behaviour of the Telenoid coherent and consistent, and the results suggest that they are favourably incline to accept the humanoid's behaviour as tuned to theirs and moved by the commitment to meet their demands.

This chapter presents a complete descriptive analysis of this preliminary research, which can be taken as a first basis to build a scale to assess the efficacy of human-humanoid interaction as their natural-like character in connection with free or task driven conditions. Some findings emerge that might prove as much emphasizing the dimensions that make a human humanoid interaction socially acceptable and efficient in task driven contexts as also relevant for future research. Future works will analyze whether items cluster within and across the constructs, whether they constitute a cognitive continuum or two specialized scaffolds for interaction, whether the features they represent only mutually reinforce each other and give a joint contribute to interaction.

Chapter 4

Case Study 2: Robot and Elderly People

The Intelligent use of the telenoid with Elderly people is a interesting applied study on communication with human and humanoid robot. A Preliminary work was made with the Analysis of a Dialog between Grandparent and Nephew at a nursing home for the elderly. The Target of the analysis was to find some kind of repetitive pattern inside the conversation between an elderly person and an interlocutor, that can be analyzable and reproducible. In particular, I have tried to study the capabilities of the humanoid of contextualizing a conversation after having created a suitable system. Conducting the studies in this direction, the humanoid robot Telenoid has been used for the interaction with the user. It is a special humanoid robot since its minimal aspect and it has been designed for appearing as a human being without instilling fear in the user. The purpose is to create a humanoid robotic architecture with the ability to allow the humanoid robot to be involved in an empathic dialogue with elder human, recognizing the context and the topic of the current conversation. For reach this goal the project need to address three distinct main issues: conceptualize the conversations (*Topic Recognition*), answer to questions (*Question Answering*) and become acquainted of new concepts (*Learning*).

So, the real challenge is to have a common environment where human and humanoid robot could engage a proficous "dialogue" in order to share the "sense of co-presence", exhibiting a natural (Iwamura et al., 2011) (Powers et al., 2005) and empathic interaction with human beings, supporting a pseudo realistic conversation and expressing human-like body emotions.

It has become indispensable that the human-humanoid communication should be as efficient as possible. In contrast to other robots which are in other environments and are able to perform only some basic operations in a specific reality, the humanoid must be in a position to interact with the human and to update its own knowledge day by day and learning through natural language or through the use of texts and, finally, self-learning during the normal interaction with people, such as happen in the baby learning processes.

In order to obtain a good affinity matching with humans, the robot can exhibit a smoothly natural transition in his emotion, changes during the interaction with humans taking also into account the previous generated emotions. The proposed emotional cognitive architecture of humanoid robot is based on the creation of a probabilistic emotional conceptual space automatically induced from external environment. The approach is based on the application of the paradigm of Latent Semantic Analysis (Chella et al., 2008) and the architecture of the presented system is inspired to the layered approach illustrated in Chella et al. 1998 (Chella et al., 1998).

In particular, I have conducted human-robot interactions test with elderly people with no prior interaction experience with robot. During the experiment, the person engaged stimulated conversation with the humanoid robot and he receives, as feedback, the human-like presence of the robot, laying the basis for the generation of an efficient and grateful interaction with the sense of don't be alone.

Several test sessions have been done using the Telenoid with elderly people. In relation to the tests done, it has been possible to evaluate the reliability of the system.

4.1 Telenoid Dialogue System

The Architecture of the proposed system named as TDS (*Telenoid Dialogue System*) is organized in three main hight-level areas, each of them related to a specific logic functionality designed:

- Knowledge Acquisition Area: it creates the corpus of knowledge base, retrieves informations, processes and represents them in the form of vectors in a conceptual space, and in a complex schematic architecture.
- Emotional-Cognitive and Behavioural Area: it constitutes the "conceptual space" where emotions, behaviors, personality of the robot, together with the perceptual data are mapped as emotional knoxels and conceptual konoxels. The coding of the personality distorts the perception of reality of the robot making the "thinking area".
- Robot Control Area: it analyzes perceptual data coming from the sensors, manages the external behaviours of the robot by generating the most coherent action in according to the emotional state inferred; implements the humanoid behaviours as actions, and it is capable of accepting requests of

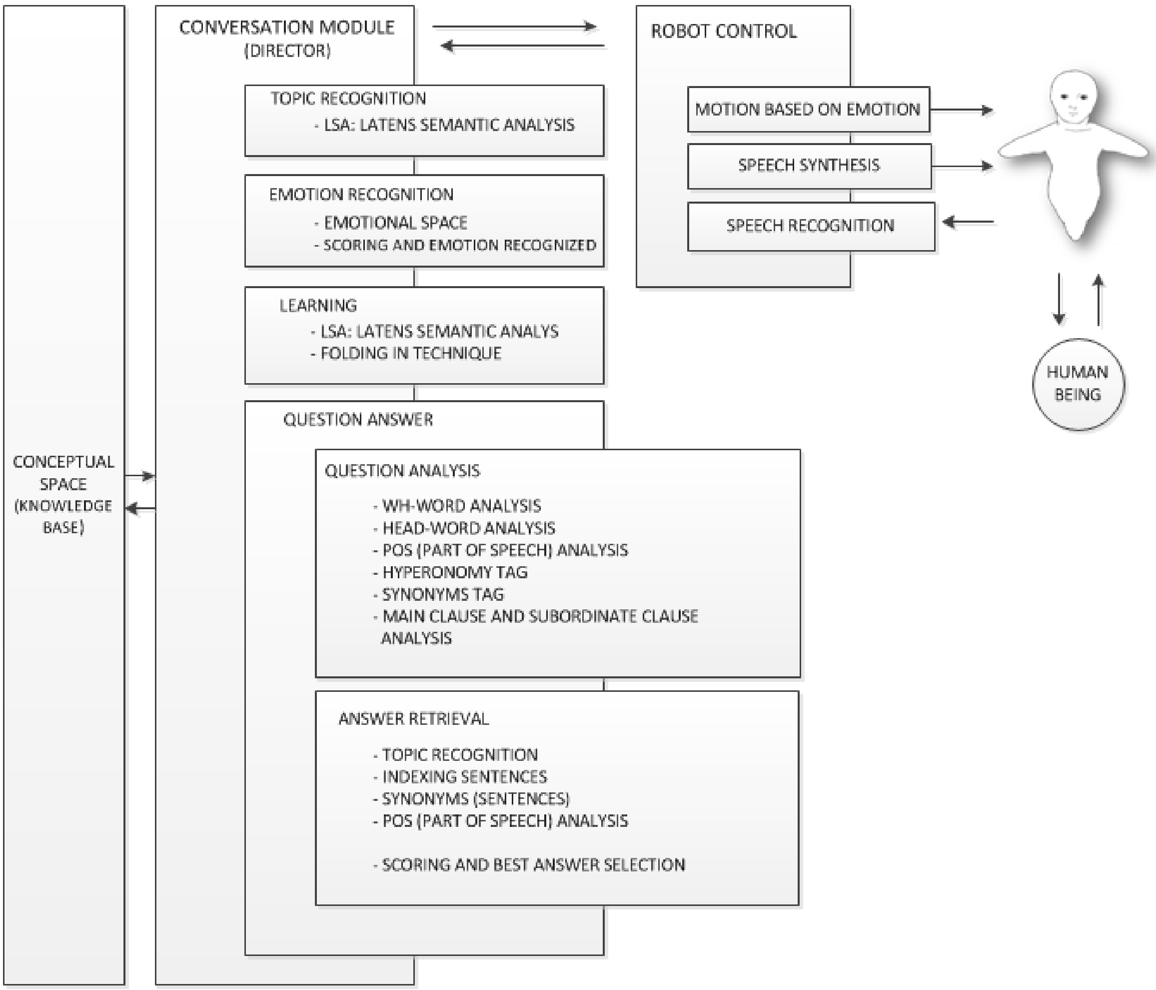

Figure 4.1: Overview of the Architecture of Telenoid Robot.

behaviours and executing them by controlling the joints of the robot. The actions of the behaviours are implemented as interpolation of different poses of the robot.

An overview of the architecture of the system is shown in figure 4.1.

The principal modules of the architecture are: the module *Document Preparation* of knowledge base is related to the role of documents extrapolating from the network, cleaning and indexing of the knowledge base. The module *Question Analysis* builds the question, analyses and finds the important constructs necessary for the correct answer. The module *Answer Retrieval* extract the best candidate answer, using a scoring system based on boolean algebra and SVM. The module *EmotionRecognition* allows the system to detect the emotions correlated with the answers of the human users (Chella et al., 2011b) in according to general mood of the dialogue session (Chella et al., 2011c). A finite state machine, as shown in figure 4.10, describes the logical states that allows the dialogue, between human user and the Telenoid robot.

4.2 Knowledge Acquisition

4.2.1 The Conceptual Space

The modules related to topic recognition and knowledge learning, take advantages of using the conceptual space to carry out their functions. To create the conceptual space the algorithm of the Latent Semantic Analysis (LSA) has been used (Chella et al., 2013). The LSA is a technique which allows the creation of vector representation of text from which we want to extract the semantic content. So, we can use the vector representation of the knowledge for the calculation of the similarity of pairs of concepts expressed as sentences, by comparing their vector representations. The key idea is to build, starting from a large amount of documents, a vector space with a lower dimension called Latent Semantic Vector Space in which a given word has a strong connection with the topic which it refers to. In doing so, we obtain an algebraic structure which allows us to use all the proper tools of linear algebra. Initially, Latent Semantic Analysis will analyze the words within documents (our topics) bringing in a matrix, called as "matrix of the co-occurrence", having in the cells, the frequency with which a particular word occurs in the various documents. We will have a lot of lines to represent the different words found and as many columns to represent documents which have been analyzed. The co-occurrence matrix thus obtained undergoes a process of decomposition through the Singular Value Decomposition usually indicated by the acronym SVD. From the decomposition are obtained three matrices: U, S and V such that

$$A = U\Sigma V^T \tag{4.1}$$

Let us suppose that **A**'s singular values are ranked in decreasing order. Let K be a positive integer with K< N, and let **U** be the M $\times$ R matrix obtained from **U** by suppressing the last N - K columns, Σ the matrix obtained from Σ by suppressing the last N - K rows and the last N - K columns and $\tilde{\mathbf{V}}$ be the N $\times$ R matrix obtained from **V** by suppressing the last N - K columns. Then

$$\tilde{A} = \tilde{U}\tilde{\Sigma}\tilde{V^T} \tag{4.2}$$

is a M $\times$ N matrix of rank R. It can be shown (Agostaro et al., 2005), that, according to the illustrated procedure,$\tilde{A}$ is the best rank K approximation of **A** (among the M × N matrices) with respect to the Hellinger distance, defined by

$$d_H(X,Y) = \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}(\sqrt{x_{ij}} - \sqrt{y_{ij}})^2} \tag{4.3}$$

which allows to interpret the TSVD technique as a statistical estimator. The two matrices $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$ obtained after truncated decomposition reflect a breakdown of the original relationships identified by **A** into linearly independent vectors. The K columns of $\tilde{\mathbf{U}}$ constitute a basis of a semantic space S, which can be interpreted as a parameter K indicating the *dimensionality* of the "Conceptual Space".

4.2.2 The Emotional Space Creation

The emotional area allows the humanoid robot to discover emotional analogies between the current emotional status and the previous knowledge of "*Emotion*" of the robot; both of them are mapped in the semantic space of emotional state. An ad-hoc corpus of documents dealing with emotions has been built and used in order to infer a semantic space of emotional states. Emotional states have been coded in this space using proper subsets of verbal description of emotional situations (environmental stimuli, context, spoken message, and so on.) that evoke them. This represents the knowledge base of the robot about emotional state. Environmental incoming stimuli are encoded in natural language words and subsequently mapped in this space in order to find empathic analogies. The process of the data collection in order to make the induction of the emotional space has been obtained trough an accurate selection of excerpts associated to feelings. So, a large amount of sentences has been selected from several documents that coming from the poetry literature, from the theatrical production and from the way to say of general speaking that are normally used for describe with natural language an emotion or the description of situation, smells, images that can reminds a specific feel. If, I want do an example of data collection for the happiness feel, I can maps the sentence :

- "Happiness is something everyone wants to have"
- "You may be successful and have a lot of money but without happiness it will be meaningless"
- "You may be successful and have a lot of money but without happiness it will be meaningless"
- "Let's share my birthday cake"

- "Happiness is my goal"
- "Happiness is given by the fruit of your labor and money that this can give you to support your family"
- "To look at the sunny side of everything"

If I consider the previous sentences individually they express a concept or describe a situation or a mode of live where the human can probably can feels happiness. All this sentences have a mining if they are contextualized, in their specific domain. With this approach, I can merge all the domains and extract from every context the profound meaning that stay at the base of happiness. So if I want to elaborate the emotion that can evoke the user sentence not mapped before on the emotional space : "I won a prize in the lottery" I obtain an hight perceptual of happiness . If I want to do a qualitative analysis of the relation inducted from this approach I can observe that if Happiness is "something everyone wants to have" and if Happiness is "You may be successful and have a lot of money but without happiness it will be meaningless" and if the Happiness is "Given by the fruit of your labor and money that this can give you to support your family" than I can conclude that a human can fells himself happy if win a prize at lottery because the prize is linked to concepts of money, of to have something (all things that generally make a person happy). I can apply the same analysis for the other emotions. I have selected the following emotional expressions: *sadness*, *fear*, *anger*, *joy*, *surprise*, *love* and the *neutral* state has also been considered. The architecture of the Emotional Inference Module is based on the creation of a probabilistic emotional conceptual space automatically induced from data. The methodology is inspired to the Latent Semantic Analysis (LSA) technique. The approach is based on the application of the Truncated Singular Value Decomposition (TSVD), preceded by specific pre-processing phases that consent to give a probabilistic interpretation of the induced vector space. The goal is to realize a mapping between the M words and the N documents, verbally describing emotions, into a continuous vector space S, where each word, as well as each emotion is associated to a emotional knoxel in S. In analogy with knoxels in conceptual spaces, an emotional knoxel is a vector in the probabilistic emotional space; from the conceptual point of view, it is the epistemologically basic element at the considered level of analysis. A matrix has been organized where the 6 emotional states excluded the neutral state and have also been coded in according to the procedure illustrated in the previous section that leads to the construction of a probabilistic emotional space. A corpus of 7 documents, six for emotions, equally distributed among the six states, and one for the personality of the robot, has been built. This set of documents represents the affective knowledge base about the emotions of the robot. Each document has been pre-processed with a filtering chain in order to remove all words in literature called "stopwords" that do not carry semantic information like articles, prepositions and so on and, also, all the words not presents in the english vocabulary in order to remove all the words that contains mistakes. Let **A** be the $M \times N$matrix whose (i, j)-th entry is the square root of the sample probability of the *i-th* word belonging to the *j-th* document. The Singular Value Decomposition of the matrix **A** is performed, so that **A** is decomposed in the product of three matrices: a column-orthonormal $M \times N$ matrix **U**, a column-orthonormal $N \times M$ matrix **V** and a $N \times N$ diagonal matrix Σ, whose elements are called singular values of **A**.

$$A = U\Sigma V^T \tag{4.4}$$

According to the LSA technique a 25000 × 6 terms-documents matrix (A) has been created where M= 25000 is the number of words that describes the emotional states and N=6 is the number of documents (six Emotions mapped).

The generic entry $a_{i,j}$ of the matrix is the square root of the sample probability of the *i-th* word belonging to the *j-th* document.

Let us suppose that A's singular values are ranked in decreasing order. Let K be a positive integer with K<N, and let U_k be the M×k matrix obtained from U by suppressing the last N - k columns, Σ_k the matrix obtained from Σ by suppressing the last N - k rows and the last N - k columns and V_k be the N×k matrix obtained from V by suppressing the last N-k columns. Then

$$A_k = U_k \Sigma_k V_k^T \tag{4.5}$$

is a M×N matrix of rank k, called matrix A_k.

The TSVD technique, with K=4, has been applied to A in order to obtain its best probability amplitude approximation ψ. This process leads to the construction of a K=4 dimensional conceptual space of emotions. The axes of S represent the "*fundamental*" emotional concepts automatically induced by TSVD procedure arising from the data.

In the obtained space S , the n_i*-th* document for each emotional state corresponds to one of the six "basic emotion". Instead, the *Personality* P has been projected in S using the folding-in technique, where each excerpt is coded as the sum of the vectors representing the terms composing it. As a result, the *j-th* excerpt belonging to the subset corresponding to the Personality state **P** is represented in S by an associated vector p_i in S.

According to this technique the *j-th* document of the set M - 1 corresponding to the emotion E_i is coded in S as a vector $d_j^{(Ei)}$ with j= [0- (M-1)] and the Personality document corresponding to Personality P is coded in S as a vector $d_j^{(P)}$ with j = M ; Each emotion E_i and Personality P appears represented in S by the Areas described by the set of vectors that code the relative documents associated. The basic idea is to use the knowledge base for build the emotion space and after this process use it, in order to map up the concept of Personality.

4.3 Emotional-Cognitive and Behavioral

4.3.1 The Emotional State Inference

I have assumed that inference of an Emotional State is based on inputs that coming by the merge of four important factors:

- Personality - it consist, in natural language, of the traits that constitute the personality of the robot, fallows the principles with some simplification of the Theory of Personality Traits of Gordon Allport (Allport, 1937).
- Environment - it describe the physical place where the dialogue is happening. For example if I can describe the Environment as "sunny room with big windows that going in rays of light", probably I might assume that the state of anime is more oriented to be happy .
- Events - they describe the main action of a possible scene. For example the simply fact of "Encounter between a grandmother and his nephew" might arouse a state of anime more oriented to be happy .
- Sentences - all the words that constitute a dialogue between two interlocutors.

The inputs from the Personality, the Environment, the Event and the Sentences of dialogue are coded in natural language words or phrases describing them and projected in the conceptual space using the folding-in technique.

The input representative of the Personality and representative the Environment are *"time-invariant"* and are coded, respectively, as a vector p and as a vector h.

The input representative of the Event and the input representative of sentence of dialogue are *"time-depending"* and are coded, respectively, as vector *d(t)* and as vector *a(t).*

All these vectors are merged together as a weighted sum in a single vector f*(t) that synthesizes the inputs stimuli used for the inference of the emotion, at instant t:

$$f^*(t) = \alpha \cdot p + \beta \cdot h + \gamma \cdot d(t) + \delta \cdot a(t) \tag{4.6}$$

where α, β, γ and δ are weights that allow assigning different relevance to the specific inputs coming from each one of the four factors. Finally, in order to avoid the bad effect caused by the fast change of mood derived of a punctual study of the emotion during the dialogue I introduced a sort of "*smoothing input*" taking in account, also, the contribution of the emotional states inferred in the very close past history.

We can define f(t) as the vector result as weighted sum in a single vector of the vectors of the emotions inferred in the very close past time as follows:

$$f(t) = omega \cdot f^*(t-2) + theta \cdot f^*(t-1) + zeta \cdot f^*(t) \tag{4.7}$$

where *omega*, *theta* and *zeta* are weights that allow assigning different relevance to the specific vectors resulting respectively at *time(t-2)* , *time(t-1)* and *time(t)*.

The emotional semantic similarity between the vector f(t) and the vectors E1 , E2 , ..., E6 , that code the six emotions in S, plus the "neutral" state, can be evaluated using the cosine similarity measured between each $d_j^{(Ei)}$ and f(t):

This procedure represents the common process that arises in human beings, when reality is "filtered" and interpreted by personality.

The emotional semantic similarity between the vector f(t) and the knoxels that code the six emotions in S, plus the "neutral" state, can be evaluated using the cosine similarity measure between each $d_j^{(Ei)}$ and f(t):

$$sim(f(t), d_j^{(Ei)}) = \frac{f(t) \cdot d_j^{(Ei)}}{||f(t)|| \cdot ||d_j^{(Ei)}||} \tag{4.8}$$

A higher value of $sim(f(t), d_j^{(Ei)})$ corresponds to a higher value of similarity between the emotion evoked from the input and the emotion E_i associated with the vector $d_j^{(Ei)}$. The semantic similarity measure is calculated between $f(t)$ and each $d_j^{(Ei)}$. The vector $d_j^{(Ei)}$, which maximizes the quantity expressed in the formula, will be the inferred emotional state E_i.

4.3.2 The Topic Recognition

The topic recognition module allows the contextualization of a dialogue. The human-humanoid dialogue through the activation of the module of the speech recognition is converted into text. Subsequently every word which constitutes the text undergoes a filtering process. Such a process consists of 4 phases: deleting the text of the words shorter than three characters, deleting words placed among the stop words, deleting words not in the vocabulary and finally word stemming. The word stemming consists of the reduction in common roots of the words having the same meaning but expressed in a different way. The words which have passed the phase of filtering contribute to the creation of text encoding. In particular, every word will be associated to a vector encoding. The vector encoding of the text is the vector sum of the homologous components of vectors of the individual words.

The vector thus obtained will be compared with the vector encoding of documents in the knowledge base through the use of the standard distance between vectors. The distance which is minimal allows us to determine which topic is more inherent with the topic faced by the user. To improve the performance of the system and ensure greater accuracy in the results obtained it is necessary in order to set a threshold and the amount of information needed to identify the encoding of the vector. If the minimum distance obtained exceeds the permitted threshold, the system will not provide any topic. In this way, it is possible to avoid the obtained results which are too generals. If the topic recognition failed for three consecutive times, the module will activate the brain form concerning the learning processes.

4.3.3 The Behavioural SubModule

The main functionality of this area is to activate a behaviour, which is coherent with the inferred emotional state E_i induced by the humanoid. Behaviour is described by a sequence of primitive actions, sent directly to the robot actuators. In order to give to the robot a non-monotonous, non-deterministic, and non-boring response, each emotional state E_i is related with a group of possible responses according to current information about the environment (for example the perceived object). One of these behaviours is selected evaluating a score w_{bi} associated to each one of them. The score is calculated as:

$$w_{bi} = \alpha_w r + \beta_w dt \tag{4.9}$$

where r is a random value ranging from 0 to 1, dt is the time elapsed by the instant at which the behaviour b_i has been executed and the instant at which this valuation is effected; α_w and β_w are respectively the weight assigned to the random *value* and to the time. The response with the highest weight is selected. Since the stimulus is weighted, also the reaction will be executed with the same intensity. This corresponds to quicker, faster or slower movements of parts of the body. If the emotional state is classified as "neutral" a standard behaviour (to lie down, to sleep, and so on) is randomly selected.

4.3.4 Knowledge Learning SubModule

To allow Telenoid to learn new concepts without having to create the knowledge base each time, the module "knowledge learning" has been provided. The addition of new concepts without the generation of a new conceptual space, it is crucial to ensure a quick and constantly updated system. The technique used to add new concepts to existing space is the "*Folding-In*".

From the human interaction with the robot, the vector encodings of the words of the dialogue will be extrapolated from the Semantic Conceptual Space. The

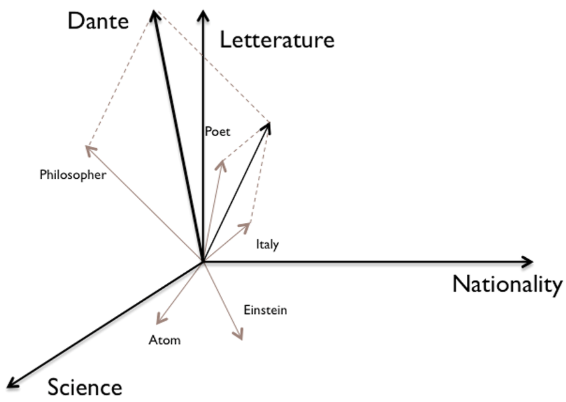

Figure 4.2: LSA: Folding-In Technique

sum of these encodings will generate a new vector containing the coding vector of the text which refers to the new topic added.

4.4 Question Answering Module

In order to make the interaction between human and humanoid as naturally as possible the system has been equipped with a module called "Question Answering" which allows to obtain the best answers to questions posed in natural language. The process of returning the response in according to the context and the topic of the current conversation between a human and the robot takes place through the following three fundamental steps:

- Document Preparation: plays the role of analysis and indexing of the knowledge base.
- Question Analysis: is responsible of the preprocessing and the analysis of the question and, of the constructs identification useful for the determination of the answer.
- Answer Retrieval: makes the discovery and the extraction of the best answer candidate to question done.

4.4.1 Knowledge Base

It is equipped the robot of a very small knowledge base that can support a very simple conversation with an elderly person. In according to general culture of the person it is possible choose the knowledge that it is better to do at the robot in

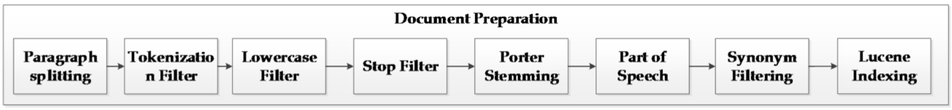

Figure 4.3: Preprocessing Pipeline of Knowledge Base

order to sustain a initial conversation. So after this first preliminary and accurate choices, it is easy, makes the Knowledge base directly by internet through discovery information algorithms. The Knowledge base is organized and stored by form of documents subdivided in corpus of texts. Each document represents a general topic about an conversation argument,and it may be different from person to person. After a few conversation trials, thanks to learning module, step by step, the Knowledge Base grows ever more, learning about the new topics that recur from time to time on specific elderly person, obtaining in this way a sort of knowledge base tailored on the subject. So isn't necessary equip the robot of all Knowledge of the world, because its interlocutor not have it. In this way, it is possible, avoid all the issues of informational retrieval derived of the big data manipulation.

4.4.2 Document Preparation

The phase of document preparation plays a crucial role for the retrieval phase. In this phase, the raw data of the knowledge base are filtered, analyzed,indexed and organized in according to the strict rules of the proposed model. I have chosen to represent the knowledge base in two separate but linked knowledge dimensions. The first dimension the knowledge base is modelled following an algebraic model through the mapping of the raw data in a complex conceptual space (see the paragraph 4.2.1). The second dimension the knowledge base is modelled in a structured model based on files and indexes.

Documents indexing

This data modelling is based on the fuse and the union of the potentiality deriving by several technologies, such as Lucene Framework, Wordnet database and the Stanford Parser Framework . The main goal of this structured model is covert the raw data in informations easily available and accessible.

The preparation of the document are performed in sequence for each topic, following the pipeline shown in figure 4.3.

In a first pre-processing step, it is necessary subdivide into every small parts the texts of the knowledge base documents, enucleating from these, the sub-concepts and analizing the main sentences and the sub-sentences, elaborate them, adding tags and other informations induced by a lexical and semantic analysis.

The first phase on the preprocessing step consists in a subdivision of the topic into paragraphs. A paragraph contains the minimum information of complete sense that it is possible extract from a long text. Each paragraph, that terminate with a point, indicates the end of a period of self-contained and complete sense.

The system elaborates all words presents in each paragraph by removing all non-alphanumeric characters, removing all the stop-words, and proceeds with the stemming operation. For each word in a document, Lucene Engine creates a separate token. Every Lucene token has its own position in the token stream. This position remains relative to its previous token and is sorted as a position incremental factor. The Type attribute string is assigned by lexical analizer that defines the lexical or syntactic class of the lucene token.

Documents and fields are Lucene's fundamental units of indexing and searching. A document is Lucene's atomic unit of indexing and searching. It is a container that holds one or more fields, which, in turn, contain the real content. Each field has a name to identify it, a text or binary value, and a series of detailed options that describe what Lucene should do with the field value.

A Token is an occurrence of a term from the text of a field. It consists of a term's text, the start and end offset of the term in the text of the field, and a type string.The start and end offsets permit to applications to re-associate a token with its source text to display highlighted query terms in a document browser.

As a design choice, I have chosen to create a separate Lucene index for each topic of the knowledge base in order to increase the flexibility of Telenoid dialogue system (TDS) for two simple reasons: the first, through the use of independent indices, the insertion, modification or removal of a topic from the corpus does not require the reconstruction of entire index, with considerable savings in terms of execution time; the second,through the use of independent indices, it is possible have more smaller index, instead of to have a unicum big index. Thanks to this design choices, at retrieval time, it is possible focus the research of the best answer on restrict index size, obtaining, as will be shown, in experimental results paragrapher, better change to find the best answer possible. (on which topic index focus the research of the best answer is delegated to the topic recognition module, that give us the first information on the conversation topic, that correspond with its topic index).

As shown in Figure 4.4 for each topic is present an index and each index is composed of a set of documents (corresponding to the individual paragraphs in which the topic was decomposed), each one of which, contains the following fields:

- Id: the id of the document.
- Contents: contains the words filtered by TokenStream and for each word are injected synonyms.

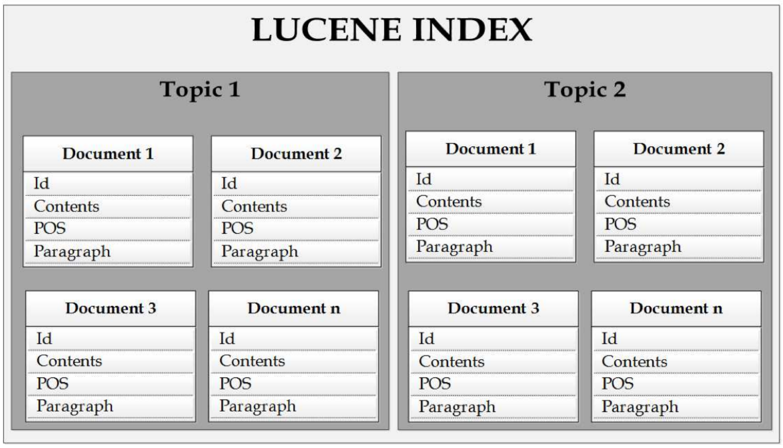

Figure 4.4: Lucene Index Fields (for each topic)

- POS: contains the parts of speech of each word feature of the paragraph.
- Paragraph: contains the original paragraph, unedited, unfiltered, and not tagged; This field is present to allow a fast system response in the case of matching demand with the current paragraph.

(About Lucene Engine, it is a powerful Java library that lets us easily add document retrieval to any application. In recent years Lucene has become exceptionally popular and is now the most widely used information retrieval library)

Wordnet Index

To extend the range semantics of words is useful to consider also synonyms, so in the discovery phase, the system is able to find answers to questions that contain not only the exactly words in the index. Using synonyms, it is possible, also, to find the answers to questions with the same sense, expressed in different terms respect the terms used in the formulation of the question. To obtain synonyms I used a dictionary of synonyms, in which each word has its synonyms. In particular, to take into account the polysemy I used the potentiality offers by Wordnet to expand the words with their synonyms. Wordnet is a lexical database developed for the English language in which the nouns, verbs, adjectives and adverbs are grouped into neighbouring groups for cognitive meaning those synsets. Each of these expresses a distinct concept, and they are interconnected for meaning through relations of conceptual-semantic and lexical. In order to expand the terms I need to create a Lucene index that contains the dictionary Wordnet through the use of class Syns2index. The distribution of Wordnet contains files prolog, with information

Field	IdfpoPSVBNtxx#txxDtxx	Norm	Value
syn	Idfp--S--Ni08--------	0.4375	coil
syn	Idfp--S--Ni08--------	0.4375	helic
syn	Idfp--S--Ni08--------	0.4375	spiral
syn	Idfp--S--Ni08--------	0.4375	volut
syn	Idfp--S--Ni08--------	0.4375	whorl
word	Idfp--S--Ni08--------	1.0	turbin

Figure 4.5: Lucene Index of Synonyms

corresponding to the synsets and the sense of the terms contained within Wordnet. The class Syns2Index have access to these files, it retrieves the content and generates an index Wordnet as the structure shown in Figure 4.5.

For each document Lucene index are two types of fields: the field *word* that contains the original term and the field *syn* that contains one of the synonyms.

(About WordNet, is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations.)

POS Tagging

The part-of-speech tagging (POS tagging), is the process of grammatical tagging or word-category disambiguation, marking up a word in a text (corpus) as corresponding to a particular part of speech, to each word (such as noun, verb, adjective) taking in account the relationship with adjacent and related words in a phrase, sentence, or paragraph. So, the POS tagging process is a division of words in a sentence, based on the syntactic role assumed by the words within it. The POS tagging operation is important for understand the sentence meaning on the basis of the organization and arrangement of phrases that compose it. In order to be able, better, to classify the paragraphs, I need to add information about the parts of speech in all sentences that compose them.

To perform this process, I use the implementation, available as part of the Stanford CoreNLP, with an implementation of the model described in (Manning et al.).The library contains two models tagger trained for English based on the Penn Treebank.

The class TagText is used for annotate each paragraph with its syntactic information by exploiting the MaxentTagger and getting a structure like this:

$$This_D T\ is_V BZ\ a_D T\ test_N N \tag{4.10}$$

As is shown in the example (4.10) the part of speech are tagged next each word.

00 adj.all all adjective clusters
01 adj.pert relational adjectives (pertainyms)
02 adv.all all adverbs
03 noun.Tops unique beginner for nouns
04 noun.act nouns denoting acts or actions
05 noun.animal nouns denoting animals
06 noun.artifact nouns denoting man-made objects
07 noun.attribute nouns denoting attributes of people
08 noun.body nouns denoting body parts
09 noun.cognition nouns denoting cognitive processes
10 noun.communication nouns denoting communicative
11 noun.event nouns denoting natural events
12 noun.feeling nouns denoting feelings
13 noun.food nouns denoting foods and drinks
14 noun.group nouns denoting groupings
15 noun.location nouns denoting spatial position
16 noun.motive nouns denoting goals
17 noun.object nouns denoting natural objects
18 noun.person **nouns denoting people**
19 noun.phenomenon nouns denoting natural
20 noun.plant nouns denoting plants
21 noun.possession nouns denoting possession
22 noun.process nouns denoting natural processes
23 noun.quantity nouns denoting
24 noun.relation nouns denoting relations
25 noun.shape nouns denoting dimensional shapes
26 noun.state nouns denoting stable states
27 noun.substance nouns denoting substances
28 noun.time nouns denoting time relations
29 verb.body verbs of grooming and bodily care
30 verb.change verbs of size, temperature change
31 verb.cognition verbs of thinking, judging
32 verb.communication verbs of telling, asking
33 verb.competition verbs of fighting, athletic activities
34 verb.consumption verbs of eating and drinking
35 verb.contact verbs of touching, hitting, tying,
36 verb.creation verbs of sewing, baking, painting,
37 verb.emotion verbs of feeling
38 verb.motion verbs of walking, flying, swimming
39 verb.perception verbs of seeing, hearing, feeling
40 verb.possession verbs of buying, selling, owning
41 verb.social verbs of political and social activities
42 verb.stative verbs of being, having, spatial relations
43 verb.weather verbs of raining, snowing, thawing
44 adj.ppl participial adjectives

Figure 4.6: Lexical Wordnet Category

On the basis of the syntactic category found, the strings are tokenizzed in single words and passed to the class OnlyPosTag that provides to make the *Lemmatization*, giving to each word its canonical form for example *(is -> be)*.

Subsequently an additional step of POS tagging is performed, using the wordnet's lexical categories. Wordnet make a division into four lexical categories: *noun, adjective, verb and adverb* but suffers of a serious problem of disambiguation then, for example, a word like "can" could be understood both as a verb (ex. can be) or as a name (ex. pepsi can).

Wordnet organizes nouns, verbs, adjectives and adverbs in synset that are express according to their syntactic category. The names and verbs are divided based on semantic fields; adjectives are classified into descriptive adjectives and adjectives relational.

It is possible avoid this kind of disambiguation, analyzing the information obtained on previous Lucene tagging phase, allowing, in this way, to identify the correct lexical category among the 44 made available by Wordnet shown in Figure 4.6.

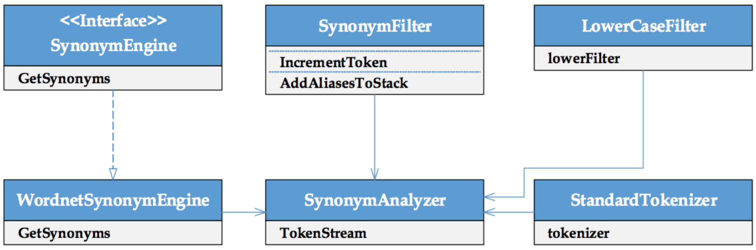

Figure 4.7: Synonym Retrieval Schema

After the step of tagging, a annotated sentence is obtained as fallow:

$$This_D T \; is_V BZ_v erb.stative \; a_D T \; test_N N_n oun.Event \tag{4.11}$$

Where the information resulting from MaxentTagger are integrated with those obtained by the lexical Wordnet's category.

WordnetSynonymEngine

The Wordnet Synonym Engine implements the interface that identifies and get the synonyms extracted from Lucene index and contains a representation of the dictionary Wordnet.

In figure 4.7, is shown, the Architecture Schema of Synonym Retrieval Engine.

The GetSynonyms method returns an array of strings containing synonyms of the string passed as argument. This method instantiates the object AllDocCollector, that inherits from Collector to collect the results of the research, and instantiates a "Searcher" that performs a search on the index Wordnet, returning each word that met the list of synonyms correspondents. The class extends the class "SynonymAnalyzer" which deals with the creation of the flow of the token. Inside the method creates a chain of filters, respectively StandardTokenizer, LowerCaseFilter and SynonymFilter carrying out the necessary analysis operations. The StandardTokenizer is a sophisticated tokenizer based on grammar and JFLEX. The LowerCaseFilter up to turn all characters in lowercase characters (lowercase). The SynonymFilter class receives an input token stream (pre-treated by previous filters) and SynonymEngine. The class through the method "getSynonyms" of Engine Wordnet gets a list of synonyms for the current word.

The synonyms are placed in the same position as the original word from incrementToken method, making the pop from the stack of synonyms. For example, for the word *Helic* will be identified synonyms shown in Figure 4.8 that are placed in the same position pf *Helic* Lucene index.

#	Score	Doc. Id	syn	word
0	4,1909	9303	coil helic spir	turbin
1	3,5922	12358	coil helic heli:	volut
2	2,9935	29861	coil corkscre\	spiral
3	2,3948	1930	curl curlicu gy	coil
4	2,3948	18267	coil curl curlic	whorl

Figure 4.8: Injection of Synonyms in Lucene Index example for the Word "Helic"

Thw injection of synonyms allows to extend the capability of detecting a correlation not only with the words contained in the index, but, also, with all its synonyms. The inclusion of synonyms in the same position of the original word, also, allows the system to perform searches based on phraseQuery without changing the structure of the sentence.

4.4.3 Question Analysis

When a Question is made by human during a dialogue, it is very important to analize it, at search of all the informations and data structure, that can help to find the best answer. In a cert sense, it performs the same process that happen in human being when he understand the question and give the answer. The Informations considered essential are the *Wh-word*, the *Head-word* and the *Part of Speech*.

- The Wh-Word: identify the type of question, analyzing the interrogative particle (a list is shown on table 4.1).
- The Head-word: identify the subject of the sentence.
- The POS: identify syntactic informations describing the role played by the words of the question.

For example, if we assume that the robot receives the following question (4.12):

$$From\ the\ beginning\ of\ the\ 18th\ century\ how\ many\ battles\ were\ fought\ ? \tag{4.12}$$

The system proceeds to analyze the question using the libraries "Stanford CoreNLP" and with process of decomposition of a period in main clause and subordinate clauses. In the specific, it make the stem of words, identify the parts of speech, extract names of persons, normalize dates and other numerical quantities, mark the sentence structure in terms of relations between terms and analize the decomposition of a period in main clause and subordinate clauses.

Wh Word	
When	temporal information
Where	information of place
Why	motivation, reason
Who	information about people, objects, groups, or artifacts
How many/much	quantity information
How far/distant	distance information
How Long	temporal information or duration
How	information of mode
What	information existential
Other Question	others

Table 4.1: Wh-Word List

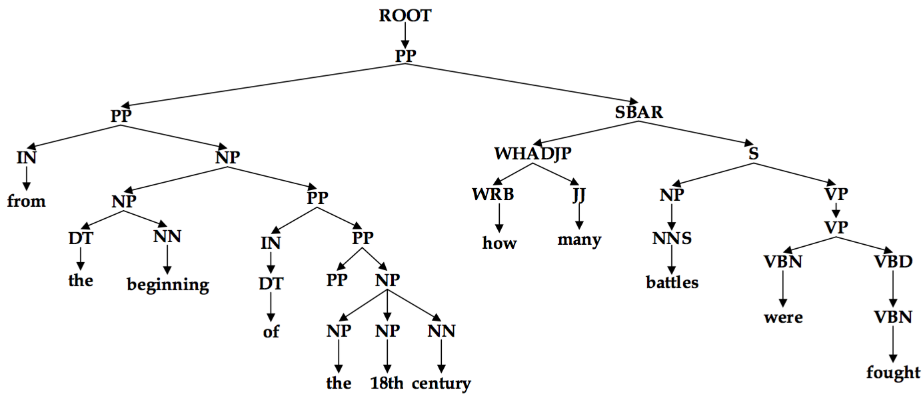

Figure 4.9: Syntactic Tree of the Users Question

Table 4.2: TregexPattern Rules

Rule	Description
[»WP»WHNP[»SBAR|»SBARQ]]	Explore the branch with What, Who or Where
[»SBARQ|SBAR][»(WRB $ JJ) | » (JJ $ WRB)|»(RB$WRB)]]	Explore the branch with How many / How much
[>WRB » WHADVP [»SBAR | »SBARQ]	Explore the branch containing How or Why.
[>WRB » WHADJ [»SBAR | »SBARQ]	Explore the branch containing where or when

The result of the question analysis is annotated in a tree with the structure, shown in figure 4.9.

In the tree, there are, the informations about the verbal POS and the syntactic POS . In the Example, the POS as DT, NN, NP refers to the *meaning* that takes the phrase in the current sentence, others such as PP, SBAR, WHADJP represent instead the *noun phrases* of which it is composed.

As shown in Figure 4.9, to the leaf nodes of the tree correspond to the terms of the question, instead the internal nodes define the syntactic structure of the sentence.

The questions, in example, consists of two parts:

the left branch:

$$from\ the\ beginning\ of\ the\ 18th\ century \tag{4.13}$$

and the right branch:

$$how\ many\ battles\ were\ fought \tag{4.14}$$

which in this case (the right branch [4.14] contains all the informative content of the question, so will be in this branch that we will better proceed with analysis of the *Wh-word* and *Head-word.*

For the detection of wh-word, the system uses the class called TregexPattern, available from libraries provided by the Stanford. This class allows my to explore the annotated tree with POS, highlighting the particular nodes inside. It was experimentally verified that the 5 rules (given in Table 4.2) appear to be sufficient to capture the wh-word correctly.

Parent Non Terminal	Priority List
FRAG	{rightExceptPunct}
PP	{ right , IN , TO , VBG , VBN , RP , FW , VBN , RP , FW , JJ },
S	{ left , TO , VP , S , FRAG , SBAR , ADJP , JJP , UCP , NP }
SBAR	{ left , WHNP , WHPP , WHADVP , WHADJP , IN , DT , S , SQ , SINV , SBAR , FRAG }
SINV	{ left , VBZ , VBD , VBP , VB , MD , VP , S , SINV , ADJP , JJP , NP }
SBARQ	{ left , SQ , S , SINV , SBARQ , FRAG }
SQ	{ left , VBZ , VBD , VBP , VB , MD , AUX , AUXG , VP , SQ }

Table 4.3: Collins Head-word rules

The Collins Head Finder - Headword

The head-word identification is delegated by the class "collinsHeadFinder", following the rules, shown in the Table 4.3, for the identification of Collins Head-word. In the specific, each rule search the correspondence of the node (if not leaf) with the head of the rule.

In according to the original Collins rules, I decided to modify them giving priority, in the discovery process to the part of branch containing the wh-word, because, in this part of branch there will be contained the subject of the sentence of our interest, meaning that in this part of the sentence is contained, what I want knows, of the question.

HyperonomyTag

The Head-word identified by the Head finder is passed to the class "Hyperonomy-Tag" in order to set category lexical word associated with the Head. The system performs a lemmatization, reducing the head-word to the basic form (is -> be). Through the library "CoreMap" of Stanford University, it is possible connects the dictionary Wordnet and find the synset corresponding to the word.Since the subject of the sentence is typically a name, we will limit its search to the lexical category POS.NOUNS.

In the following lines will be shown the entire process of answer retrieval, step by step, from the question to the best answer candidate through the aid of an example.

WH-WORD	POS
When	noun.time, noun.quantity
Where	noun.location, verb.stative
Who	noun.person, noun.organization, noun.animal, noun.group, noun.artifact
How many/ much	noun.quantity
How far/ distant	noun.location, noun.time, noun.quantity
How long	noun quantity

Table 4.4: relation between POS and Wh-Word

Suppose that the question is :

$$From\ the\ beginning\ of\ the\ 18th\ century\ how\ many\ battles\ were\ fought\ ? \tag{4.15}$$

The system annotates the question's phrase with the part of speech tag and identifies the Wh-Word, the Head-word and its lexical category. This elaboration produce the following result:

$$QT : 5\ HW : battles\ POS : noun.act \tag{4.16}$$

The first value (QT) indicates the Wh-word's id, in this case *how many*. This information is very useful because it indicates to the system that it is looking for an answer containing the informations of quantity. The second value (HW) (battles) is the Head-word of the phrase identified by the Head finder. The third value (POS) (noun.act), finally, is the lexical category of Wordnet synsets associated with the head-word.

The question sentence is subjected to a stemming process, and a common words (stop-word) removal process and, finally, in the injection of synonyms with WordnetSynonymEngine. After all these operations, the query will be presented as:

(begin start sourc showtim rootag root outset origin offset kickoff get first commenc) 18th (centuri hundr c) mani (battl struggl fight engag conflict combat bellicos) were fought

4.4.4 Answer Retrieval

The module *Answer Retrieval* extracts the best candidate answer using a scoring system based on boolean algebra and SVM. The question pre-processed, during

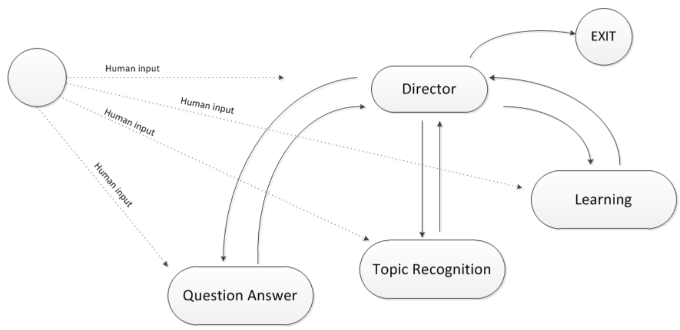

Figure 4.10: The Finite State Machine of Telenoid Dialogue System.

the Question Analysis phase, with tag and synonym is the input of the Lucene IndexSearcher engine. A "QueryParser" parses the user's query string and constructs a "*Lucene Query Object*", which is passed on to IndexSearcher.search() as the input. Based on this Query object and the pre-built Lucene index (made in Document Preparation Phase see the paragraph 4.4.2), IndexSearcher.search() identifies the matching documents and returns them as an "TopDocs" objects in the result with the best score. In this case the "TopDocs" objects return five of the best paragraphs candidate to be the answer (found with the index search and sorted in descending order in according to lucene score). The process of "answer retrieval" takes place analyzing the documents available to the system. With the topic recognition module is possible find "*What we are talking about*" and localize the right *Lucene Index* on which to proceed to search the best answer candidate to question posed in natural language.

Identified the correct index, corresponding to conversation topic, is shown in the table 4.4, the relations between *Wh-word* and the *Part Of Speech*, used as search criteria for Lucene IndexSearcher.

The last phase, is using the emotion recognition module, in order to understand the general mood of the dialogue and take in consideration, also, this kind of information in the process of discovery of the best 5 paragraphs (our answers) candidate to be the final answer of the robot to a user's question.

4.5 The Finite State Machine of Telenoid Dialogue System

It is described the logical schema of the functionalities used from the human user during his interaction with the robot. As reported in figure 4.10, the finite state machine is composed by three main states: *Question Answering State*, *Topic Recognition State* and *Learning State*.

4.5.1 Question Answering State

The user submits a question and the Telenoid through its system discovers the topic closest to the question using the conceptual space. After the topic is well defined, the search inside the conceptual space is obtained using the index factor related to the topic and the robot obtains the best answer for the user's question. Therefore, there are several cases that they could happens during the interaction:

a) The achieved topic is correct and the robot's answer is correct; b) The achieved topic is correct but the robot's answer is incorrect; c) The achieved topic is incorrect and the robot don't produce an answer; d) It isn't possible to achieve the topic related to the user's question.

If the answer is correct (*case a*) the system remains in a state of *wait* until the user chooses to go ahead with the next action. If the identified topic is correct but the answer is incorrect (*case b*), the system asks to the user if he wants to provide new knowledge to the robot. If the user accepts this request, the *learning* module is activated from the robot system and the new incoming information about the topic from a user is saved inside the conceptual space. At the end of the learning phase, the system returns to the wait state until the user chooses the next operation to be performed. If the identified topic is incorrect (*case c*) the user provides the system with a description of the topic he is talking about. In this case, the system calls the topic recognition module and it identifies the topic by providing the new correct answer. Then the user has the opportunity to assess the correctness of the answer similar to the case previously illustrated.

Finally, if the topic of the user's question isn't identified (*case d*), the system goes to state of *learning* and the robot asks the user if he wants to teach him new information, so the system transits to learning state.

4.5.2 Topic Recognition State

If the user chooses the *topic recognition* option, the robot asks the user the talking point. The user responds by providing the description of a topic and the system will search the information of topic described by the user starting to talk about it. During the interaction with the robot, the user is given the possibility to stop the dialogue of the robot. Finally, the system asks the user to rate the quality of the response. There are two possible typical situations: 1) The topic is correctly identified: The system returns to the *wait* state until the user will make another choice; 2) The topic is not correctly identified: the system asks the user if he wants to talk about a new topic and if user accept this option, the system goes to *learning* state.

Figure 4.11: A snapshot of elder user with Telenoid inside a Care Facility.

4.5.3 Learning State

In the *learning* state the user has the possibility to teach something of new to the robot, tolking about a new topic with its related information content. After the new information is submitted, the system evaluates if the content of the topics is quite enough. The acceptability of the information text needs to be measured using a minimum threshold, and in the wrong case the system returns an error and moves to the initial state. If the content is satisfactory, the system asks the user if he wants to pursue in the transaction of learning or if he likes to go back to the initial state. Based on the user's choice the system continues with the insertion of knowledge or the system returns to its initial state.

4.6 Experimental Results

In order to prove the positive impact on elderly people of the Telenoid robot, several sessions test has been made. Some of experiments session conducted with a "field work approach" in their real environment (Care Facility), as showed in figure 4.11. However, data regarding people's natural reactions to these humanoids cannot be collected easily in such environments. I believe that field environments, although uncontrolled, are important for improving the spontaneity of human-humanoid communication.

To validate the performance of Telenoid system, I have conducted different tests with goals focused to the evaluation of the following two key points:

- Evaluation and Performance of Telenoid Dialogue System with a variable value of dimensionality k of the conceptual space.
- Comparing of TDS benchmark with QAnus system (QAnus is a Question Answer system (Ng and Kan, 2012), developed by University of Singapore) benchmark.

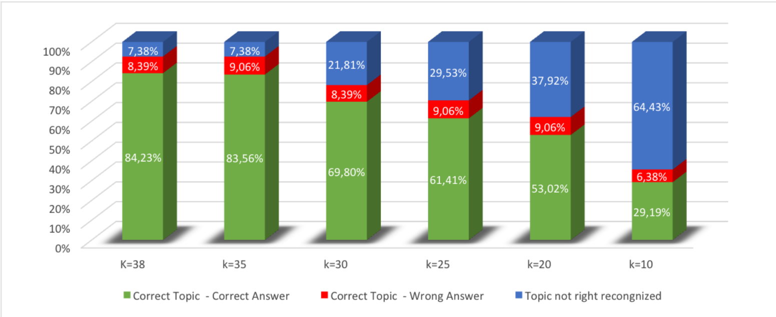

Figure 4.12: Question Answer Results of Telenoid Dialogue System.

I used a conceptual space composed with all the words took from the documents related to the topics. The topics are used in a structured not free context related to daily lives and dialogues of the elderly.

4.6.1 Evaluation of Telenoid Dialogue System

In order to evaluate the Telenoid dialogue system, a conceptual space formed with 38 topics that produced a total amount of 18115 words, has been made. For each of 38 topics I have submitted 9 questions to the system, and, for each question, I verified the correct match between the answer found by the system and a golden answer (the best answer expected). The golden answer was manually found after the analysis of the question submitted to system and the knowledge base of the robot.

I found during the experiments, three typical situations:

- The system revealed the correct topic of the question and it was able to give enough information in order to formulate the correct answer.
- The system revealed the correct topic of the question and it was not able to give enough information in order to formulate the correct answer.
- The topic found by the system was incorrect.

To assess the impact of the k dimensions of the Conceptual Space with respect to the successful rate of the system, I repeated the test with the same set of questions varying the value of K. At the beginning, I selected a value of K equal to the dimensions of the conceptual space. Later, I decided to proceed with test with a lesser value of K.

I considered a total amount of 298 questions submitted to the robot. In order to test the flexibility and the efficiency of the robotic system, user question with

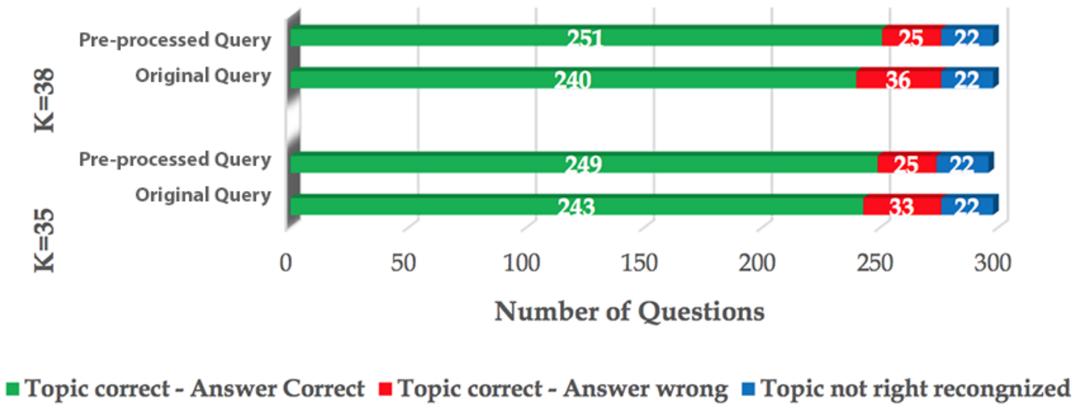

Figure 4.13: Comparison between preprocessed and not preprocessed query in TDS

a reduced number of words have been chosen. This choice is dictated by the fact of wanting to test the system in the worst case possible, starting from the consideration that, in general a short question have information content lower than a longer and, this fact, is translated in a higher difficult to find the right answer.

The higher success rate of the test was obtained with a value of K equal to 38 which represents the maximal dimension of the conceptual space (the second better value was 35 with a 0.8% positive reduction of the rate). In this case it is possible assess that the percentage of correct topic recognition is directly linked to the value of K because the structure of the question has effect to the compression factor of the space.

With K equal to 38 and 298 questions submitted, the system is able to identify correctly both the topic and the answer for about the 84% of the totals results (251 on 298 answers total (*using a user question that was preprocessed with synonyms injection, POS, head word and wh-word, parse trees*). For 25 questions (9% of the user queries), the system is able to identify the correct topic inside the question but is unable to *find the correct answer for the user*. Finally on 22 questions (7% of the user queries), the system is not able to identify the topic correctly. The figure 4.12 summarizes the results distribution previously described. It is also interesting to note that the average of wrong answers remains fixed on about 26.

I have, also, tested the impact on results submitting to system, the same question, two time: the first time, the question is preprocessed by the question analysis and the second time the question is not preprocessed (is the original user question). The goal of the test is evaluate, the success rate of the robotic system, using a conceptual space with *K=38* and with *K=35* to vary of the pre-elaboration method of the user query. As results, I obtained, as shown in figure 4.13, respectively for the

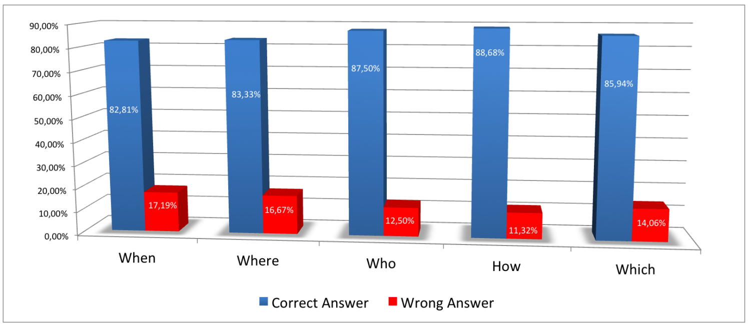

Figure 4.14: The Distribution of The Correct Answers For The 5 Wh-Words.

k value equal 38, for the preprocessed query 251 correct answers, equal to 84,22 % of the success and for the original query, 240 correct answer, equal to 80,54 % of the success, with an increments of success for the preprocessed query compared to original query, respectively of 4.38% and 2.64%.

I also, tested, the distribution of the correct answers in the cases, when, in the user question is present one of the 5 wh-words (*when, where, who, how, which*). With K equal to 38 and 100 questions submitted,the distribution of correct answers as showed in figure 4.14. Analysing the graph, we can observe, the distribution of correct answers respect the wh-words used on the question. The system produce better results for the wh-word *how* and *who*, respectively with 88,68 % and 87,50 % of success than the other wh-words.

These results indicate that the heuristic used for the preprocessing of the query is particularly suitable for the extraction of quantitative and persons data. A lower value in the case of *where* questions, is attributable at not indexing by wordnet of the names of places.

4.6.2 Telenoid Dialogue System vs QAnus System

The last series of tests was related to the comparison between TDS and QAnus in term of *performance* and *recall* parameters.

The first test involved the comparison between the performances. I tested several conceptual spaces, consisting respectively of 38,35,30,25 and 20 topics. Each selected space has been loaded inside QAnus and TDS and 10 random questions had be submitted to the both the TDS and QAnus systems. The response times of the two systems for each space have been calculated. Due to the architecture of QAnus, it was not possible to decompose the total response time in term of its main procedure phases. The comparison of the obtained results are shown in figure 4.15.

The response times in QAnus are always maintained in the order of 10 seconds

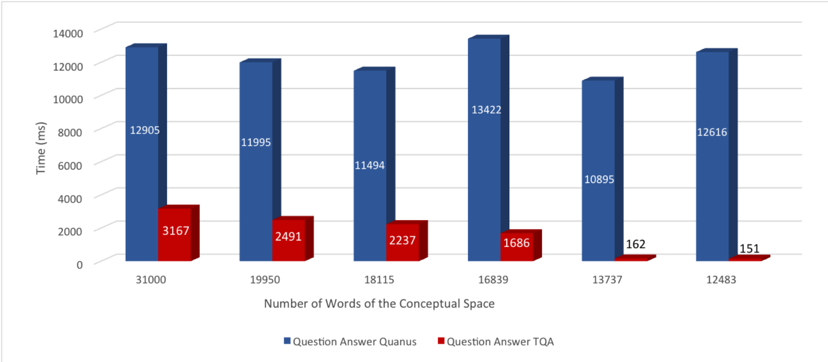

Figure 4.15: Telenoid Dialogue System vs QAnus System.

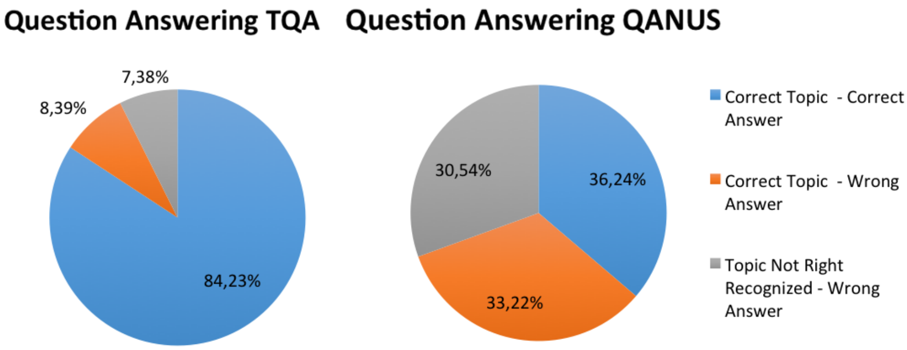

Figure 4.16: Comparison of Question Answering Results between Telenoid Dialogue System and QAnus.

and even in the worst case, when we used the biggest conceptual space (31000 words), the response times of our system were better with a time response reduction by around over 70%. In the best case, with a smallest conceptual space (12000 words), TDS presents a response time of just 151 ms against 12.5 seconds required by QAnus. With these times of our system, it is possible to realize a real-time interaction between Telenoid and the human user. Also with regard to the rate of recall the results obtained by our system are definitely positive.

In 36% of the cases QAnus produced the correct answer; in 33% of the cases the topic was correctly identified, but was not produced the correct answer to user question. Finally in the 31% of the cases QAnus was not able to identify the correct topic. Comparing these results with those of our system we can underline an increment of correct answers in TDS equal to 56.8% with 251 answers correctly identified, compared with the 108 identified by QAnus. Also the number of topics incorrectly identified was drastically reduced from 91 to just 22, with an improvement of 79.98%.

Chapter 5

Case Study 3: Robot and Autistic Children

Supported by the good results obtained by the study on the Acceptance and believability of the Telenoid robot, I was interested to investigate the possibility to have a humanoid robot that is both *avatar* and *partner* (Breazeal, 2003b) and to evaluate the utility and the functionality of a humanoid robot in the processes of Social Integration among Children with ASD and their classmates. The work described in this chapter is part of a multidisciplinary project resulted of the collaboration between the roboticslab of the University of Palermo and the local neuroscience public healthcare unit focused on Autistic syndrome. Thanks to their support, I was able to evaluate my research on subjects with social disadvantages as Autistic Children. Everyone might agree that in our modern society stigma and prejudice have a huge impact and affect on countless numbers of groups of people. The vision that common persons may have towards the marginalized people in our society is devastating, this feeling, is even more serious in small groups, as that of a school class. This is particularly true when considering a class where one or more children are affected by a mental illness such as autism. An autistic child, either at the lowest, or, at the highest edge of the spectrum, is not capable of offering cognitive and emotional feedbacks to peers and adults. In these conditions, ideas and thoughts on the child disease, on its consequences, on the relationship within the other children members of the class, that are continuously amplified and often cause themselves tension and distress. High levels of distress in these contexts both decrease quality of life often delaying the natural processes of interaction between autistic children and their classmates, preventing altogether sometimes the simple (verbal and non-verbal) interactions, imposing, in this way, significant barriers to communication. Behavior analysts have recently conceptualized such prejudices or stereotypes as a stimulus network, rejection being the final response to a mechanism of language known as derived relations (Hayes et al., 2002). So, with this

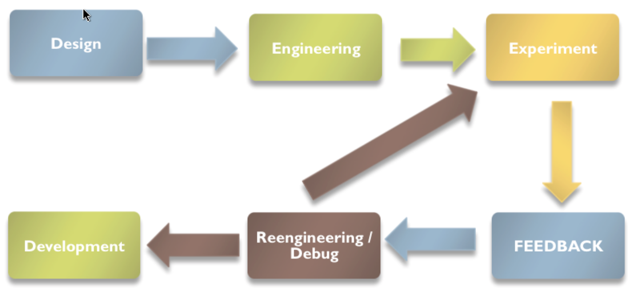

Figure 5.1: Disign-Try-Fix Method

new approach, the Telenoid robot guided by the proposed cognitive architecture overcome the limitations of existing cognitive architecture (Samsonovich, 2011) because it will be able to *simulate the real life computational equivalents of human mind* (Samsonovich, 2013). In the near future, is my opinion that can exist a therapist robot equipped with these cognitive capabilities and it is able to place side by side with a real human specialist during a therapy session.

5.0.3 Methods

The Methods used for analize the children's behavior and the possible improvements in the interaction with the classmates evolved during the laboratory session, is based on the collection of the informations through Observation, Participation and Notation of Social Phenomena during the session and the Recording of performance through audio-visual material by a specialized team of therapists and psychologists.

In the side of the engineering process, in order to understand, how the robot can help this people and how design the complex mechanism that support the robotics architecture for reach this kind of goals, it was used the method of Design - Try - and Fix. As is shown in the figure 5.1, following the indication coming by therapist and psychologists, the first architecture is designed, engineered and tried during the laboratory session with the children. After the feedbacks that coming from children, from the psychologists and by myself, there is a fix session of the eventual bugs and an adapting of the robotic system to new requirements; at the end there is the final development of the robotic framework as support of the project for the autistic children.

The psychologist Observations was annotate during the semi-structured work sessions subdivided in laboratory. The activities performed with the robot during each lab session have been designed and tailored on the real needs of the two autistic children involved in the project.

A first preliminary cognitive analysis of our autistic children was made by psychologists in the months preceding the beginning of the lab session. The main deficit encountered in the two boys involved in the pilot project are summarized as follow:

- Awkward or inadequate interactions: lead them to not know how to be good playmates or conversation partner because having activities and restricted and limited interests, compared to their peers, they can not share with them, remaining so, inevitably more isolated.
- Deficit of social reciprocity: They have difficulty to initiate interaction or to enter into relationship with others.
- Marked decrease of socio-relational and communication with others; a parallel internal retreat.
- Problems in the use of social skills to connect with other people because they may seem to be in their world and may be hard for them to share a common focus with another person about the same object or event-known as joint attention.

The age of the our two candidate is for both 11 years old.

5.0.4 Experimental setup

The experimental sessions was made directly in the classroom used by our autistic children with his classmates, so they do not need to adapt they-self to a new environments that do not know having side effects to the experiments.

The classroom previous of the beginning of the session test is equipped with two fixed video cameras, for recording all happen in the lab session. In the center of the room is placed the Nao or Telenoid robot in according with the specific lab and the task of the day, and a second control location is placed outside the main classroom (only when is used the telenoid robot), or in the corridor or in a second neighbour classroom.

The children that who take part in the experiments are free to move on the classroom, touch the robot and following (or not) all the indication that coming by the robot.

Other therapists or hidden observers can compile evaluation sheets during sessions, and these data, can also be added to a common repository and used for successive analysis.

Figure 5.2: Lab 1: Children use Telenoid Robot as a Proxy

5.0.5 Lab 1: Children use Telenoid Robot as a Proxy

The experimental setup of the laboratory 1, as shown in the figure 5.2, is designed to promote the interaction among children through the use of the telenoid robot as *communications mediator*. The basic idea is help the autistic child to initiate a first interaction with others classmates, thanks to the help a human-like robot. The final long term goal is, make sure, that the child continues to interact with classmates after the ending of session chat when the robot (the facilitator of the conversation) has been removed.

It is well documented that children with ASD, under some circumstances, have been shown to respond and interact better with robotic systems than with humans. This may be due to the robots'more simplified and predictable nature, which may allow them to be less intimidating and confusing interact through the robot than interact directly with humans. The robots as mediator socio-communicative, compared to humans, is able to establish a different way to communicate, more close to the social sphere of an autistic. Indeed, the Telenoid robot reduces the immediacy of the relationship, possibly making it more tolerable, such as happen in a skype chat.

For theoretical reasons, the perceptual awareness of sharing a common environment of a autistic child with the robot is assumed to be of momentous importance for the interesting subjects that ascribe intentions and actions to the robot for communicate with others.

In this laboratory, the goal is promote the interaction between three children (one is an autistic child); so the telenoid robot is placed in the classroom corner,

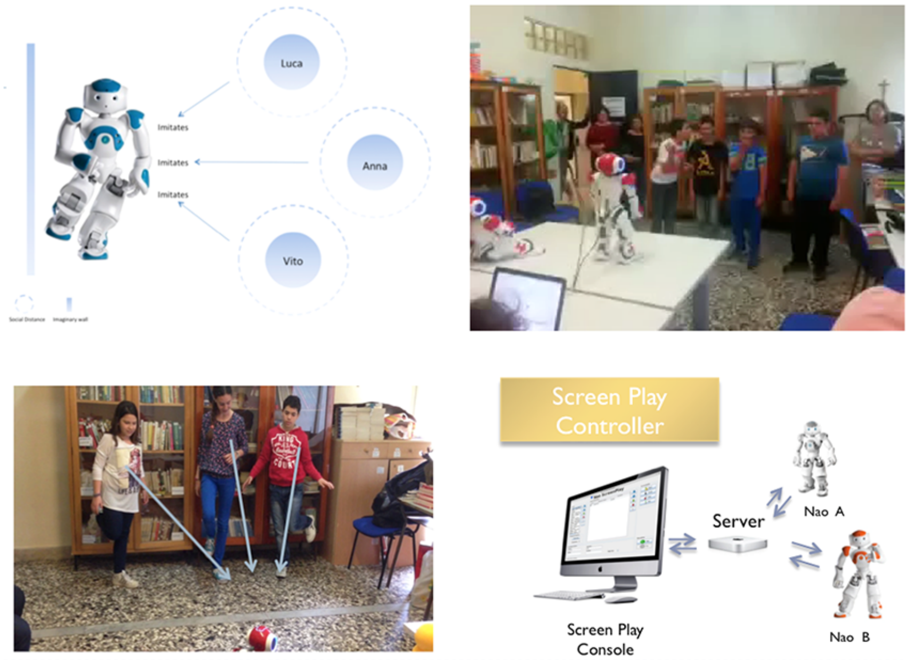

Figure 5.3: Lab 2: Children imitate the Robot

two children are sitting in front to him, the third is in a separate room in front to the robot control system for the teleoperation. The teleoperation system are so described: a webcam, a set of headphones, a face recognition system and a control panel, which are all connected to a laptop connected to the Internet. It permits to a person to control, easier, the robot just with the voice and the body movements. The webcam captures the operator's face and head movements, which will be synchronized with the movement of the android's face and head through the face recognition system. The child's voice is passed on to the receiver of robot through a microphone. The android also performs automatic behaviors, independent of the operator, such as breathing and eye blinking in order to make it seem more life-like during the conversation. With this kind of control system, the autistic child can talk with his classmate through the robot, using in this way the robot as mediator and avoiding the deep difficulties that would sustain during a direct communication. The architecture used for the framework is based on Client-Server paradigm.

5.0.6 Lab 2: Children Imitate the Robot

The robot, placed in the center of classroom and the children are seated on semi-circle in front of the him, as shown in figure 5.3, performs some pre-recorded

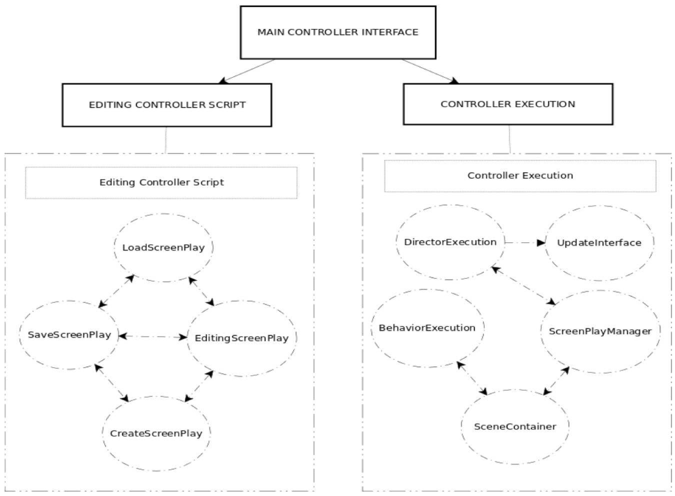

Figure 5.4: The Schema of the System:The Robotic Script Maker

behaviours. The goal of the children is emulate the movements of robot. During this activity, the attention is focused on the robot. The basic idea is, to have 3 children (one of this is our autistic child) that interact to each other stimulated by difficulty of imitate the robot. So each of them can ask help to the other for completing the task.

For supporting a better interaction between the group and for stimulating the communicative processes, I design and developed a framework for the easy construction of a screenplay based on the combinations of complex robotic behaviors. In this way is very simple and fast, make a new scene of a robotic choreography that the children can imitate.

The system, as shown in Figure 5.4, is composed of two macro-blocks: Editing Controller Script and the Controller for the block execution of the screenplay. The editing script block manages the scenes of the choreography (*script*) with a easy system for editing the complex behaviors. Each of them is made by a sequence of more simple behaviors activated in a time-line. The controller of Blocking Execution is responsible for the management and the execution of the scenes. It manages the behaviors, so that a new scene begins only when all the behaviors of the previous scene has been completed. It, also, manages all the connections between the robot and the clients and it implements the logic control of the robot.

The architecture used for the framework developed is to type Client-Server.

5.0.7 Lab 3: The Robot Imitates Children

The experimental setup of the laboratory 3, as shown in the figure 5.5, is designed to promote physical interaction among children through the use of social imitation

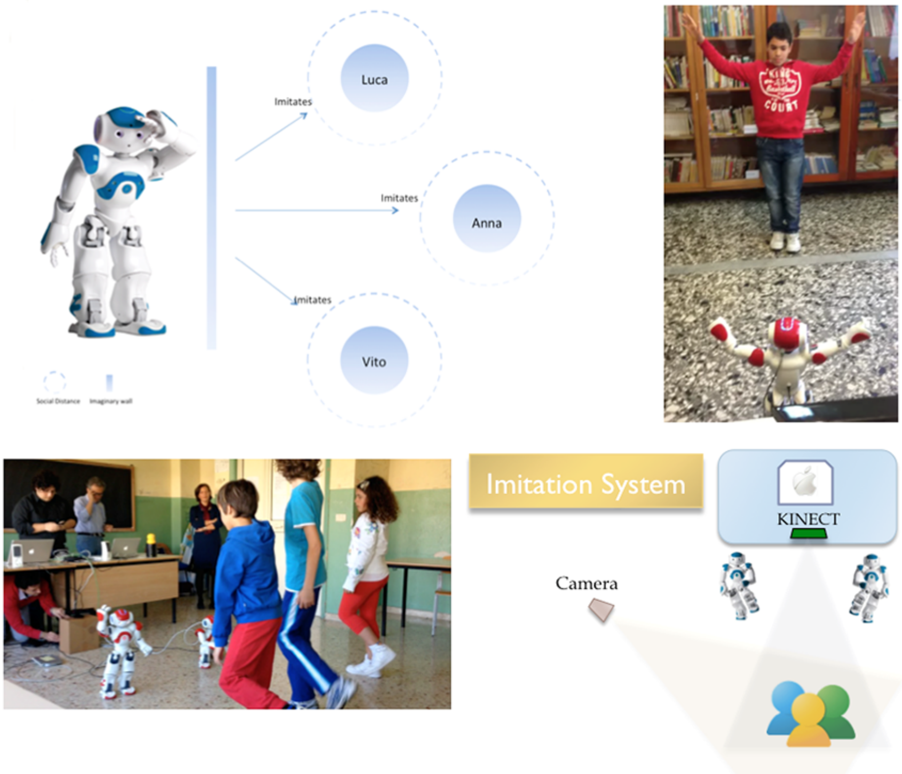

Figure 5.5: Lab 3: The Robot imitates Children

mechanism.

The basic idea is to have 3 children that communicate to each other with a not verbal communication. The child's goal is to create some possible movements that can be reproduced by the robot. If this kind of task can result easy for a normal child, is not the same for an autistic child. In this way, the lab task try to encourage the child to show some form of initiative. In this laboratory the nao robot is placed in the center of the classroom, the children are in front at the robot. A Xbox 360 kinect is placed behind the robot in a position hidden to children but able to catch their movements. In order to improve human-robot interactions specifically related to autism treatment, a system in which humans and robots can interact has been created, through the non verbal language of the mime. In particular, a robot imitation learning platform was developed using the Xbox 360 Kinect and the Aldebaran NAO humanoid robot. The robot imitated movements made by a human based on a real time full body tracking using a kinect sensor.

These developments will not only advance and improve interaction between robots and humans and provide a method of interfacing readily available tech-

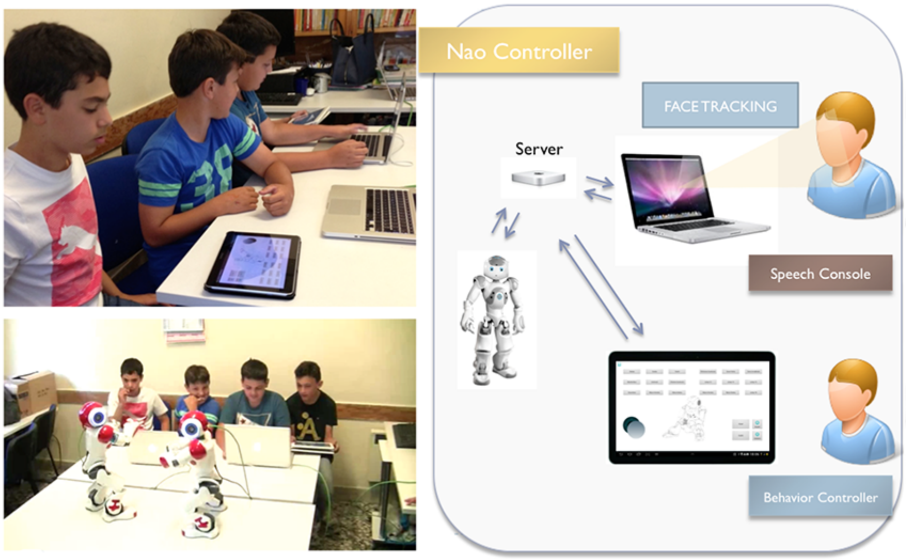

Figure 5.6: Lab 4: Children converse with each other using robots as an Avatar

nologies, but will also pave the way that allow the robot to better respond to the child's body language and as a supplement to responding to explicit motions.

Using the client-server network interface, the data are sent from the Kinect to a server, where a python module will read and interpret the data informations and translate them into commands for the robot's actuators. NAO will then execute the movements in real-time.

5.0.8 Lab 4: Children Speak Using The Robot as Avatar

In support of this laboratory, I have designed and developed an architecture for the Remote control, working using a Cloud network, of a humanoid robot through the aid of a mobile device.

The experimental setup of the laboratory 4, as shown in the figure 5.6, is designed to promote physical communication among children through the use of the interaction of two robot as avatar.

The basic idea is to have 4 children that communicate to each other through the use of two humanoid robots. In this laboratory the Children are divided in pairs, and each pair controls a robot.

The control system of each robot is designed to be used by two children. A child placed in front of a computer, through a system of facial tracking can control the movement of the robot's head and through a system of chatting on the keyboard can decide what to do to say to the robot. The other child through a motor control system of the robot, implemented inside a Google Android tablet, can control and choose the movements of the robot. The possible behaviors are: stand up, walk, indicating with his hands freely move through a joystick.

Having divided the robot control in two parts, the children are encouraged to work together to choose the best robot control strategies. For example,it happened that the first child controls the robot's head, getting okay, in according with the second child that make him walk.

The architecture which is at the basis of the face tracking system and the motor control system of the robot, is a client-server architecture. This allows me to make the system scalable, free in the choice of technologies for each client, and independent of the physical place where the robots can operate. A plausible scenario could be that the pairs of children can be found in two different places and control the robot through the Internet.

5.0.9 Observations and Results

Robot, as a social communicative mediator, is able to establish a new different way of communication with patients with DSA. The initial sessions is devoted to familiarization of the children with the robot, and also served as observation time of the spontaneous reactions of the child when interacts with the robot. During the familiarization phase, it is possible to identify verbal and non-verbal expressions of the child. Collecting a repertoire of preselected social situations can be allows the therapist to realize new situations as a consequence of the real time interaction between Robot and the child and the interactions between the children involved in the lab sessions. The first element to be evaluated during the session is the capacity of the child to interact with others. Factors such as the spontaneity, the time of the interaction (*counted in seconds*), or how many time the child begin a interaction with others, can be considered by the therapist an indicator of the "*goodness*" of the *integration* . At the end of the experimental period with the robot, the observation by the psychologists, has been very positive. They observed that for the autistic subject that have low degree of initiative using the Telenoid Robot has showed more interest to others. They oberseved, also, that Teleoperating the robot, the child, changed prosody and language, acquiring more initiatives and assuming the role of *Initiator* of the first interaction in the Group. In the specific, the autistic subject (Eros is his name) at the end of the experimental session acquired more spontaneity and more initiative on the approach in the conversation. The other autistic subject (Nicolo is his name) that was poor listener to others at the end of the experimental session acquired more self control resulting less prevaricating during the interaction and less repetitive. In any case, for both subjects social stimuli, the shared Attention with Motivation and Interest was increased. They had the opportunity to "stay together" and this had facilitate the creation of new social relationship. In conclusion it is emerged from the observations of the teacher and parents that the autistic children, using the robot as a "facilitator", have the new acquired relationship capability to other generalized contexts of their life.

Chapter 6

Case study 4: Robot and ALS Patient

Supported by the good results obtained by the study on the Acceptance and Believability of the Telenoid robot, I am interested to investigate the possibility to have a humanoid robot that can support the ALS Patient's Parents (or *caregivers*) especially when patients are left alone, most of the time, during the day.

Assistive social robots, a particular type of assistive robotics designed for social interaction with humans, could play an important role respect to the health of quadriplegic and anarthric (*locked-in*) ALS patients

Objectives Assistive social robots are believed to be useful in ALS patients for two reasons, a functional one (the robot can be seen as a continuation of the body of the patient that can fulfill all of its primary needs) and an affective one (the robot as the Telenoid (from our studies that is perceived to be) can be used as an avatar of his assistant 24/24, satisfying his needs as secondary to not feel alone or feel his loved ones nearby.).

Such robots are developed to function as an interface for the patients for communicate with the external world (*means parents, siblings or all their interlocutor*) with digital technology, and to help increase the quality of life of the patients by providing companionship, respectively.

The project aims to combine a Brain-computer interface (BCI) controlled humanoid robot into the conventional home management of quadriplegic and anarthric (locked-in) ALS patients and a communication system that trought a Brain-computer interface (BCI) and a humanoid robot as avatar, can help increase the quality of life of the locked-in patients.

In the study, reported in this chapter, a robot will move through the environment, will grasp and will carry objects and will show the places it visits via a wireless webcam whilst being autonomously controlled by the patient through a BCI. With this approach, the robot will restore the basic autonomous motion of the paralyzed patient enabling him to extend his presence beyond the restricted boundaries of his bed and can communicate with other people by speaking directly to the robot, in this way, perceiving them next to him (the patient isn't alone in the room).

I do expect to improve the quality of home care and to achieve outstanding effects on quality of life of patients and caregivers.

In this thesis is illustrated the architecture for a human-humanoid interaction based on EEG - Brain Computer Interface (EEG-BCI) for neurological patients affected by ALS locked-in syndrome.

I developed an architecture composed by four main areas. The first area analyses the user's brain signals. The second cognitive area elaborates the signals measuring the intention and attention and converts into behaviors for the robot. The third area manages the execution of the behavior of the robot. I have conducted experiments using a bio-signal amplifier and two humanoid robot: Nao and Telenoid.

In consideration of the undamaged cognitive status in most ALS patients and the increasing improvement of humanoid robot technology, I hypothesize that a dedicated BCI-robotic platform might give an unprecedented contribute to the management of these patients, offering independence, increasing the quality of life and psychosocial status of locked in patients and represents an exciting multidisciplinary challenge.

6.0.10 The proposed System

The entire schema of the proposed system, as shown in figure 6.1, consists of *visual user interface, brain signal acquisition, BCI architecture.*

The user through a user interface receives visual stimuli that are perceived differently by each user's brain. The user's interface is a visual 4x4 symbolic matrix where each element represents, robot behaviours, emotions felt, or a phrases that the patient want communicate.

A real-time EEG catcher, used as a brain signal acquisition module, synchronizes the provided stimuli and the registration of the user's mental signals. The BCI architecture creates a cognitive link between the user's mental signals with the humanoid robot assisting the user in achieving desired intentions through the robot behaviours.

The BCI architecture, as shown in figure 6.2, is divided in five main modules: signal processing, cognitive, communication, behavioural and Motion.

- The *Signal Processing Module* processes the user's brain raw signals acquired by EEG cap;

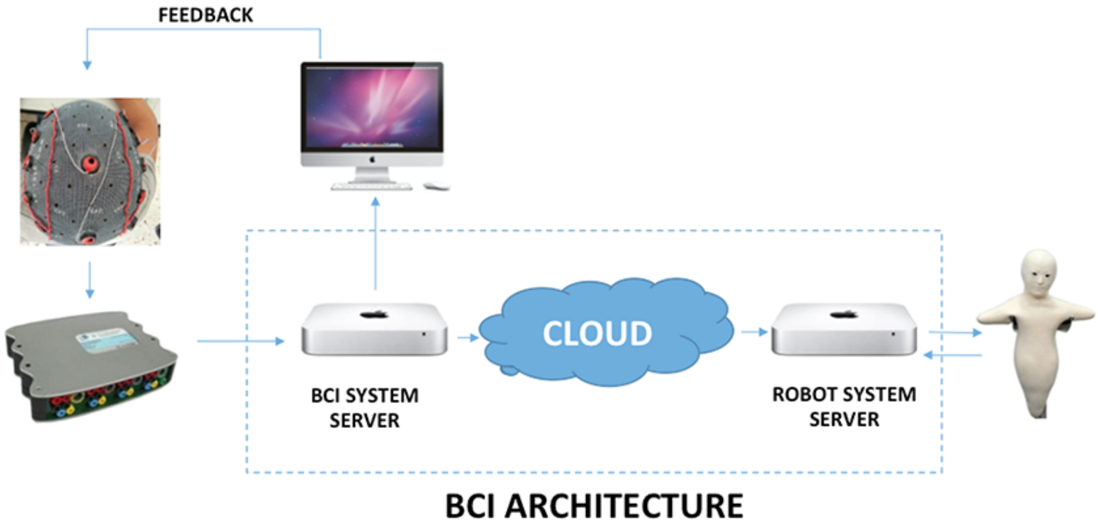

Figure 6.1: Overall Schema of the BCI System

- The *Cognitive Module* acquires the desired behavior related to the signals and modulates it through the measurement of the concentration (*Attention*) and willing (*Intention*) of the user.
- *The Communication Module* creates a communication channel between all components of the architecture in according to the different methods of connection of each module.
- The *Robot Behavioural Module* activates a non deterministic customized behaviour related to the Cognitive Module.
- The *Robot Motion Module* implements the humanoid behaviours executing them by controlling the joints of the robot. The actions of the behaviors are implemented as interpolation of different poses of the robot.

6.0.11 Experimental Setup

The experimental setup, as shown in the figure 6.3, is designed to promote physical interaction of ALS patients and their parents (or *caregivers*) through the use of the robot as mediator.

The experiments will be performed in a ALS-dedicated room directly to patient's house. In this space the complete EEG-computer-robot system will be installed allowing the patient practice with it directly from his bed. The basic idea is to have the telenoid robot next to the bed of the locked-in patient, a monitor in front of the patient, where he can see the bci interface for the communication

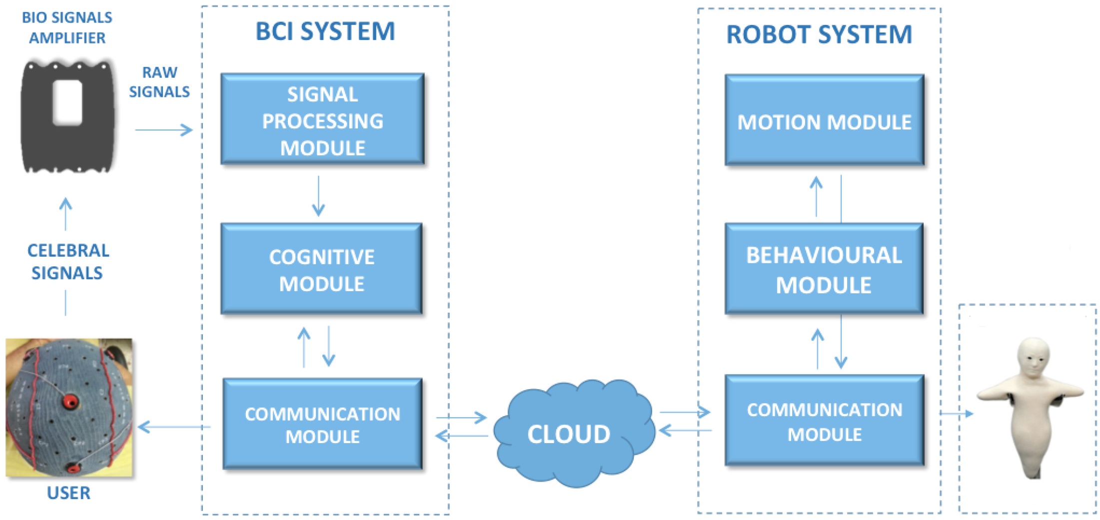

Figure 6.2: The Overview of the System Architecture

system and a big tv where he can see his interlocutor located in a remote place thanks to a video camera. In the remote place will be the caregiver that using the telenoid teleoperation system can speak with the patient. The components of the teleoperation system are a webcam, a set of headphones, a face recognition system and a control panel, which are all connected to a laptop hooked up to the Internet. The webcam captures the caregiver's face and head movements, which will be synchronized with the movement of the android's face and head through the face recognition system. The caregiver's voice is passed on to the receiver though a set of headphones. The android also performs automatic behaviors, independent of the operator, such as breathing and eye blinking in order to make it seem more life-like during the conversation. All words and phrases (the thinking) produced by the patient are processed by bci system and sent to a system of voice playback so that they can be heard by caregiver. Whit this kind of communication system, it is possible to give to a locked-in patient the opportunity to can speak again and, at the same time, he has the sensation to not remain alone.

The EEG device is connected through USB to the Data PC which is dedicated to the analysing of signals and the management of the user application. The Data PC is connected to an experimental monitor used by experimenters to control and supervise the system and a user monitor, used by the patient to run the provided task. The Data PC is connected also to the BCI server, used to acquire the signal coming from the EEG device, and to send them over the network to a listening Bci server, connected to the Telenoid robot which it will translate the provided command into an high level behaviour in the robot.

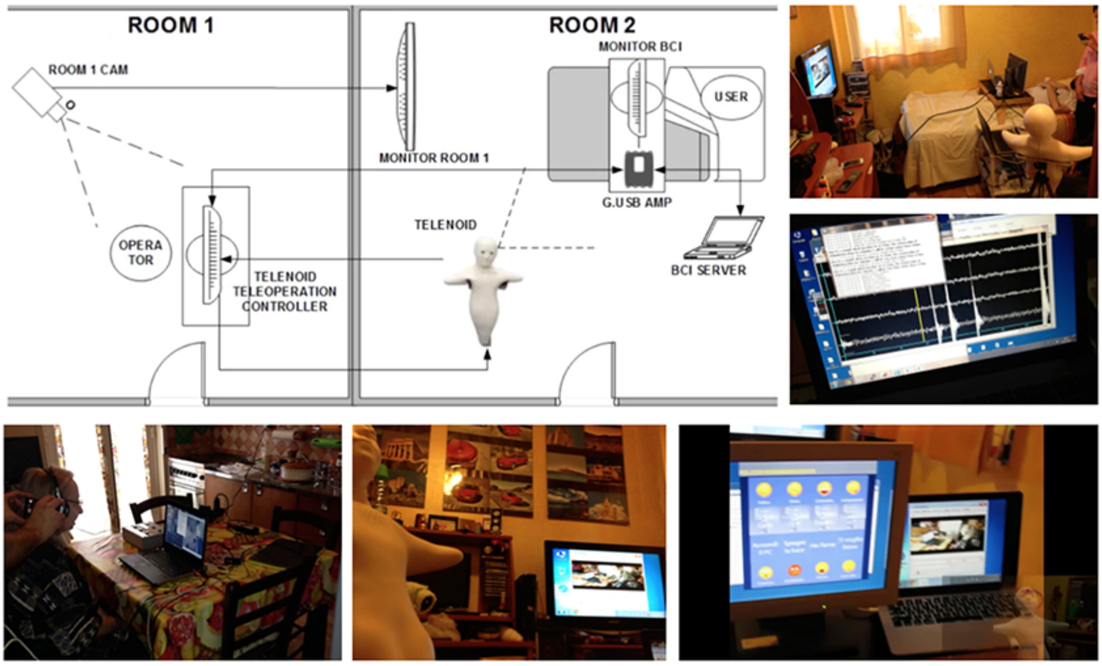

Figure 6.3: BCI System: Experiment session with ALS Patient

6.0.12 Preliminary Experimental Result

Two non-demented locked-in ALS patients, as shown in figure 6.5, with valid eye movements have been recruited. The communication capacity of each patient has been verified with ETRAN (*eye transfer*) panel or an eye-tracking computer device in order to obtain informed consent for the study and to verify the expected outcomes. Only one patient at a time has been submitted to the experimental sessions. Before starting the training session, each patient has been first submitted to specific tests to explore motivation and expected results and at the end of the session, perceived fatigue, subjective satisfaction, goal achievement and quality of life has been explored, through specific questionnaires. Data will be also correlated with clinical and demographic variables to identify factors likely affecting the propensity of the patient to successfully carry out the training session and the effectiveness of the utilization of the BCI- system. All patients have been introduced to the interaction setting structure with the humanoid robot. All interactions were videotaped.

In order to study the relations between a locked-in patient and a humanoid robot I focused on two important variables that may represent specific targets of BCI-robot technology: 1) Loneliness; 2) Assistance.

1. Social assistive robotics improve adherence to rehabilitation programs in

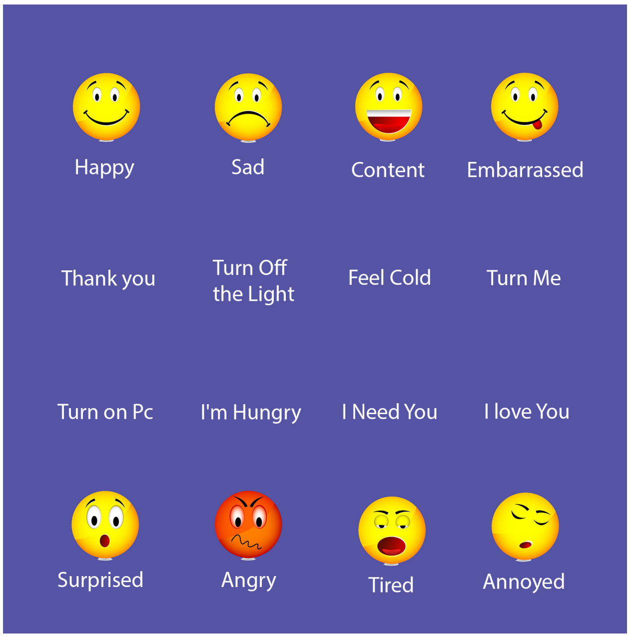

Figure 6.4: BCI Communication Interface

disabling diseases. From this perspective, we will assess the capacity of ad hoc fashioned humanoid-machine in providing companionship and encouragement to locked-in patients. The robot will be made to be able to receive the emotional states of the patient and to reply accordingly to them with a reinforcing feedback.

2. The clinical status and communication impairment of locked-in subjects require an all-day round dedicated assistance for any need. I evaluate the effects of the patient-controlled robot on the burden of the principal caregiver. Each variable will be assessed using specific yes/no questionnaires.

The BCI system was finally tested in a normal session of interaction with the patient. The patient has the possibility to use an interface for the communication, as shown in figure 6.4. In this interface are represented 8 sentences tailored on the need's patient and 8 emotional feeling useful for a easy and fast remote communication.

To verify the effect of a patient-controlled humanoid robot on motor substitution and independence in ALS patients. I tested the ability of patients affected by quadriplegia to operate a humanoid robot for grasping objects, movement and scene-shooting. Moreover, due the possibility for humanoid robots to express facial emotions and to imitate human behaviour, I evaluated the efficacy of these robot in giving coaching, encouragement and companionship to disable subjects.

I evaluated the effect of the BCI-controlled humanoid robot on the quality of life of patients and caregivers through the administration of simple satisfaction

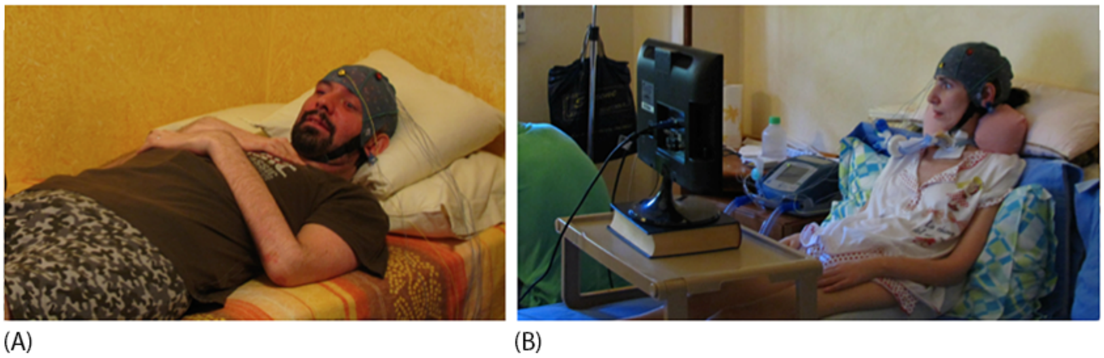

Figure 6.5: Patients recruited:
(A) 35 years old; ALS (3 years); Locked-in; Use eye-gaze for communication and entertainment.
(B) 44 years old; ALS (9 years); Locked-in; In invasive ventilation; 4 years Use the eye-gaze for internet, email and video reading.

questionnaire also aim to analyse the possible change in the patients assistance needs and to identify the variables (*social, cultural and demographic*) specifically related to a successful Human-Humanoid relationship. All data collected showed a marked improvement of the living conditions of the patient. Also have shown significant improvements on interim analysis of the impact of attention, motivation, training and learning in the individual patient.

My experimental results shows linking between cognition and emotions in their interaction between patients and caregiver through the robot used as a *proxy*.

In consideration of the undamaged cognitive status in most ALS patients and the increasing improvement of humanoid robot technology, I hypothesize that a dedicated BCI-robotic platform might give an unprecedented contribute to the management of these patients offering independence, increasing the quality of life and psychosocial status of locked-in patients and it represents an exciting multidisciplinary challenge.

At the present time no therapeutic interventions may modify the course of ALS towards the complete paralysis. The development of a BCI-robotic platform devoted to care represent a farsighted chance of taking advantage from high technology to promote the human health.

The use of a BCI-humanoid robot system has the potentiality to restore the communication lose for a completely locked-in patient and extend his presence at every place where the robot webcam can reach. Consequently, this system has the potentiality to restitute to the intact mind of patients the autonomy of the communication of their care and change completely their participation to the society.

Conclusions

The present thesis tittle "*Investigating Perceptual Features For A Natural Human Humanoid Robot Interaction Inside A Spontaneous Setting*" shows a descriptive analysis of the main perceptual and social and cognitive features of natural conditions of agent interaction, which can be specified by agent in human-humanoid robot interaction. These perceptual and cognitive features underlie the ordinary human-human interaction in daily life contexts and therefore can be exploited in the human-humanoid interaction because they could be seen as a sort of specialized implicit competences.

The sense of "*togetherness*" between persons and a Humanoid robot is "*inherently social*" (Biocca, 1992) and is highly connected with the concepts of particular behavior defined "*sensible*" (Chella et al., 2011a) because it is capable to express cognitive functionality.

The research was focused on the social dimensions that allow to obtain in general a successful interaction. Some of these dimensions were extracted to be embedded in a controlled interaction setup and then they were studied in real-life case studies. In particular, I was interested to investigate the possibility to use a social robot with subject with disadvantage, (mental and psychic disability), such as Elderly people, Autistic Children, Autistic young adult and ALS patients, and with normal people but that live every day with the problematic of autism as parents of autistic children.

I found that the sense of a shared environment is substantive for obtaining a satisfying interaction, where the distances between agents are construed as perceptual-motor proxies of the regions where intentions and actions are available at a glance. The characteristics of the present thesis aims at specifying the cognitive contribution of features that arise during the interaction to the perception of its consistent purposed driven behaviour.

One crucial result of this research is that subjects are significantly inclined to perceive the Telenoid as *cooperative* and *competent* even though this favorableness.

Since the users perceived the behaviour of the Telenoid coherent and consistent, this results suggest that they are favourably incline to accept the humanoid's behaviour as tuned to theirs and moved by the commitment to meet their demands.

So, I presented in this thesis a humanoid robotic system capable of exploiting a knowledge base through the use of latent semantic analysis technique in order to contextualize the dialogue that exists between humans and a humanoid robot. In this way, I have a Humanoid Robot that can exhibit a natural and empathic interaction with human beings, supporting a pseudorealistic conversation and expressing human-like body emotions.

Some findings emerge that might prove as much emphasizing the dimensions that make a human humanoid interaction socially acceptable and efficient in task driven contexts as also relevant for future research.

Future works will analyze whether items cluster within and across the constructs, whether they constitute a cognitive continuum or two specialized scaffolds for interaction, whether the features they represent only mutually reinforce each other and give a joint contribute to interaction.

In the near future, we will measure the level of each sub-scales of empathy (Davis, 1983) for demonstrating the efficacy of Telenoid as Empathic Embodied Communication Technology (EECT) and we will use these results in order to amplify the non verbal interaction of the robot through the use of tailored emotional body stimuli on a specific user.

In this way Telenoid, (Sorbello et al., 2014), could be used as a new medium able to perform empathic dialogue for understanding the needs and desires of such users and as a stimulus in order to evoke reflections, images and emotions. My research project, I hope, will be part of the research contribution to realize a society preserved through mutual assistance to take care of elderly and not fully self-sufficient peoples.

Bibliography

R. Adolphs. Neural systems for recognizing emotion. *Current opinion in neurobiology*, 12(2):169–177, 2002.

F. Agostaro, A. Augello, G. Pilato, G. Vassallo, and S. Gaglio. A conversational agent based on a conceptual interpretation of a data driven semantic space. *AI* IA 2005: Advances in Artificial Intelligence*, pages 381–392, 2005.

Gordon Willard Allport. Personality: A psychological interpretation. 1937.

SW Anderson, A. Bechara, H. Damasio, D. Tranel, and AR Damasio. Adolphs, r.(1999). social cognition and the human brain. trends in cognitive science, 3, 469-479. allen, c.(2004). *Behavioral and Brain Sciences*, 1991.

Salvatore M Anzalone, Yuichiro Yoshikawa, Hiroshi Ishiguro, Emanuele Menegatti, Enrico Pagello, and Rosario Sorbello. Towards partners profiling in human robot interaction contexts. In *Simulation, Modeling, and Programming for Autonomous Robots*, pages 4–15. Springer, 2012.

M. Argyle and M. Cook. Gaze and mutual gaze. 1976.

M. Argyle and J. Dean. Eye-contact, distance and affiliation. *Sociometry*, pages 289–304, 1965.

M. Argyle, R. Ingham, F. Alkema, and M. McCallin. The different functions of gaze. *Semiotica*, 7(1):19–32, 1973.

Magda B. Arnold. Emotion and personality. *Psychological aspects*, Vol. I, 1960.

G. Balistreri, S. Nishio, R. Sorbello, A. Chella, and H. Ishiguro. Natural human robot meta-communication through the integration of android's sensors with environment embedded sensors. *Frontiers in Artificial Intelligence and Applications*, IOS Press vol 233:26–37, 2011a.

S. Balistreri, G.and Nishio, R. Sorbello, and H. Ishiguro. Integrating built-in sensors of an android with sensors embedded in the environment for studying a more natural human-robot interaction. *Lecture Notes in Computer Science,Springer Berlin / Heidelberg*, 6934:432–437, 2011b.

Jacob Benesty, M Mohan Sondhi, and Yiteng Huang. *Springer handbook of speech processing*. Springer, 2008.

F. Biocca. Communication within virtual reality: Creating a space for research. *Journal of Communication*, 42(4):5–22, 1992.

C. Breazeal. Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1):119–155, 2003a.

Cynthia Breazeal. Toward sociable robots. *Robotics and Autonomous Systems*, 42 (3):167–175, 2003b.

Cynthia L Breazeal. *Designing sociable robots*. MIT press, 2004.

Laurie Carr, Marco Iacoboni, Marie-Charlotte Dubeau, John C Mazziotta, and Gian Luigi Lenzi. Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the national Academy of Sciences*, 100(9):5497–5502, 2003.

A. Chella, M. Frixione, and S. Gaglio. An architecture for autonomous agents exploiting conceptual representations. *Robotics and Autonomous Systems*, 25 (3):231–240, 1998.

A. Chella, G. Pilato, R. Sorbello, G. Vassallo, F. Cinquegrani, and S. Anzalone. An emphatic humanoid robot with emotional latent semantic behavior. *Simulation, Modeling, and Programming for Autonomous Robots*, pages 234–245, 2008.

A. Chella, C. Lebiere, D. Noelle, and A. Samsonovich. On a roadmap to biologically inspired cognitive agents. *Frontiers in Artificial Intelligence and Applications*, IOS Press vol 233:453–460, 2011a.

A. Chella, R. Sorbello, G. Pilato, G. Balistreri, S. Anzalone, and M. Giardina. An innovative mobile phone based system for humanoid robot expressing emotions and personality. In *BICA*, pages 57–63, 2011b.

A. Chella, R. Sorbello, G. Pilato, G. Vassallo, G. Balistreri, and M. Giardina. An architecture with a mobile phone interface for the interaction of a human with a humanoid robot expressing emotions and personality. *AI* IA 2011: Artificial Intelligence Around Man and Beyond*, pages 117–126, 2011c.

A. Chella, R. Sorbello, G. Pilato, G. Vassallo, and M. Giardina. A new humanoid architecture for social interaction between human and a robot expressing human-like emotions using an android mobile device as interface. In *Biologically Inspired Cognitive Architectures 2012*, pages 95–103. Springer, 2013.

R. Chrisley. Synthetic phenomenology. *International Journal of Machine Consciousness*, 1(1):53–70, 2009.

Sakhia Darjaa, Milos Cernak, Marián Trnka, Milan Rusko, and Róbert Sabo. Effective triphone mapping for acoustic modeling in speech recognition. In *INTERSPEECH*, pages 1717–1720, 2011.

K. Dautenhahn. The art of designing socially intelligent agents: Science, fiction, and the human in the loop. *Applied Artificial Intelligence, 12*, 7(8):573–617, 1998.

Mark H Davis. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology*, 44(1):113, 1983.

N. Durlach and M. Slater. Presence in shared virtual environments and virtual togetherness. *Presence: Teleoperators & Virtual Environments*, 9(2):214–217, 2000.

J.M. Fellous. From human emotions to robot emotions. *Architectures for Modeling Emotion: Cross-Disciplinary Foundations, American Association for Artificial Intelligence*, pages 39–46, 2004.

E. Goffman. *Behavior in public places: Notes on the social organization of gatherings.* Press of Glencoe, 1966.

Steven C Hayes, Reville Niccolls, Akihiko Masuda, and Alyssa K Rye. Prejudice, terrorism, and behavior therapy. *Cognitive and Behavioral Practice*, 9(4):296–301, 2002.

F. Hegel, T. Spexard, B. Wrede, G. Horstmann, and T. Vogt. Playing a different imitation game: Interaction with an empathic android robot. In *Humanoid Robots, 2006 6th IEEE-RAS International Conference on*, pages 56–61. IEEE, 2006.

F. Heider. *The psychology of interpersonal relations.* Lawrence Erlbaum, 1982.

E. Hudlicka. Reasons for emotions. *Integrated models of cognition systems*, 1:263, 2007.

Luca Iocchi, Emanuele Menegatti, Andrea Bonarini, Matteo Matteucci, Enrico Pagello, Luigia Carlucci Aiello, Daniele Nardi, Fulvio Mastrogiovanni, Antonio Sgorbissa, Renato Zaccaria, et al. Development of intelligent service robots. *Intelligenza Artificiale*, 7(2):139–152, 2013.

H. Ishiguro. Android science–toward a new cross-interdisciplinary framework. *Robotics Research*, 28:118–127, 2007.

H. Ishiguro, S. Nishio, A. Chella, R. Sorbello, G. Balistreri, M. Giardina, and C. Calì. Perceptual social dimensions of human-humanoid robot interaction. *12th International Conference on Intelligent Autonomous Systems*, on press in Advances in Intelligent and Soft Computing ISSN: 1867-5662, 2012.

H. Ishiguro, S. Nishio, A. Chella, R. Sorbello, G. Balistreri, M. Giardina, and C. Calí. Investigating perceptual features for a natural human-humanoid robot interaction inside a spontaneous setting. In *Biologically Inspired Cognitive Architectures 2012*, pages 167–174. Springer, 2013.

Y. Iwamura, M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita. Do elderly people prefer a conversational humanoid as a shopping assistant partner in supermarkets? In *Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on*, pages 449–457, March 2011.

Peter Jackson and Isabelle Moulinier. *Natural language processing for online applications: Text retrieval, extraction and categorization*, volume 5. John Benjamins Publishing, 2007.

P.L. Jackson and J. Decety. Motor cognition: a new paradigm to study self-other interactions. *Current Opinion in Neurobiology*, 2004.

T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro. Interactive robots as social partners and peer tutors for children: A field trial. *Human-Computer Interaction*, 19(1):61–84, 2004.

T. Kanda, T. Miyashita, T. Osada, Y. Haikawa, and H. Ishiguro. Analysis of humanoid appearances in human–robot interaction. *Robotics, IEEE Transactions on*, 24(3):725–735, 2008.

Shotaro Kobayashi, Susumu Tamagawa, Takeshi Morita, and Takahira Yamaguchi. Intelligent humanoid robot with japanese wikipedia ontology and robot action ontology. In *Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on*, pages 417–424. IEEE, 2011.

K. Koffka. *Principles of Gestalt psychology.* New York, 1955.

W. Köhler. *Gestalt psychology: An introduction to new concepts in modern psychology.* Liveright Publishing Corporation, 1992.

Hideki Kozima, Cocoro Nakagawa, and Hiroyuki Yano. Can a robot empathize with people? *Artificial Life and Robotics*, 8(1):83–88, 2004.

Florian Kraft, Kevin Kilgour, Rainer Saam, S Stuker, M Wolfel, Tamim Asfour, and Alex Waibel. Towards social integration of humanoid robots by conversational concept learning. In *Humanoid Robots (Humanoids), 2010 10th IEEE-RAS International Conference on*, pages 352–357. IEEE, 2010.

Akinobu Lee and Tatsuya Kawahara. Recent development of open-source speech recognition engine julius. In *Proceedings: APSIPA ASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference*, pages 131–137. Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference, International Organizing Committee, 2009.

K. Lewin, F.T. Heider, and G.M. Heider. Principles of topological psychology. 1936.

Christopher D Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J Bethard, and David McClosky. The stanford corenlp natural language processing toolkit.

E. Menegatti, G. Silvestri, E. Pagello, N. Greggio, A. Cisternino, F. Mazzanti, R. Sorbello, and A. Chella. 3d models of humanoid soccer robot in usarsim and robotics studio simulators. *International Journal of Humanoid Robotics*, Vol. 5 (03):523 – 543, 2008.

G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. Von Hofsten, K. Rosander, M. Lopes, J. Santos-Victor, et al. The icub humanoid robot: An open-systems platform for research in cognitive development. *Neural Networks*, 23(8-9), 2010.

T. Minato, M. Shimada, H. Ishiguro, and S. Itakura. Development of an android robot for studying human-robot interaction. *Innovations in applied artificial intelligence*, pages 424–434, 2004.

Catrin Misselhorn. Empathy with inanimate objects and the uncanny valley. *Minds and Machines*, 19(3):345–359, 2009.

J. Monceaux, J. Becker, C. Boudier, and A. Mazel. Demonstration: first steps in emotional expression of the humanoid robot nao. In *ACM Proceedings of the 2009 international conference on Multimodal interfaces*, pages 235–236, 2009.

B. Mutlu, J. Forlizzi, and J. Hodgins. A storytelling robot: Modeling and evaluation of human-like gaze behavior. In *Humanoid Robots, 2006 6th IEEE-RAS International Conference on*, pages 518–523. IEEE, 2006.

Bilge Mutlu, Fumitaka Yamaoka, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. Nonverbal leakage in robots: communication of intentions through seemingly unintentional behavior. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, pages 69–76. ACM, 2009.

Jun-Ping Ng and Min-Yen Kan. Qanus: An open-source question-answering platform, 2012.

S. NISHIO and H. ISHIGURO. Attitude change induced by different appearances of interaction agents. *Int. J. Mach. Conscious.*, 3(01):115–126, 2011.

E. Oztop, D.W. Franklin, T. Chaminade, and G. Cheng. Human-humanoid interaction: is a humanoid robot perceived as a human? *International Journal of Humanoid Robotics*, 2(4):537, 2005.

André Pereira, Iolanda Leite, Samuel Mascarenhas, Carlos Martinho, and Ana Paiva. Using empathy to improve human-robot relationships. In *Human-Robot Personal Relationships*, pages 130–138. Springer, 2011.

K.L. Phan, T. Wager, S.F. Taylor, and I. Liberzon. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in pet and fmri. *Neuroimage*, 16(2):331–348, 2002.

Mary L Phillips, Wayne C Drevets, Scott L Rauch, and Richard Lane. Neurobiology of emotion perception i: The neural basis of normal emotion perception. *Biological psychiatry*, 54(5):504–514, 2003.

D. Pietroni, G.A. Van Kleef, C.K.W. De Dreu, and S. Pagliaro. Emotions as strategic information: Effects of other's emotional expressions on fixed-pie perception, demands, and integrative behavior in negotiation. *Journal of Experimental Social Psychology*, 2008.

M. Poel, D. Heylen, A. Nijholt, M. Meulemans, and A. Van Breemen. Gaze behaviour, believability, likability and the icat. *AI & Society*, 24(1):61–73, 2009.

Aaron Powers, Adam DI Kramer, Shirlene Lim, Jean Kuo, Sau-lai Lee, and Sara Kiesler. Eliciting information from people with a gendered humanoid robot. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pages 158–163. IEEE, 2005.

A.B. Prevedini, G. Presti, E. Rabitti, G. Miselli, and P. Moderato. Acceptance and commitment therapy (act): the foundation of the therapeutic model and an overview of its contribution to the treatment of patients with chronic physical diseases. *G Ital Med Lav Erg*, 33:1, 2011.

Alexei V Samsonovich. Comparative analysis of implemented cognitive architectures. In *BICA*, pages 469–479, 2011.

Alexei V Samsonovich. Extending cognitive architectures. In *Biologically Inspired Cognitive Architectures 2012*, pages 41–49. Springer, 2013.

Byoung-Kyun Shim, Yoo-Ki Cho, Jong-Baem Won, and Sung-Hyun Han. A study on real-time control of mobile robot with based on voice command. In *Control, Automation and Systems (ICCAS), 2011 11th International Conference on*, pages 1102–1103. IEEE, 2011.

M. Shimada, T. Minato, S. Itakura, and H. Ishiguro. Evaluation of android using unconscious recognition. In *Humanoid Robots, 2006 6th IEEE-RAS International Conference on*, pages 157–162. IEEE, 2006.

Michihiro Shimada and Takayuki Kanda. What is the appropriate speech rate for a communication robot? *Interaction Studies*, 13(3):408–435, 2012.

M. Slater, A. Sadagic, M. Usoh, and R. Schroeder. Small-group behavior in a virtual and real environment: A comparative study. *Presence: Teleoperators & Virtual Environments*, 9(1):37–51, 2000.

Rosario Sorbello, Antonio Chella, Carmelo Calí, Marcello Giardina, Shuichi Nishio, and Hiroshi Ishiguro. Telenoid android robot as an embodied perceptual social regulation medium engaging natural human–humanoid interaction. *Robotics and Autonomous Systems*, 2014.

Ilona Straub, Shuichi Nishio, and Hiroshi Ishiguro. From an object to a subject-transitions of an android robot into a social being. In *RO-MAN, 2012 IEEE*, pages 821–826. IEEE, 2012.

Adriana Tapus and Maja J Mataric. Emulating empathy in socially assistive robotics. In *AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics*, pages 93–96, 2007.

Paul Thagard and Cameron Shelley. Emotional analogies and analogical inference. *The analogical mind: Perspectives from cognitive science*, pages 335–362, 2001.

K.R. Thórisson. Gandalf: an embodied humanoid capable of real-time multimodal dialogue with people. In *First ACM International Conference on Autonomous Agents*, pages 5–8, 1997.

Ioannis Toptsis, Shuyin Li, Britta Wrede, and Gernot A Fink. A multi-modal dialog system for a mobile robot. In *INTERSPEECH*, 2004.

O. Torres, J. Cassell, and S. Prevost. Modeling gaze behavior as a function of discourse structure. In *First International Workshop on Human-Computer Conversation*. Citeseer, 1997.

R. Vertegaal, R. Slagter, G. van der Veer, and A. Nijholt. Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 301–308. ACM, 2001.

L. Xie, Z.L. Wang, W. Wang, and G.C. Yu. Emotional gait generation for a humanoid robot. *International Journal of Automation and Computing*, 7(1): 64–69, 2010.

Ryuji Yamazaki, Shuichi Nishio, Kohei Ogawa, Hiroshi Ishiguro, Kohei Matsumura, Kensuke Koda, and Tsutomu Fujinami. How does telenoid affect the communication between children in classroom setting? In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 351–366. ACM, 2012.

S. Zhao. Toward a taxonomy of copresence. *Presence: Teleoperators & Virtual Environments*, 12(5):445–455, 2003.